



HAL
open science

Analyse statistique de données issues de batteries en usage réel sur des véhicules électriques, pour la compréhension, l'estimation et la gestion des phénomènes de vieillissement

Anthony Barré

► To cite this version:

Anthony Barré. Analyse statistique de données issues de batteries en usage réel sur des véhicules électriques, pour la compréhension, l'estimation et la gestion des phénomènes de vieillissement. Autre. Université de Grenoble, 2014. Français. NNT : 2014GRENT064 . tel-01305938

HAL Id: tel-01305938

<https://theses.hal.science/tel-01305938>

Submitted on 22 Apr 2016

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Signal, Image, Parole, Télécoms**

Arrêté ministériel : 7 août 2006

Présentée par

Anthony BARRÉ

Thèse dirigée par **Delphine RIU**
et encadrée par **Frédéric SUARD** et **Mathias GÉRARD**

préparée au sein du **Laboratoire d'Analyse des Données et d'Intelligence des Systèmes (CEA LIST)**
et du **Laboratoire de Modélisation et de Monitoring (CEA LITEN)**
dans l'**École Doctorale EEATS**

**Analyse statistique de données
issues de batteries en usage réel
sur des véhicules électriques,**
pour la compréhension, l'estimation et la
gestion des phénomènes de vieillissement

Thèse soutenue publiquement le **17 Octobre 2014**,
devant le jury composé de :

Latifa OUKHELLOU

HDR, Directrice de recherche, IFSTTAR,

Rapporteur

Pascal VENET

Professeur, Université Claude Bernard, Lyon,

Rapporteur

Marie-Cécile PÉRA

Professeur, Université de Franche-Comté, Belfort,

Présidente du jury

Rody EL CHAMMAS

PhD Project leader, Toyota Motor Europe,

Examineur

Jean-Michel LOUBES

Professeur, IMT, Toulouse,

Examineur

Delphine RIU

HDR, Maître de Conférence, G2Elab, Grenoble,

Directrice de thèse

Frédéric SUARD

PhD Ingénieur Chercheur, CEA LIST, Saclay,

Encadrant de thèse

Mathias GÉRARD

PhD Ingénieur Chercheur, CEA LITEN, Grenoble,

Encadrant de thèse



Remerciements

Je tiens en premier lieu à remercier chacun de mes encadrants, Frédéric Suard et Mathias Gérard, pour avoir proposé ce sujet de recherche passionnant, pour leur complémentarité et leur confiance commune. Ces trois années auprès de vous m'ont permis de me développer et d'évoluer dans de parfaites conditions.

Je souhaitais remercier Anthony Larue pour m'avoir accordé sa confiance à l'issue des entretiens préliminaires. Je remercie ma directrice de thèse, Delphine Riu pour son suivi régulier de l'évolution de mes travaux.

Je remercie sincèrement Latifa Oukhellou et Pascal Venet d'avoir accepté de rapporter cette thèse, pour le temps consacré à la lecture attentive de mes travaux et pour les remarques pertinentes en ayant découlé.

Je tiens ensuite à remercier Marie-Cécile Péra, Jean-Michel Loubes et Rody El Chammas pour avoir accepté de composer le jury et de participer à la soutenance et pour nos discussions constructives. Merci à vous, et aux personnes lisant ce manuscrit, pour l'intérêt porté à cette thèse.

Il est important pour moi de remercier ici les personnes m'ayant permis de me diriger sur ce parcours. Merci donc aux professeurs ou autres professionnels rencontrés à Angers et à Toulouse qui ont été à la base de ces travaux.

L'occasion m'est donnée de remercier tous les membres des équipes LIST et LITEN que j'ai rencontré lors de ces trois années. Merci aux personnes ayant de près ou de loin participé à mes travaux.

Je voudrais remercier de manière plus personnelle Frédéric et Mathias pour avoir fait en sorte de toujours travailler dans la bonne humeur. Merci aussi pour les plaisirs de la vie quotidienne à Yannick, Zoulficar, Aurélie, Lucile, Crédo, Khrykhry, Philippe, Julie, Sawsan, Étienne... pour le partage réussi de nos bureaux successifs, ou pour nos rigolades indispensables.

Je souhaite également remercier mes amis et ma famille ayant fait le déplacement pour me faire l'honneur d'assister à ma soutenance. Donc un grand merci à Yannick, Julie, Aurélie, Ramon, Maha et bien sûr à mes parents. Sur cette note plus personnelle je remercie mes parents pour leurs encouragements et pour leur confiance sans faille au fil de ces années, merci pour tout ! Enfin, merci à Maha pour son aide pratique et morale sans lesquelles ses travaux, entre autres, n'auraient pu être possibles.

Table des matières

Acronymes	7
Liste des symboles	9
Introduction générale	11
1 Contexte et problématique	15
1.1 Les véhicules électriques - Transports	16
1.1.1 Essor des véhicules électriques	16
1.1.2 Limitations des véhicules électriques	19
1.2 Les batteries lithium-ion	22
1.2.1 Présentation et fonctionnement	23
1.2.2 Données caractéristiques d'une batterie	25
1.2.3 Phénomènes de vieillissement	28
1.2.4 Facteurs globaux de vieillissement	30
1.3 Estimation du vieillissement	34
1.3.1 Indicateurs de vieillissement	34
1.3.2 Modélisation des phénomènes de vieillissement	35
1.3.3 Bilan de l'estimation des phénomènes de vieillissement	42
1.4 Conclusion des problématiques liées à l'estimation du vieillissement d'une batterie	44
1.4.1 Point de vue utilisateur sur le vieillissement	44
1.4.2 Bilan des solutions existantes	46
1.4.3 Orientation des travaux	46
2 Analyse hors ligne des données de batterie	49
2.1 Objectifs et problématiques	50
2.2 Présentation des données	50
2.2.1 Expérimentation ForeWheel	50
2.2.2 Expérimentation AX - SAXO	53
2.2.3 Synthèse des expérimentations	55
2.2.4 Données descriptives	57
2.3 Analyse des causes de vieillissement	59
2.3.1 Méthodologie d'analyse des données	60
2.3.2 Application aux données ForeWheel	66
2.3.3 Conclusion de l'analyse des données	71
2.4 Vers un estimateur du niveau de dégradation	71

2.4.1	Modèle statistique	71
2.4.2	Résultats et interprétations	73
2.5	Déduction du temps de vie restant d'une batterie	77
2.5.1	Méthodologie	77
2.5.2	Application et résultats	78
2.6	Conclusions - Discussions	79
3	Analyse en ligne des données d'une batterie	83
3.1	Problématique de l'analyse en ligne	84
3.2	Étude de signaux issus de mesures embarquées	84
3.2.1	Analyse des signaux	84
3.2.2	Extraction de motifs	87
3.2.3	Interprétation visuelle	90
3.3	Approche de classification de l'état de santé d'une batterie	92
3.3.1	Choix d'une méthodologie de comparaison des signaux adaptée	93
3.3.2	Méthodes de classification	99
3.3.3	Résultats des méthodes de classification	102
3.4	Vers une estimation continue de l'état de santé d'une batterie	107
3.4.1	Méthodologie	107
3.4.2	Résultats	109
3.5	Déduction du temps de vie restant d'une batterie	112
3.5.1	Méthodologie	112
3.5.2	Application et résultats	114
3.6	Conclusion - Discussion	117
4	Vers le suivi d'une flotte de véhicules électriques	119
4.1	Objectifs et problématique	120
4.2	Classification des usages	121
4.2.1	Méthodologie	121
4.2.2	Application de la classification	127
4.3	Aide aux préconisations d'usages	130
4.3.1	Méthodologie	130
4.3.2	Application de cette méthodologie	133
4.4	Mise en œuvre de modèles d'estimation sur une flotte	138
4.4.1	Modèles	139
4.4.2	Résultats sur une flotte	140
4.5	Gestion de la santé des batteries d'une flotte	145
4.5.1	Méthodologie de prise en compte de check-up	146
4.5.2	Stratégies d'optimisation	148
4.5.3	Exemple d'application	151
4.5.4	Perspectives	154
4.6	Conclusion - Discussion	155
	Conclusion générale	157
	Bibliographie	165
	Valorisations liées aux travaux	177

Annexes	180
A Cellules A123 Systems	181
B Cellules LiFeBatt	183
C Relevance Vector Machines	185
D Filtered Shape Averaging	187
D.1 Introduction	187
D.2 Contexte	187
D.2.1 Distance euclidienne	188
D.2.2 Dynamic Time Warping	188
D.2.3 Méthodes de moyenne	189
D.2.4 Motivation	190
D.3 Méthode proposée	190
D.3.1 Moyenne basée sur la Dynamic Time Warping	190
D.3.2 Fonction de lissage par noyau Gaussien	192
D.3.3 Complexité	192
D.4 Expérimentations sur la base de données UCR	193

Acronymes

<i>pdf</i>	Probability Density Function
ACP	Analyse en Composantes Principales
ANN	Artificial Neural Networks
ARMA	Auto-Regressive Moving Average
BEV	Battery Electric Vehicle
BMS	Battery Management System
CAH	Classification Ascendante Hiérarchique
CEA	Commissariat à l'Énergie Atomique et aux Énergies Alternatives
CU	Check-Up
DFT	Density Functional Theory
DOD	Depth Of Discharge
DTW	Dynamic Time Warping
EC	Ethylene Carbonate
EEL	Estimation En Ligne
EHL	Estimation Hors Ligne
EIS	Electrochemical Impedance Spectroscopy
EKF	Extended Kalman Filter
EOL	End Of Life
FSA	Filtered Shape Averaging
GA	Global Alignment
GDTW	Gaussian Dynamic Time Warping
GPS	Global Positioning System
IEA	Electric Vehicles Initiative
KMC	Kinetic Monte Carlo
kNN	k-Nearest Neighbors

LARS	Least Angle Regression
LASSO	Last Absolute Shrinkage and Selection Operator
LFP	Lithium Fer Phosphate
Li-ion	lithium-ion
LMO	Lithium Manganèse Oxyde
NiMH	nickel-metal
NN	Neural Networks
PF	Filtre Particulaire
PKE	Positive Kinetic Energy
RBPF	Rao-Blackwellized Particle Filter
RUL	Remaining Useful Life
RVM	Relevance Vector Machine
SEI	Solid Electrolyte Interphase
SOC	State Of Charge
SOF	State Of Function
SOH	State Of Health
SROT	Silver's Rule Of Thumb
SVM	Support Vector Machine
TGA	Triangular Global Alignment
VE	Véhicule Électrique

Liste des symboles

B^{train}	Base d'apprentissage $B^{train} = \{(x_1, t_1), \dots, (x_n, t_n)\}$ composée de n motifs x associés à leurs cibles t
B^{test}	Base de test $B^{test} = \{(x_1, t_1), \dots, (x_m, t_m)\}$ composée de m motifs x associés à leurs cibles t
Bat_1	Batterie n°1
Bat_2	Batterie n°2
Bat_3	Batterie n°3
B	Nombre d'échantillons générés par bootstrap
C_n	Capacité nominale d'une batterie
C_j	Cluster j
$\text{corr}(X, Y)$	Coefficient de corrélation entre deux variables X et Y
C-rate	Régime de charge ou décharge
d_{DTW}	Distance Dynamic Time Warping
d_{GA}	Distance Global Alignment
$d_{DTW_{SC}}$	Distance DTW adaptée avec les contraintes de Sakoe & Chiba
d_{ED}	Distance euclidienne
$D(V_1, V_2)$	Distance entre les usages de deux véhicules V_1 et V_2
ΔSOC	Variation d'état de charge
$H(X)$	Entropie de Shannon d'une variable X
f_g	Fenêtre glissante
Fl	Flotte de véhicule
$I(X; Y)$	Information mutuelle entre deux variables X et Y
IC	Intervalle de confiance

K_{GDTW}	Noyau associé à la distance DTW
K_{Gauss}	Noyau Gaussien
K_{GA}	Noyau associé à la distance Global Alignment
K_{SC}	Noyau associé à la distance Dynamic Time Warping, avec les contraintes de Sakoe & Chiba
K_{TGA}	Noyau associé à la distance Global Alignment avec une bande de contraintes de Sakoe & Chiba, définie par le paramètre T
P_c	Période d'expérimentation considérée
S	Période totale d'expérimentation
P	Polynôme de degré d
τ_c	Seuil définissant l'impact d'une variable
S_t	Seuil $S_t = (t_{min}, t_{max})$ sur la durée d'un motif permettant l'extraction d'un motif
S_v	Seuil $S_v = (v_{init}, v_{term})$ sur la vitesse d'un motif permettant l'extraction d'un motif
S	Couple de stratégies choisies
$\tau(M_1^v, M_2^v)$	Taux de recouvrement de deux fonctions de densité M_1^v et M_2^v
Utilisation	Correspond aux conditions de vie d'une batterie, comprenant les stockages, roulages et les charges
Usage	Correspond aux Utilisations et aux conditions environnementales
$V_j(v)$	Variable v mesurée sur le véhicule j
V_j	Véhicule j

Introduction générale

La récente augmentation du nombre de véhicules électriques disponibles à l'achat et la diversité importante des constructeurs automobiles proposant au moins un modèle électrique, illustrent le développement actuel de ce marché. Cette accélération de la production de véhicules électriques s'explique notamment par une combinaison de différents facteurs avec diverses influences selon la sensibilité des analystes. Ces facteurs sont liés, sans ordre d'importance, aux restrictions qui légifèrent les émissions des pollutions locales et globales, à l'augmentation du prix du pétrole ou bien à la prise de conscience écologique.

Cependant, le déploiement du marché des véhicules électriques est actuellement freiné par les limitations perçues par les acheteurs comme des contraintes importantes pour leurs usages. Actuellement, ces limitations sont essentiellement causées par les systèmes de stockage électrique, qui impactent directement les performances des véhicules. À ce jour, les batteries lithium-ion (Li-ion) proposent l'une des solutions technico-économique les plus compatibles avec cette application, principalement par sa densité énergétique élevée. Les performances des batteries sont par conséquent au centre de nombreuses études ayant pour objectif de résoudre les problématiques liées aux limitations identifiées par les utilisateurs. L'amélioration de la performance des batteries est de première importance pour les constructeurs automobiles qui souhaitent rester compétitifs sur ce segment de marché par rapport à de nouvelles solutions innovantes, telles que les piles à combustible.

Les attentes des acheteurs définissent un cahier des charges de performances automobiles guidé par les caractéristiques des véhicules thermiques. Ces lignes directrices engendrent un développement des batteries mené par des problématiques de coût, de poids, de performance énergétique et de durée de vie. L'amélioration significative de ces verrous auraient très certainement pour conséquence une augmentation très nette des ventes de véhicules électriques qui proposeraient une alternative viable aux véhicules thermiques et hybrides.

Le travail de recherche exposé ici se place dans le cadre de la maîtrise de la durée de vie et la gestion des batteries Li-ion. Cette problématique spécifique est un enjeu majeur dans la commercialisation des véhicules électriques. En effet, les performances initiales des batteries Li-ion, en deçà des critères nécessaires, déclinent au fil du temps et des usages, de par le vieillissement de ces composants.

Les batteries Li-ion sont des systèmes de stockage électrique réversible complexes impliquant des phénomènes de vieillissement multi-causaux, difficilement quantifiables. Les approches actuellement proposées dans la littérature permettent de résoudre

certaines problématiques relatives au vieillissement, telles que la compréhension électrochimique des phénomènes de dégradation intervenant lors du vieillissement d'une batterie. Néanmoins, de nombreuses interrogations liées à l'estimation de l'état de santé d'une batterie subsistent.

L'étude détaillée ici utilise une approche, rarement utilisée à ce jour dans ce type de problématique, basée sur l'exploitation des données mesurées et enregistrées lors d'usages réels de véhicules électriques. Cette quantité de données propose une base d'information importante nécessitant d'être analysée et traitée pour en extraire des informations liées au vieillissement des batteries. Une telle approche permet d'extraire les informations contenues dans les données et les signaux mesurés sans considérer de modèle extérieur amenant potentiellement un biais dans l'étude. Ce travail développe une telle méthodologie, jamais appliquée aux problématiques relatives au vieillissement des batteries en usage sur un véhicule électrique, et pouvant donc mettre en évidence des phénomènes non identifiés pour les présenter ensuite à des experts afin de les valider.

Les objectifs de cette recherche sont la compréhension, l'analyse, l'estimation et la gestion des phénomènes de vieillissement des batteries en usage réel, qui nécessitent des investigations approfondies. Ces besoins sont guidés par les besoins des utilisateurs de véhicules électriques, dans le cadre d'une gestion de flotte ou de maintenance, et forment le cœur de l'étude proposée ici.

Ce travail se focalise sur l'analyse des conditions d'usage d'une batterie et leurs influences sur les phénomènes de vieillissement, ainsi que sur l'estimation de l'état de santé d'une batterie, lors d'usages réels. Ces problématiques sont traitées dans le cadre d'un usage individuel d'un véhicule électrique ainsi que lors d'une exploitation d'une flotte de véhicules. Ces travaux sont ici présentés au sein de quatre parties. Le fonctionnement des méthodologies utilisées sera en premier lieu détaillé dans le cadre général, puis appliqué à la problématique traitée dans un second temps.

La première partie de ces travaux présente le contexte justifiant ces travaux de recherche explicités dans la suite de cette étude. L'évolution du marché des véhicules électriques y est décryptée, les limitations actuelles liées au vieillissement des batteries sont explicitées. Le vieillissement d'une batterie est défini et les phénomènes intervenant sont détaillés et interprétés pour comprendre la complexité de ce procédé. Les problématiques traitées dans la suite de l'étude sont définies dans cette partie, et les solutions existantes y sont discutées, les contraintes et les limites des approches existantes sont analysées. Les lignes directrices des méthodologies proposées dans cette étude sont posées avec pour objectif de répondre spécifiquement aux problématiques établies par l'utilisation des données issues de la vie de la batterie.

La deuxième partie détaille les différentes expérimentations permettant l'obtention des données employées dans l'ensemble de ces travaux. L'historique de ces données est par la suite exploité dans le cadre d'une analyse approfondie, hors ligne, pour l'identification des interactions entre les variables et la quantification de leurs impacts respectifs sur les phénomènes de vieillissement d'une batterie. Les phénomènes de vieillissement d'une batterie sont associés aux variables les impactant, pour fournir des interprétations sur les conditions d'expérimentation influençant le vieillissement des batteries. Des modèles d'estimation statistique du niveau de vieillissement d'une

batterie sont construits et validés à partir de la base de données à disposition, démontrant la pertinence de l'approche proposée.

La troisième partie se focalise sur un traitement en ligne des mesures. Ces travaux sont guidés par les besoins d'un indicateur fiable du niveau de vieillissement de la batterie, calculé en embarqué, lors de ces usages réels. Cette partie détaille le principe de la méthodologie proposée, ainsi que la réflexion dirigeant cette approche. La méthode utilisée exploite directement le comportement des signaux mesurés sur la batterie. Ceux-ci contiennent des informations relatives au niveau de performance de la batterie. Une méthodologie est développée autour de cet axe pour obtenir un estimateur instantané du niveau de vieillissement, uniquement à partir de signaux mesurés sur la batterie. La précision de cette méthodologie est évaluée selon différents critères et pour différentes batteries, fournissant un niveau de confiance prometteur.

La quatrième partie aborde les problématiques relatives à la gestion d'une flotte de véhicules électriques. L'étude d'une flotte engendre des problématiques spécifiques, différentes de celles liées aux véhicules individuels, il est donc nécessaire d'établir des méthodologies dédiées. Pour ce faire, les outils présentés dans la deuxième et la troisième partie sont appliqués au contexte de la gestion d'une flotte de véhicules. La combinaison et l'interaction entre les différentes méthodologies sont détaillées, formant une cohésion entre tous les outils développés dans ces travaux. Les usages des différents véhicules sont regroupés pour proposer un contrôle individualisé des véhicules de la part d'un gestionnaire de flotte. De plus, les conditions d'expérimentation responsables d'un vieillissement important des batteries sont identifiées pour chacun des véhicules, pour former un outil d'aide à la gestion de flotte. L'estimation du niveau de vieillissement est également abordée par une mise à jour des résultats en fonction des possibilités de mesures exactes des performances des batteries.

Finalement, l'extension des approches développées dans ces travaux est abordée par le biais d'améliorations et de développements des outils évoqués ou bien par la proposition de nouveaux domaines d'applications.

Chapitre 1

Contexte et problématique

1.1 Les véhicules électriques - Transports	16
1.1.1 Essor des véhicules électriques	16
1.1.2 Limitations des véhicules électriques	19
1.2 Les batteries lithium-ion	22
1.2.1 Présentation et fonctionnement	23
1.2.2 Données caractéristiques d'une batterie	25
1.2.3 Phénomènes de vieillissement	28
1.2.3.1 Effets du vieillissement à l'électrode négative	28
1.2.3.2 Effets du vieillissement à l'électrode positive	30
1.2.3.3 Conséquences des phénomènes de vieillissement	30
1.2.4 Facteurs globaux de vieillissement	30
1.2.4.1 Vieillessement calendaire	31
1.2.4.2 Vieillessement cyclique	32
1.2.4.3 Conclusion des facteurs de vieillissement	33
1.3 Estimation du vieillissement	34
1.3.1 Indicateurs de vieillissement	34
1.3.2 Modélisation des phénomènes de vieillissement	35
1.3.2.1 Modèles électrochimiques	35
1.3.2.2 Modèles de circuit équivalent	37
1.3.2.3 Modèles semi-empiriques	38
1.3.2.4 Modèles analytiques	40
1.3.2.5 Méthodes statistiques	41
1.3.3 Bilan de l'estimation des phénomènes de vieillissement	42
1.3.3.1 Évaluation des méthodes d'estimations	42
1.3.3.2 Bilan des méthodes décrites	43
1.4 Conclusion des problématiques liées à l'estimation du vieillissement d'une batterie	44
1.4.1 Point de vue utilisateur sur le vieillissement	44
1.4.1.1 Utilisation d'un Véhicule Électrique (VE) personnel	45
1.4.1.2 Gestionnaire d'une flotte de VE	45
1.4.2 Bilan des solutions existantes	46
1.4.3 Orientation des travaux	46

1.1 Les véhicules électriques - Transports

1.1.1 Essor des véhicules électriques

Les premiers véhicules électriques (VE) sont apparus au milieu du 19^{ème} siècle, lorsque l'électricité était le choix de propulsion privilégié. Depuis quelques années l'impact environnemental et la raréfaction du pétrole implique une augmentation de son coût, ont engendré un regain d'intérêt pour la propulsion électrique. Les VE ont commencés à être réintroduits commercialement, en France, au début des années 1990. Une liste non exhaustive des VE commercialisés lors des dernières années, illustrant les évolutions de leurs caractéristiques, est présentée table 1.1. On peut remarquer la nette amélioration récente des performances estimées des VE, ainsi que la diversité accrue des constructeurs proposant un modèle électrique.

Véhicule	Batterie	Vitesse max (km.h ⁻¹)	Autonomie (km)	Année
Citroën AX	NiCd	91	80	1995
G.M. EV1	NiMH	130	150	1996
Citroën Saxo	NiCd	91	75	1997
G.M. EV1	Plomb	130	225	1999
Tesla Roadster	Li-Cobalt	201	394	2008
BMW Mini-e	Li-Manganèse	150	160	2009
Nissan Leaf	Li-ion	145	175	2010
Renault Fluence ZE	Li-ion	135	160	2011
Mitsubishi iMEV	Li-ion	130	130	2011
Wheego Whip	Li-ion	105	160	2011
Tesla Serie S	Li-ion	210	480	2012
Renault Zoe	Li-ion	135	210	2012
Peugeot iOn	Li-ion	130	150	2012
Citroën C-ZERO	Li-ion	130	150	2012
BMW i3	Li-ion	150	160	2013
Honda Fit EV	Li-ion	150	200	2014
Toyota RAV4 EV	Li-ion	145	160	2014
Volkswagen e-Golf	Li-ion	135	275	2015
Tesla Model X	Li-ion	??	370	2016

TABLE 1.1 – Comparatif des caractéristiques de quelques véhicules électriques commercialisés (Li-ion est spécifié dans les cas où les détails sur la composition de ces batteries ne sont pas fournis)

Après un démarrage des ventes lent, les ventes de véhicules électriques (hybrides ou 100 % électriques) connaissent depuis 2012 une nette accélération. Les ventes aux États-Unis ont par exemple été doublées entre 2012 et 2013. Cette tendance semble se poursuivre en 2014 par une augmentation constatée de 66 % des ventes lors des 4 premiers mois par rapport à la même période en 2013 (figure 1.1). Il est important de remarquer que les ventes de Battery Electric Vehicle (BEV) ont récemment dépassé celles des véhicules hybrides aux États-Unis, qui représentent désormais 4 % du mar-

ché total des véhicules neufs, démontrant la part non négligeable des VE et hybrides dans le marché des ventes automobiles.

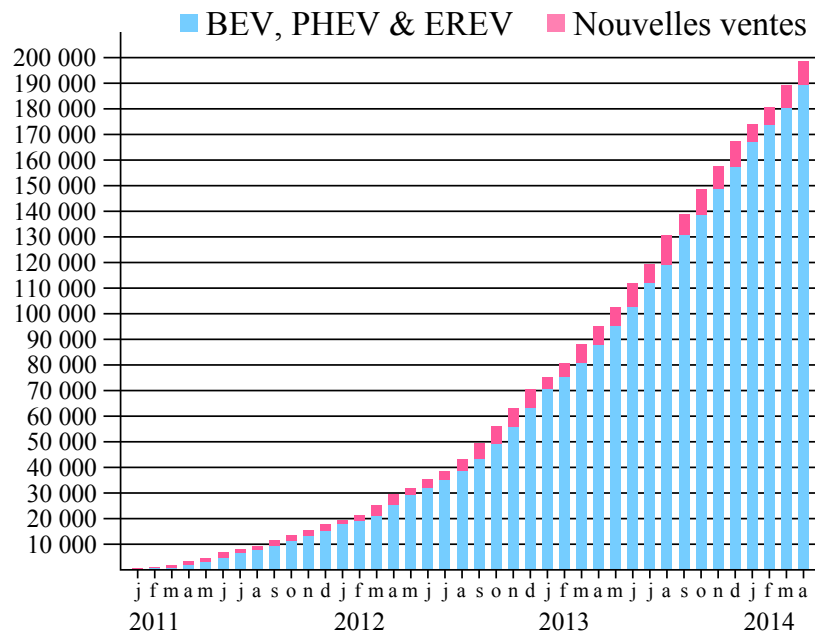


FIGURE 1.1 – Évolution des ventes de véhicules électriques et hybrides aux États-Unis [1] - Battery EV (BEV) - Plug-In Hybrid EV (PHEV) - Extended Range EV (EREV)

Il est également intéressant de remarquer la présence des VE dans l'ensemble des régions majeures à travers le monde (figure 1.2). D'autre part, le public séduit par l'achat de ce type de véhicules correspond à toutes les classes de la population et à tous les âges (figure 1.3). Ces remarques permettent d'envisager un développement plus important du marché des VE dans les années à venir.

De manière raisonnable, certaines études estiment une proportion de véhicules électriques (VE) représentant environ 10 % du marché total des nouveaux véhicules individuels d'ici 2020 (figure 1.4), considérant une augmentation annuelle de 35 % des ventes de VE [2]. Il faut noter que les estimations les plus optimistes présentent des ventes croissantes de 70 % tous les ans jusqu'en 2020 [4].

Ces estimations sont confortées par d'autres analyses. Les Nations Unies estiment par exemple que la proportion des personnes vivant dans de grandes villes atteindra 70 % en 2050, ce qui correspond au marché cible des VE. D'autre part, les achats de VE se font majoritairement lors d'un remplacement d'un ancien véhicule (figure 1.5). Or, ce marché spécifique est cyclique et la quantité de véhicules remplacés chaque année présente une opportunité intéressante pour les constructeurs de VE.

L'actuelle inflation des ventes reste cependant à mettre en relation avec les nombreuses aides dont profitent les VE. Celles-ci n'étant probablement pas destinées à être pérennisées. D'autre part, les constructeurs réduisent considérablement leurs prix de vente de manière à attirer de nouveaux clients. BMW a par exemple annoncé, durant l'année 2013, un prix de vente étonnamment bas pour sa récente BMW i3 de 35 000 €, les analystes estimant auparavant un prix au plus bas de 40 000 €. Plus récemment, Ford a fait de même en réduisant le prix de la Focus EV de 3 000 €, ou encore la Nissan Leaf baissant son prix initial de 4 700 €, ce qui correspond à des véhicules neufs

EVI MEMBER COUNTRIES HELD OVER 90% OF WORLD ELECTRIC VEHICLE (EV) STOCK IN 2012

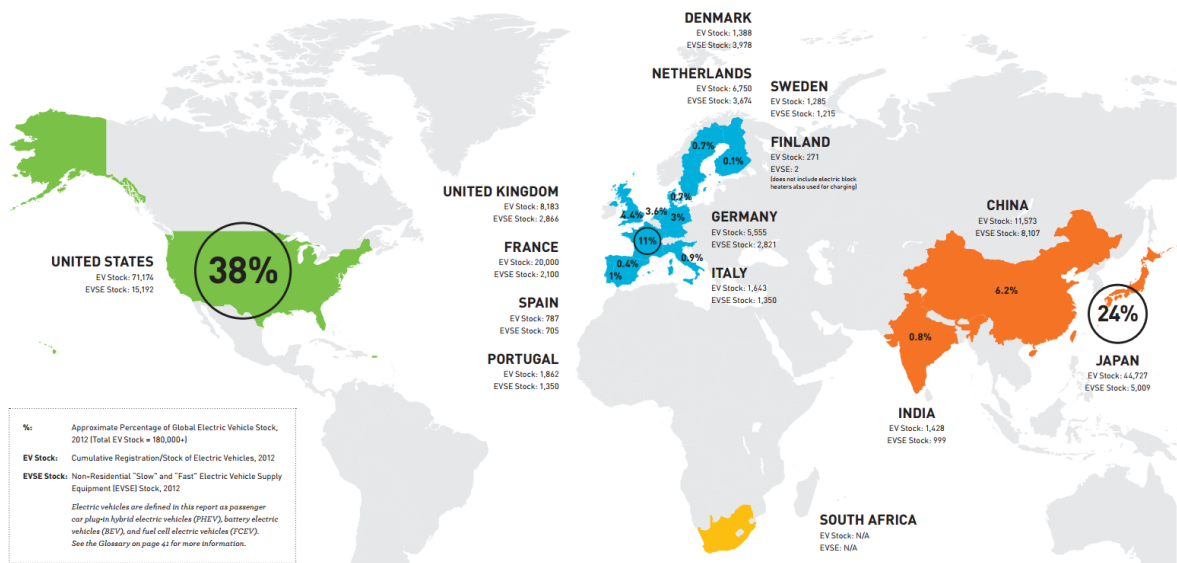


FIGURE 1.2 – Répartition de la présence des VE dans le monde, fin 2012 [2]

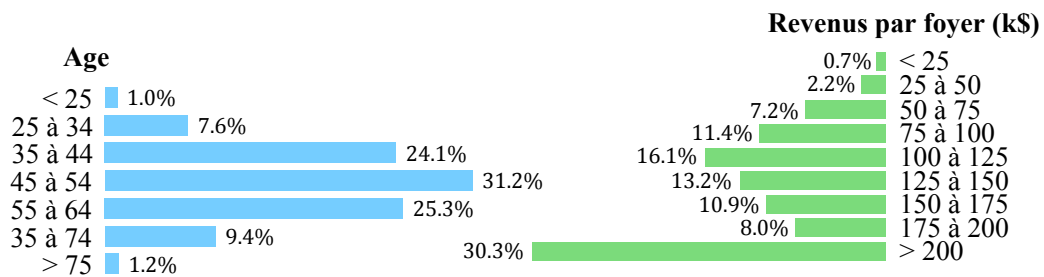


FIGURE 1.3 – Caractéristiques sociologiques des acheteurs de VE [3]

autour de 22 000 €. Bien évidemment ce sont de bonnes nouvelles pour les utilisateurs, mais cela démontre également un problème d'adéquation entre l'offre et la demande des VE. Le point positif étant néanmoins de démontrer par la pratique qu'il existe bel et bien un marché pour les VE, à un prix de vente adapté aux attentes des acheteurs.

Le principal challenge actuel des constructeurs automobiles correspond au prix des batteries qui demeurent trop chers pour le marché automobile. La réduction des coûts des batteries, pour les mêmes caractéristiques techniques placerait les VE dans un contexte favorable et les constructeurs ne seraient plus obligés de baisser leurs prix, la batterie représentant près de la moitié du coût total d'un VE. La proportion du prix d'une batterie par rapport à l'ensemble d'un véhicule est présenté en figure 1.6.

L'évolution des recherches et des technologies permettent actuellement une réduction partielle des coûts de fabrication des batteries. Les améliorations récentes sont tout de même loin des objectifs à terme nécessaires pour séduire une large partie du public. Ceux-ci oscillent pour le U.S. Advanced Battery Consortium d'un coût de 150 \$ par kWh, à un coût de 300 \$ par kWh pour le Electric Vehicles Initiative (IEA) (figure 1.7). L'atteinte d'un tel coût des batteries permettrait aux VE d'être financièrement

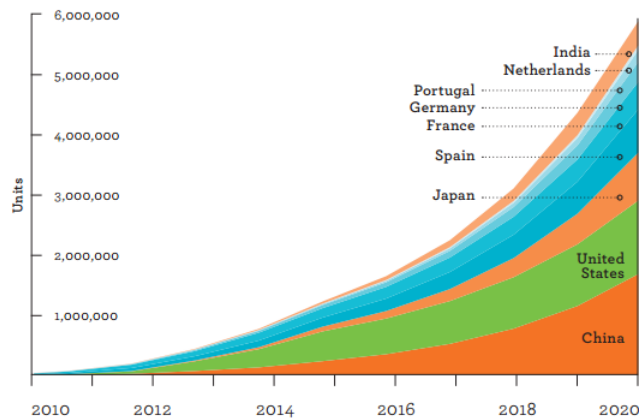


FIGURE 1.4 – Évolution estimée des stocks de EV, par pays [2]

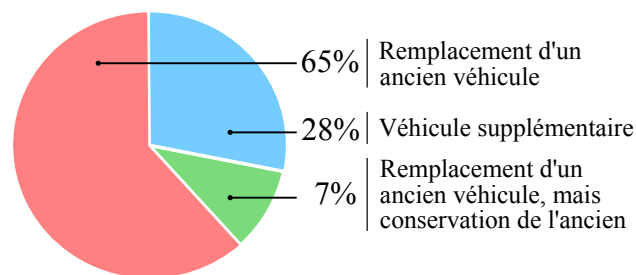


FIGURE 1.5 – Répartition des conditions d'achat d'un VE [3]

plus rentables que les véhicules thermiques, quelque soit l'intensité d'utilisation, ce qui peut développer les ventes à une grande échelle.

Il faut néanmoins noter que certains analystes estiment qu'une diminution rapide des prix des batteries, et donc par conséquent des VE, peut également avoir un effet néfaste sur leurs ventes. D'après cette hypothèse, avec des prix en baisse constante, des consommateurs prêts à acheter pourraient continuer à attendre de nouvelles réductions importantes dans le futur [5]. De plus, les prix actuels étant artificiellement abaissés pour créer un marché, il sera très difficile d'augmenter à nouveau les prix, ce qui rendra la rentabilité plus difficile à atteindre pour les constructeurs. Tant que les coûts des batteries descendront progressivement, les fabricants pourront difficilement de faire du profit sur les VE, en particulier sachant que les consommateurs ne se lanceront dans ce marché que pour un prix cohérent avec le marché global des véhicules.

1.1.2 Limitations des véhicules électriques

Une telle évolution du marché mondial des VE implique d'importantes étapes à venir pour permettre aux batteries de correspondre aux objectifs requis par les utilisateurs dans le cadre d'usage de VE, en dehors de leur prix. Les consommateurs seraient moins réticents à payer le prix actuel des VE, si les performances de ceux-ci étaient plus adaptées à leurs attentes. Ces contraintes correspondent à réduire les inconvénients actuels des VE, par rapport aux véhicules thermiques, qui constituent la référence de comparaison des acheteurs. En effet, l'image écologique des VE ne permet pas

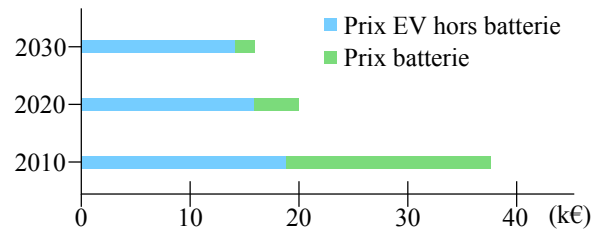


FIGURE 1.6 – Projection de la proportion du coût d’une batterie dans le prix total d’un VE [3]

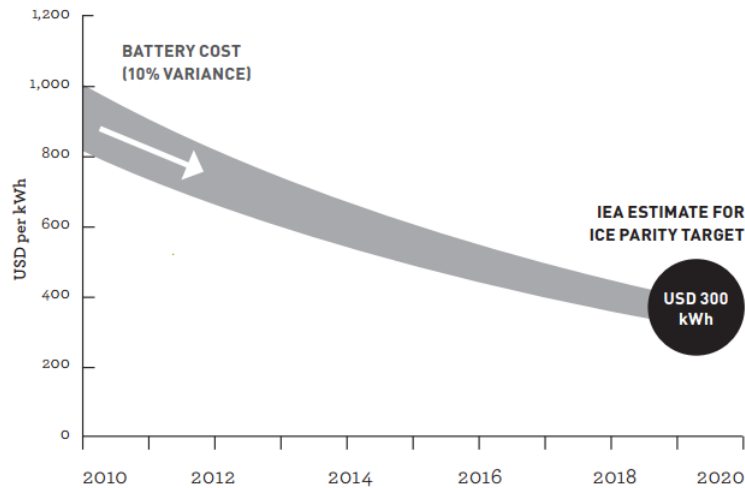


FIGURE 1.7 – Exemple d’évolution prévue du coût des batteries dans les prochaines années [2]

systématiquement d’attirer de nouveaux clients, ceux-ci étant prioritairement séduits par des économies potentielles (figure 1.8). Ceci démontre la nécessité d’un changement de discours de la part des constructeurs automobiles qui continuent d’axer leurs principaux arguments sur l’écologie. Les désavantages majeurs au développement des VE sont donc liés à leurs usages et à leurs performances. Les trois inconvénients majeurs cités par les consommateurs sont le niveau d’autonomie, le prix et le temps de charge des VE (figure 1.9). Notons que la contrainte du niveau d’autonomie est amplifiée par le manque de confiance des utilisateurs vis à vis des indicateurs correspondants. D’autre part, le long temps de recharge des VE est lié au manque d’infrastructures de recharges.

Concrètement, le développement des ventes de VE nécessite une amélioration de leurs performances. Comme nous l’avons développé précédemment, les prix des VE doivent subir une baisse conséquente dans les années à venir. Le coût vu comme un frein par les potentiels acheteurs est donc sujet à devenir un atout dans le futur.

Le temps de charge d’une batterie est aujourd’hui d’environ 6-9 heures en charge dite lente, et inférieur à une heure pour les charges rapides. Les études récentes montrent que les bornes de recharges lentes et rapides sont amenées à être de plus en plus nombreuses. Néanmoins les charges rapides sont principalement une solution temporaire mise en place pour pallier les restrictions des batteries actuelles, et atti-

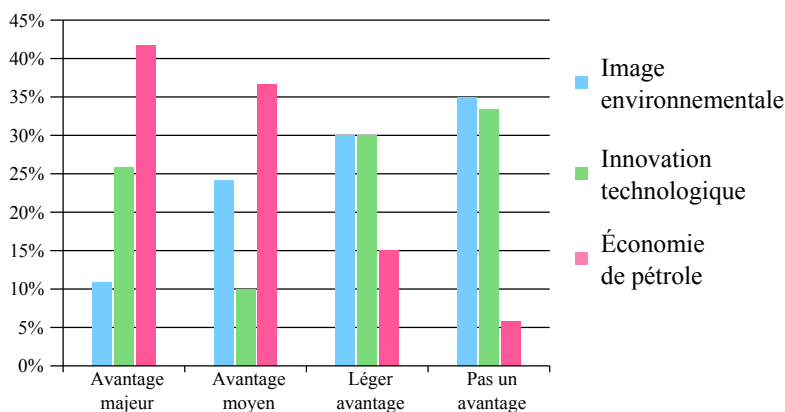


FIGURE 1.8 – Degré des bénéfices annoncés des VE [6]

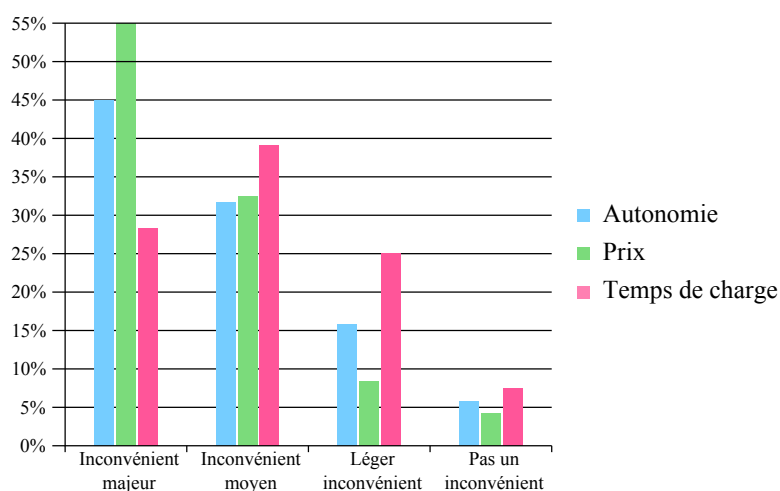


FIGURE 1.9 – Degré des inconvénients annoncés des VE, comme des barrières à l'achat [6]

rer le marché par ce qui semble être un facilitateur technologie clé [7]. L'éducation des consommateurs par rapport aux problèmes engendrés par ce type de recharges rapides, sur la santé de la batterie n'est cependant pas faite à ce jour, probablement pour ne pas évoquer un nouvel inconvénient des VE au grand public [8, 9]. Il est de plus actuellement difficile de quantifier cet impact des charges rapides sur la santé d'une batterie.

La majorité des VE sur le marché présentent aujourd'hui des autonomies maximales d'environ 150 km, qui en réalité revient le plus souvent à une autonomie de 100-120 km en usage normal, et < 100 km dans pour des usages extrêmes (par exemple avec une température négative). Cette limitation de l'autonomie représente un inconvénient majeur pour les usagers de VE. Dans un sondage américain, 75 % des personnes interrogées évoquent l'autonomie des VE comme un inconvénient important [6] des VE. Il faut noter que les États-Unis et la France comportent les populations les plus sensibles à l'autonomie maximale [10]. Néanmoins, l'usage quotidien moyen d'un véhicule aux États-Unis est de 46 km, avec une distance moyenne par trajet de 15 km [11]. Ces usages moyens des États-Unis étant les plus longs au monde, il apparaît

que les VE proposent des caractéristiques suffisantes pour une majorité des conducteurs. Cependant, tant que cette autonomie maximale reste entre les attentes actuelles et les besoins moyens, les perceptions de l'autonomie des VE et l'anxiété qu'elle induit persistera dans l'esprit des consommateurs.

Il faut également évoquer le rôle des autonomies maximales annoncées par les constructeurs comme un frein à la confiance faite par les utilisateurs. En effet, ces distances maximales parcourues avec un VE correspondent à des usages optimaux peu réalisables lors d'un usage réel. Le consommateur n'atteint donc jamais en réalité cette distance maximale, ce qui crée un manque de confiance des utilisateurs vis à vis des performances annoncées par les VE.

D'autre part, l'amélioration des batteries en terme de durée de vie, et en plage de températures de fonctionnement, permettrait de réduire les coûts d'exploitation. Les phénomènes de dégradations des batteries actuelles impliquent en effet une faible durée de vie, annoncée par les constructeurs d'environ 8 ans ou d'environ 160 000 km d'utilisation. Au fil des utilisations l'autonomie maximale de la batterie est réduite par des phénomènes de vieillissement, ce qui accroît l'anxiété des utilisateurs de VE (*range anxiety*). Ce phénomène de stress, familier des usagers de VE, se produit lors de trajets nécessitant une grande partie des performances des batteries. Le terme norvégien "*rekkeviddeangst*"¹ figure même dans la liste des mots de l'année 2013 [12]. Les utilisateurs de EV font face à un stress généré par la peur de ne pas parvenir à leur destination. La solution à cette anxiété est bien évidemment une évolution des performances de batterie, ou à plus court terme la possibilité de connaître précisément la distance pouvant être parcourue. Ceci n'est pas le cas actuellement, les indicateurs utilisés pour estimer cette distance maximale n'étant que peu fiables, notamment par la non prise en compte des auxiliaires et de l'environnement.

L'autonomie maximale d'un VE à un instant donné dépend de l'état de santé de la batterie utilisée. L'estimation du niveau de ces performances en temps réel est un sujet majeur au développement des VE. La suite de cette étude a donc pour but de comprendre les phénomènes de dégradation d'une batterie, puis d'obtenir un indicateur du niveau de vieillissement capable de fiabiliser les estimations d'autonomie faites à partir des données mesurées par la batterie et le véhicule, susceptible de réduire cette anxiété des utilisateurs.

1.2 Les batteries lithium-ion

Les batteries Li-ion sont commercialisées depuis 1991, initialement pour les utilisations mobiles telles que les téléphones ou les ordinateurs portables [13]. L'intérêt pour cette technologie s'est considérablement développé depuis et a généré de nombreuses recherches ayant pour objectif d'améliorer les performances de ces batteries [14]. Récemment, les batteries lithium-ion ont pénétré le marché des véhicules électriques et hybrides grâce à leur importante densité énergétique, ce qui a fait d'elles le candidat le plus prometteur à ce type d'application pour succéder aux batteries nickel-metal (NiMH) utilisées dans les années 1990 et au début des années 2000 [15]. Un comparatif des performances obtenues pour différentes technologies de batteries est présenté en table 1.2, et sont comparées aux objectifs affichés à moyen terme pour une utili-

1. Inquiétude liée à l'autonomie en norvégien.

sation en véhicule électrique. Ce comparatif illustre les avantages des batteries Li-ion par rapport aux autres technologies pour un usage VE.

	Plomb	NiCd	NiMH	Li-ion	Objectif
Densité énergétique (Wh/kg)	30-50	45-80	60-120	80-160	250
Nombre de cycles	200 à 300	1500	300 à 500	500 à 1000	>1000
Temps de recharge rapide à 100 %	8 à 16h	1h	2 à 4h	1 à 4h	0.5h
Auto-décharge par mois	5 %	20 %	30 %	10 %	3 %
Plage de température (°C)	-20 à 60	-40 à 60	-20 à 60	-20 à 60	-30 à 65

TABLE 1.2 – Comparatif des caractéristiques des technologies de packs batterie [16]

Le passage à la technologie lithium représente un appui majeur au développement des véhicules électriques. La plupart des grands constructeurs automobiles ont introduit des modèles hybrides ou électriques lors des deux dernières années, la quasi-totalité utilisent des batteries au lithium pour le stockage d'énergie (table 1.1). Nous plaçons donc plus précisément cette étude dans l'étude du vieillissement des batteries Li-ion, pour une application VE.

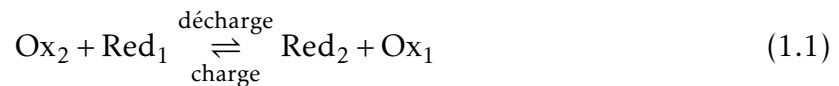
1.2.1 Présentation et fonctionnement

Une batterie peut se définir comme un convertisseur et un système de stockage d'énergie avec un dispositif oxydant-réducteur réversible, ce qui n'est pas le cas des systèmes primaires comme les piles. D'un point de vue général, une batterie est composée de plusieurs éléments appelés accumulateurs, destinés à emmagasiner l'énergie sous forme chimique, et à restituer de l'énergie électrique.

Le principe d'un tel élément de stockage d'énergie repose sur l'exploitation de deux couples oxydo-réducteurs : Ox_1/Red_1 et Ox_2/Red_2 , se trouvant aux deux électrodes différentes. Ces deux électrodes, reliées par un conducteur métallique interne, génèrent la circulation des ions qui circulent sous l'effet d'une différence de potentiel de l'électrode négative vers l'électrode positive. Les électrons, quant à eux, transitent par le circuit électrique externe provoquant une réaction électrochimique de réduction en arrivant sur l'électrode positive.

L'énergie chimique produite peut alors être convertie en électricité par l'usage d'un circuit électrique externe. Si la batterie est en charge ou en décharge nous pouvons

considérer deux cas. La réaction globale s'écrit donc :



Il est important de noter ici que l'électrode où se produit la réduction est appelée la cathode, et l'anode correspond au lieu de l'oxydation. Les électrodes positive et négative sont par abus de langage respectivement appelées cathode et anode, cependant ces définitions ne sont pas valables au sens électrochimique dans le cas d'une batterie du fait de l'alternance charge/décharge car chaque électrode change de rôle.

Les deux électrodes sont placées au sein d'un électrolyte assurant les échanges ioniques entre celles-ci. Un séparateur, conducteur ionique et isolant électrique, situé entre les deux électrodes permet également d'obtenir un milieu non conducteur entre ces dernières. La figure 1.10 illustre la composition d'une batterie.

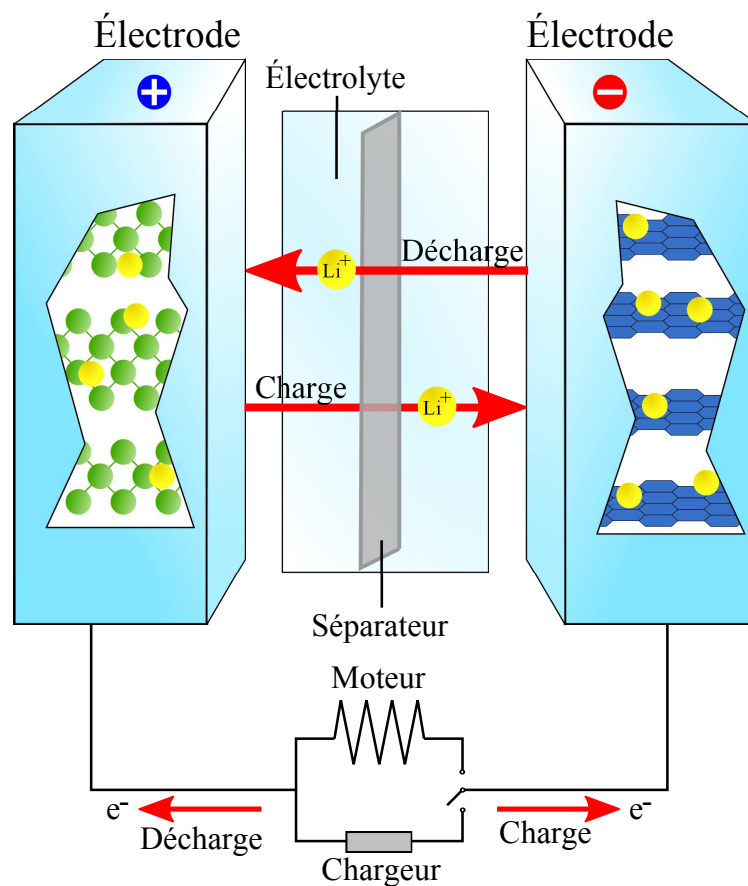


FIGURE 1.10 – Illustration de la composition et du fonctionnement d'une batterie Li-ion

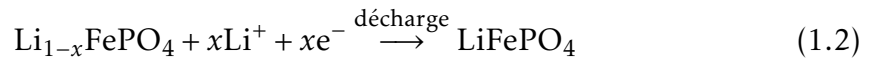
Dans le cas de batterie Li-ion, le fonctionnement repose sur l'échange réversible des ions Li⁺ entre les électrodes. Une réaction dite d'insertion ou d'intercalation intervient, les cations Li⁺ peuvent s'insérer ou se désinsérer dans une structure cristalline, induisant une perte énergétique.

Lors d'une décharge, le lithium contenu dans l'électrode négative se désinsère de la structure cristalline de ce matériau, puis se diffuse à travers l'électrolyte, pour finalement venir s'insérer dans le matériau de l'électrode positive. Cette migration d'un

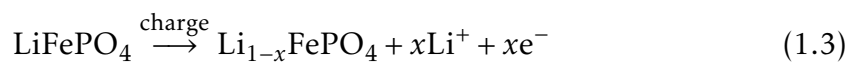
cation Li^+ au sein de la batterie est compensée par le déplacement d'un électron dans le circuit électrique externe, formant un courant électrique. Dans le cas d'une charge, le processus est inversé et les cations Li^+ se propagent de l'électrode positive pour venir s'insérer dans le matériau de l'électrode négative. La figure 1.10 représente ces réactions.

En considérant par exemple une batterie Lithium-Fer-Phosphate (LiFePO_4), nous avons les réactions électrochimiques suivantes :

– À l'électrode positive :



– À l'électrode négative :



Il existe de nombreuses batteries Li-ion, formées de matériaux d'électrode et d'électrolyte différents, ayant chacune des caractéristiques propres. Les performances d'une batterie Li-ion sont les conséquences de la composition de ces électrodes (principalement la positive). On peut avoir de nombreux matériaux de cathodes tels que le lithium cobalt ; le composé lithium, nickel, cobalt et aluminium ; le composé lithium fer phosphate ; le composé lithium nickel cobalt manganèse ou encore le lithium manganèse spinel. Diverses études ont été menées afin d'évaluer les différentes caractéristiques des matériaux possibles. Une représentation résumant les caractéristiques de quelques technologies de batteries lithium, selon divers critères de performances, de sécurité et de coût, est présentée en figure 1.11.

Nous pouvons alors remarquer le bon compromis entre sécurité et durée de vie des batteries Lithium Fer Phosphate (LFP), au contraire des batteries NMC qui proposent une grande énergie spécifique mais une durée de vie et une sécurité moindre ce qui implique une altération rapide de l'énergie spécifique. La technologie LFP elle l'une des solutions retenues, entre autres, par le Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), et que nous utilisons dans l'ensemble des expérimentations évoquées par la suite. Il est également important de noter les avantages en terme de sécurité de cette technologie, cette notion étant au cœur des exigences suite à de nombreux problèmes. En effet, les batteries LFP sont relativement sûres car l'oxydation des matériaux n'a lieu que sous des faibles potentiels (~ 3.4 V) et car elles offrent une très bonne stabilité thermique (assurée jusqu'à 70°C). De plus, ces batteries peuvent difficilement prendre feu ou exploser en cas de surcharge. Enfin, la technologie LFP ne contient pas de métaux lourds toxiques comme le plomb, le cadmium, le cobalt, ce qui rend cette chimie plus facilement recyclable.

1.2.2 Données caractéristiques d'une batterie

Afin de pouvoir détailler le fonctionnement des batteries Li-ion, cette section présente les principales notions utilisées par la suite.

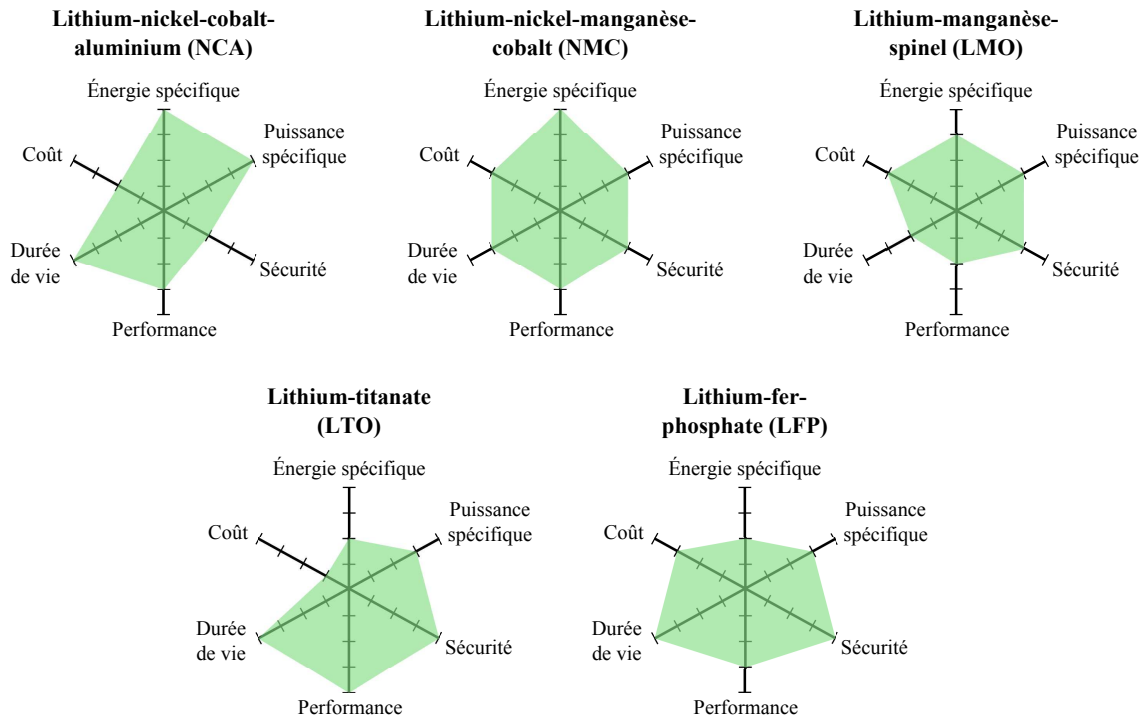


FIGURE 1.11 – Comparaison des caractéristiques pour différentes technologies de batterie Li-ion [17]

Capacité initiale : La capacité initiale d'une batterie correspond à sa charge maximale au temps t_0 (pour une batterie neuve). Cette valeur fait office de référence lorsque l'on étudie l'évolution de la capacité de la batterie en fonction du vieillissement.

Capacité nominale : La capacité nominale d'une batterie, notée C_n , s'exprime en ampère-heure (Ah) et représente la quantité de charge que peut restituer la batterie initialement et entièrement chargée lors d'une décharge complète, dans des conditions thermiques normales et avec un courant constant. Cette notion de capacité est sujette à évolution, généralement une diminution, au cours de la vie de la batterie en fonction de son vieillissement, et en fonction du régime de décharge ou de température.

Capacité stockée : La seule différence entre la capacité stockée et la capacité nominale réside dans le fait que la capacité stockée mesure la quantité de charge d'une batterie à un instant donné et ne se réfère pas à la charge maximale possible.

State Of Charge (SOC) : L'état de charge d'une batterie (SOC) est le pourcentage de quantité de charge disponible à un temps t . L'expression la plus fréquente est :

$$\text{SOC}(t) = \frac{\text{capacité stockée à } t}{\text{capacité nominale à } t} \quad (1.4)$$

L'état de charge (SOC) donne une indication très intéressante sur le niveau d'autonomie restant de la batterie. Cet indicateur est au cœur de la gestion d'une batterie. De nombreuses études ont cherché à en obtenir une estimation la plus fiable possible

[18]. Il est important de noter que le SOC ne donne aucune indication sur le niveau de vieillissement d'une batterie et reste toujours entre 0 % et 100 %. L'état de charge d'une batterie détermine la matière active présente au sein des électrodes.

On peut également exprimer le SOC à l'instant t_1 à partir de l'intégration du courant, par la formule suivante :

$$\text{SOC}(t_1) = \text{SOC}(t_0) - \frac{\int_{t_0}^{t_1} i(t) dt}{C_n(t_0)}, \quad (1.5)$$

avec $C_n(t_0)$ la capacité nominale de la batterie à t_0 .

Depth Of Discharge (DOD) : À partir du SOC, il vient intuitivement une notion de niveau de décharge. Cette valeur est représentée par le DOD simplement défini comme étant le complémentaire du SOC, autrement dit on a :

$$\text{DOD}(t) = 1 - \text{SOC}(t) \quad (1.6)$$

ΔSOC : L'autre notion directement associée au SOC est la variation d'état de charge, lors d'une charge ou d'une décharge, que l'on note ΔSOC . Cette valeur est définie comme étant le taux de capacité nominale de la batterie qui a été déchargée (resp. chargée), il est donc négatif (resp. positif). Cet indicateur est en liaison directe avec le régime imposé à la batterie.

End Of Life (EOL) : La notion de fin de vie de la batterie doit être également définie. Par convention, on dira que la batterie a atteint une limite définissant la fin de vie lorsque qu'un critère nommé EOL est franchit. Le critère EOL est un seuil fixé a priori, en dessous duquel on considère que la batterie n'est plus admissible pour une utilisation donnée. Le plus souvent, ce seuil est un critère dépendant de la capacité nominale de la batterie. Il est très souvent fixé pour une perte maximale de capacité nominale de 20 % par rapport à la capacité initiale [19, 20], en particulier pour les applications transports.

Durée de vie : La notion de durée de vie est définie comme étant la période de vie totale d'une batterie (en charge/décharge et en stockage), jusqu'à ce qu'elle franchisse le critère de fin de vie EOL défini précédemment. Cette durée peut se mesurer soit en nombre de cycles, soit en années selon les cas et les utilisations.

Impédance : L'impédance Z d'un système est la fonction de transfert entre le courant (fonction d'entrée) et la tension (fonction de sortie). Cette fonction de transfert peut être décomposée en une partie imaginaire et une réelle :

$$Z = R + jX, \quad (1.7)$$

avec R la partie dite résistive, et X la partie dite réactive.

C-rate : Le *C-rate* signifie *charge rate*, ce qui correspond à la proportion de capacité de la batterie qui peut être chargée ou déchargée en une heure. Pour une batterie de 2.1 Ah, un C-rate de $C/3 = 0.7A$ signifie que la batterie requiert 3 heures pour être entièrement chargée (ou déchargée). Ce paramètre est la conséquence directe du mode de recharge (rapide ou lente).

1.2.3 Phénomènes de vieillissement

Les premiers usages des batteries Li-ion, destinés aux utilisations portables, avaient un besoin de durée de vie peu élevé de quelques années. Or, avec les nouvelles applications de transport, les intérêts se focalisent sur les phénomènes de vieillissement, considérant les exigences des constructeurs automobiles. D'un point de vue du *design* des batteries, au-delà de toutes les contraintes de performance, certains objectifs sont clairement définis pour la durée de vie (8 ans ou 160 000 km pour la Nissan Leaf et la Volkswagen e-Golf). Il faut ici noter la réduction de ces garanties de batteries LFP, celles-ci étant définies entre 10 et 15 ans il y a quelques années [21]. La variété des garanties démontre l'incertitude des dégradations subies lors d'usages réels par une batterie. L'objectif actuel est d'améliorer les performances et les durées d'usages des batteries, ce qui nécessite une meilleure compréhension des phénomènes de vieillissement de batteries [22, 23].

L'identification du vieillissement et des phénomènes de dégradations d'une batterie est un des principaux défis de ce domaine. Ces procédés sont complexes car de nombreux facteurs provenant aussi bien de la nature des matériaux, du procédé de fabrication, de la cellule, de l'environnement que des modes d'Utilisation², et sont en interaction permanente générant divers effets de vieillissement. Cette partie présente les résultats actuels sur les mécanismes de vieillissement des batteries Li-ion, et précisent les facteurs et leurs conséquences.

Le vieillissement d'une batterie se produit en premier lieu dans la composition chimique de son électrolyte et des électrodes. Les mécanismes de dégradation diffèrent de l'électrode positive à l'électrode négative [20, 24]. L'origine de ces mécanismes peuvent être mécaniques ou électrochimiques et sont fortement dépendants de la composition des électrodes. En cours d'Utilisation d'un véhicule correspond aux conditions de vie d'une batterie, le vieillissement provoque des dégradations irréversibles sur les composants des cellules [25], ce qui génère, par exemple, une modification des propriétés structurelles, une variation de la composition chimique de l'électrolyte ou bien une perte de matériel actif par dissolution du matériel dans l'électrolyte [20].

1.2.3.1 Effets du vieillissement à l'électrode négative

La plupart des électrodes négatives sont constituées de graphite, carbone, titane ou silicone. Le choix le plus fréquent est porté sur le graphite, ce qui est important dans le cadre des propriétés de sécurité des batteries [24]. Le principal facteur de vieillissement est le développement d'une interface phase solide entre l'électrolyte et l'électrode. Celle-ci est nommée Solid Electrolyte Interphase (SEI). Cette interphase solide est créée naturellement lors de la première charge. Son rôle est de protéger l'électrode

2. Le terme d'utilisation (noté *Utilisation*) (resp. d'usage (noté *Usage*)) comprend les stockages, roulages, les charges (resp. les conditions d'Utilisation et environnementales).

négative de possibles corrosions et l'électrolyte de réductions [26]. Ce phénomène se produit essentiellement pendant le début d'un cycle, formant donc une barrière naturelle entre l'électrode négative et l'électrolyte ce qui offre une garantie de sécurité [27, 28].

La SEI n'est pas stable car les batteries Li-ion opèrent en tension en dehors de la fenêtre de stabilité électrochimique de l'électrolyte. La SEI se développe avec le temps ce qui provoque une perte continue d'ions lithium et une décomposition de l'électrolyte [29]. De plus, la perte de lithium disponible causée par des réactions secondaires à l'électrode négative en graphite ont été reportées comme la principale source de vieillissement se produisant lors des stockages d'une batterie [30].

En outre, la SEI est perméable aux ions lithium et à d'autres éléments chargés (anioniques) ou des éléments neutres (solvant) [20, 31]. De ce fait, le solvant interagit avec le graphite après diffusion à travers la SEI, ce qui induit une exfoliation du graphite [32], formant du gaz qui peut craquer la SEI et permettre son expansion [26, 33]. Néanmoins, la formation de gaz est faible et semble se produire uniquement lors les périodes de stockage à SOC élevé [34]. Il existe également une perte de surface active avec le temps, ce qui accroît l'impédance de l'électrode. La figure 1.12 illustre l'ensemble de ces phénomènes se produisant à la SEI. Ces phénomènes peuvent avoir lieu lors des Usages de la batterie.

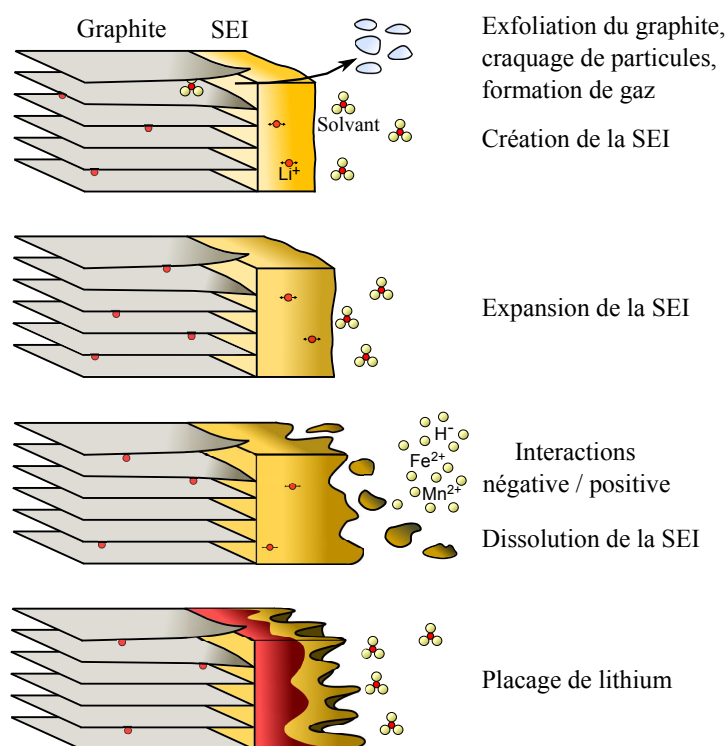


FIGURE 1.12 – Formation et développement de la SEI lors de la vie d'une batterie [35]

Notons que des conditions inadéquates d'utilisation, telles que des températures extrêmes ou des surcharges, peuvent accélérer ces processus de dégradation [36]. Ainsi, sous certaines conditions de température la SEI peut se dissoudre et créer des sels lithium moins perméables que les ions lithium, ce qui peut augmenter l'impédance de l'électrode négative [37]. Au contraire, de basses températures engendrent une perte

de diffusion du lithium au sein de la SEI et du graphite [38, 39], pouvant recouvrir l'électrode avec un placage de lithium. Il est important de noter que la formation de la SEI, son développement et le placage de lithium sont chacun responsables de la perte de lithium cyclable, dans des conditions d'utilisation mobile [40], et de l'augmentation de la résistance.

1.2.3.2 Effets du vieillissement à l'électrode positive

Il a été démontré à partir d'observations sur l'électrode positive qu'il n'existe pas de modification directe évidente de la morphologie de l'électrode positive, pour tous les niveaux d'Utilisation de la batterie [34, 41]. Ce qui confirme l'importance prépondérante de l'électrode négative dans le processus de dégradation d'une batterie [42].

Il faut cependant remarquer que l'électrode positive est sujette à de faibles altérations avec le temps, dépendantes du matériau choisi [43]. Une SEI est également créée à l'interface électrode positive/électrolyte causée par des tensions élevées à cette électrode [44], cette SEI est néanmoins complexe à détecter [45, 46].

Pour résumer, les principaux phénomènes observés à l'électrode positive sont : l'usure de la masse active, dégradation de l'électrolyte, l'oxydation de l'électrode et la formation d'une SEI, interaction entre les éléments de l'électrode positive dissous à l'intérieur de l'électrolyte et de l'électrode négative [41, 47, 48]. Tous ces effets ne sont pas indépendants et leurs interactions respectives diffèrent en fonction des matériaux employés [49].

1.2.3.3 Conséquences des phénomènes de vieillissement

Nous pouvons identifier deux effets principaux du vieillissement d'une batterie : la perte de capacité et une augmentation de sa résistance interne. Ces deux phénomènes possèdent des causes chimiques différentes. Cela implique une dépendance non linéaire de nature inconnue entre ces différents phénomènes de vieillissement du fait de l'interaction entre les conditions de vie de la batterie.

Afin d'interpréter les phénomènes de dégradations d'une batterie énumérés jusqu'ici, il est essentiel de se placer dans le cadre d'un usage de la batterie dans un véhicule électrique. En terme de performances de la batterie, les pertes de lithium cyclable et la perte de matériel actif provoquent une diminution de sa capacité. De plus, les films passifs génèrent l'augmentation de la résistance de la batterie (figure 1.13).

Lors de l'Utilisation d'un véhicule électrique, ces phénomènes de vieillissement se traduisent pour la perte de capacité par une réduction de l'autonomie disponible, et pour l'augmentation de la résistance par une diminution de la puissance maximale disponible. Cette puissance maximale peut, par exemple, être perçue comme la faculté d'accélération du véhicule. Notons que ce phénomène est moins critique sur les VE car les technologies actuelles sont assez puissantes pour ce type d'application.

1.2.4 Facteurs globaux de vieillissement

Les conditions de vie d'une batterie peuvent se diviser en deux parties distinctes : les charges/décharges et les stockages. Ces conditions correspondent à des modes de

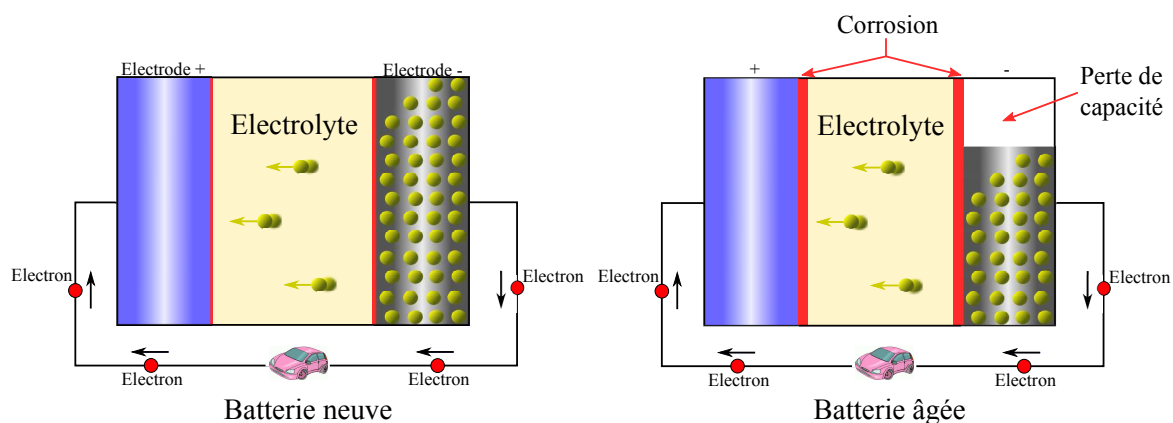


FIGURE 1.13 – Illustration des phénomènes de dégradation d'une batterie

vieillessement différents : calendaire et cyclique [50]. Le vieillissement calendaire correspond aux phénomènes et aux conséquences des stockages d'une batterie.

Au contraire, le vieillissement cyclique est associé à l'impact des périodes d'usages d'une batterie, appelées cycles (charges ou décharges).

1.2.4.1 Vieillessement calendaire

L'auto-décharge d'une batterie correspond à l'énergie perdue lors d'un stockage. Une partie est réversible et l'autre non, impliquant une perte de capacité, et une augmentation de la résistance. Le vieillissement calendaire est principalement défini par la proportion irréversible de la perte de capacité survenant lors d'un stockage d'une batterie. Le taux d'auto-décharge varie en grande proportion selon les conditions de stockage. Les effets générant la dégradation d'une batterie peuvent être accélérés ou ralentis en fonction de ces conditions.

De nombreuses études ont été menées notamment sur le vieillissement calendaire des cellules pour différents niveaux de température, et de SOC [51, 52]. Ces études mettent en avant les conditions de température comme étant la principale contrainte du vieillissement calendaire [53]. Lorsque les températures de stockages sont élevées, les réactions secondaires décrites précédemment et les dégradations sont facilitées ; la perte de lithium est également plus importante que pour des températures modérées, ce qui engendre une perte de capacité [54, 55]. À l'inverse, de faibles températures permettent de limiter le développement de ces phénomènes mais ces conditions induisent néanmoins des problèmes causés par la diminution de diffusion et altèrent la chimie de la batterie [56].

L'autre principale variable représentée dans les études du vieillissement calendaire est le niveau de charge (SOC) auquel la batterie est stockée. Pour des conditions de températures égales, mais pour différents niveaux de SOC, les cellules n'ont pas le même taux de vieillissement (figure 1.14), de plus fortes dégradations se produisant pour des SOC élevés [42]. Par définition, le SOC représente la proportion d'ions présents dans les électrodes. Cela implique directement, pour des niveaux de SOC extrêmes, un fort déséquilibre de potentiel à l'interface électrode/électrolyte, favorisant les réactions chimiques décrites précédemment.

La majorité des études calendaire explorent les conditions de stockage comme une

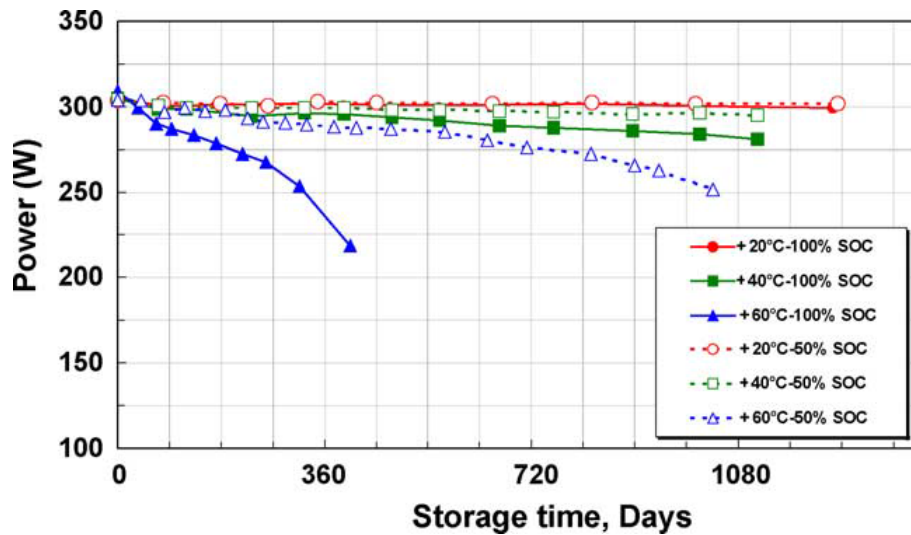


FIGURE 1.14 – Évolution de la puissance de cellules stockées à des températures entre 20°C et 60°C, et avec des SOC entre 50 % et 100 % [57]

combinaison entre les SOC et les conditions de températures. Or, ces variables altèrent simultanément la batterie, chacune avec un effet non-linéaire avec le temps, indépendamment ou non des autres variables. Les résultats des études menées évoquent un effet plus restrictif d'un haut niveau de SOC par rapport à une température élevée [57, 58]. Ces conclusions sont uniquement faites à partir de l'observation d'expérimentations menées et il demeure nécessaire de comprendre les effets de la combinaison de ces deux causes du vieillissement calendaire.

1.2.4.2 Vieillissement cyclique

Le vieillissement cyclique intervient, lui, lorsque la batterie est en charge ou en décharge. Il s'agit d'une conséquence directe du niveau et du mode d'Utilisation, et donc des températures et des sollicitations appliquées à la batterie. De nombreux facteurs sont impliqués dans ce type de vieillissement.

Lors de ces périodes, une batterie est sujette à des effets exothermiques [59] qui peuvent être privilégiés sous des températures élevées et provoquer un vieillissement prématuré. Il est cependant important de considérer également les effets des températures basses. Des études rapportent un impact direct de la température ambiante sur le vieillissement d'une batterie, mais aucune ne se porte sur la température mesurée dans la batterie directement. Cette notion reste mal comprise et requiert des investigations plus approfondies.

Mis à part ces variables, les facteurs de vieillissement cyclique dépendent du mode de cyclage de la batterie. Un facteur récurrent dans la littérature est le ΔSOC , qui représente la variation de l'état de charge se produisant lors d'un cycle. Ce facteur est primordial car il correspond à la quantité de charge fournie (resp. accumulée) par la batterie durant une décharge (resp. une charge). Des expérimentations étudiant des cellules utilisées sous des températures et des SOC initiaux identiques, pour divers ΔSOC , ont démontré une perte de puissance de la batterie lorsque la valeur de ΔSOC est élevée ($> 70\%$) [54]. De tels phénomènes sont principalement induits par la dé-

gradation de l'électrode positive et du développement de sa SEI, provoqués par les régimes élevés de charges ou de décharges.

Une autre variable impactant le vieillissement des batteries Li-ion est la tension de charge employée. Des charges sous haute tension impliquent une accélération des phénomènes de dégradation [60]. Afin d'illustrer cela, une étude montre que la durée de vie d'une batterie est réduite de moitié lorsque la tension de charge est augmentée de 0.1 V (figure 1.15). Notons que la tension de décharge influence également le processus de vieillissement à travers une augmentation de l'impédance de la batterie [61, 48].

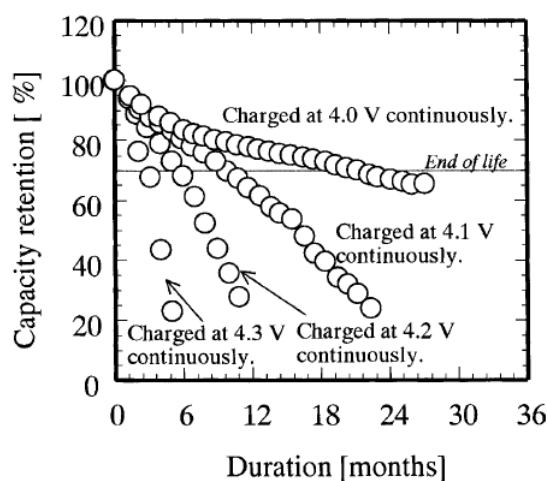


FIGURE 1.15 – Évolution de la capacité en fonction de différentes tension de charges, avec une température de 45°C [62]

Enfin, les pics de courant demandés semblent être une notion impliquée dans les dégradations subies par une batterie. Un pic de courant important génère un haut niveau d'énergie fourni ou reçu par la batterie, ce qui peut provoquer une accélération du vieillissement. Cependant, cette variable n'a pas été étudiée précisément dans la littérature actuelle, car a-priori ponctuelle lors de conduite simples supposées par les constructeurs.

1.2.4.3 Conclusion des facteurs de vieillissement

L'ensemble des facteurs présentés impactent le vieillissement d'une batterie de manière couplée pour générer à la fois une augmentation de la résistance et une perte de capacité. Notons que la plupart de ces facteurs sont dépendants des conditions externes. Dans le cadre d'une application concernant les transports, la température externe dépend du climat, tandis que le ΔSOC est, lui, une conséquence du comportement de l'utilisateur³. Le mode d'Utilisation de la batterie est très largement impliqué dans les études de vieillissement. Cela peut se traduire par exemple par le niveau de température d'une batterie qui est dépendant du profil de vitesse appliqué ou bien selon le dénivelé de la route utilisée.

De nombreuses interactions entre diverses variables interviennent dans les processus de dégradation, provenant à la fois de l'environnement et du mode d'Utilisation

3. Le ΔSOC est dépendant des charges et des trajets effectués.

de la batterie [63, 64]. Toutes ces caractéristiques des phénomènes de vieillissement d'une batterie, font de l'analyse et de la compréhension du vieillissement un challenge important restant à être relevé.

1.3 Estimation du vieillissement

1.3.1 Indicateurs de vieillissement

Afin d'évaluer et de quantifier le niveau de vieillissement d'une batterie, plusieurs indicateurs ou notions ont été créés. L'indicateur le plus utilisé dans la littérature est le State Of Health (SOH), usuellement défini à l'instant t par :

$$\text{SOH}(t) = \frac{\text{capacité nominale à } t}{\text{capacité initiale à } t_0} (\%) \quad (1.8)$$

Notons que d'autres définitions du SOH ont été proposées via le critère EOL, mais demeurent une proportion de capacité restante. Cet indicateur illustre et mesure la perte de capacité survenue sur une batterie étudiée. Les mesures de capacité doivent donc toujours être faites à une même température de référence pour obtenir une valeur fiable.

Cet indicateur SOH mesure la proportion de capacité restante d'une batterie. Ces définitions sont déduites à partir du niveau de capacité de la batterie, et par conséquent ne prennent pas en compte toutes les formes de vieillissement explicitées dans ce chapitre ; la résistance est par exemple omise. Bien que moins contraignante que la capacité dans le cadre d'un usage en VE, cette notion peut être considérée par le biais d'autres indicateurs. Il peut, par exemple, être possible de modifier le SOH défini en équation 1.8 pour calculer la proportion de la résistance actuelle par rapport à la résistance initiale. Il est important de noter que la croissance de la résistance est spécifiquement impactante dans le cadre d'applications impliquant de fortes puissances, car pouvant réduire l'accélération maximale d'un VE.

Un autre indicateur a été introduit pour quantifier l'effet des dégradations d'une batterie sur ces performances [65], il est appelé le State Of Function (SOF) et est défini par :

$$\text{SOF}(t) = \frac{|P(T, \text{SOH}, \text{SOC}) - P_{\text{cible}}|}{|P_0(T, \text{SOH}, \text{SOC}) - P_{\text{cible}}|} \quad (1.9)$$

avec P_{cible} la puissance cible désirée, $P(T, \text{SOH}, \text{SOC})$ la puissance disponible pour la conditions précisées et $P_0(T, \text{SOH}, \text{SOC})$ la puissance que peut fournir la batterie au temps t_0 . La figure 1.16 illustre le fonctionnement de cet indicateur. Cet indicateur SOF est destiné à quantifier le niveau de performance d'une batterie en fonction de son état (SOH et SOC) et des conditions externes (température), permettant de connaître les performances optimales de la batterie à un instant donné. L'intérêt du SOF est de pondérer le SOC d'une batterie par son état réel en terme de performances maximales. Il faut cependant remarquer que cet indicateur SOF n'est que très rarement employé dans la littérature concernant l'étude et l'estimation du vieillissement des batteries, sa pertinence étant ciblée pour l'identification des performances maximales en fonction des conditions d'usage.

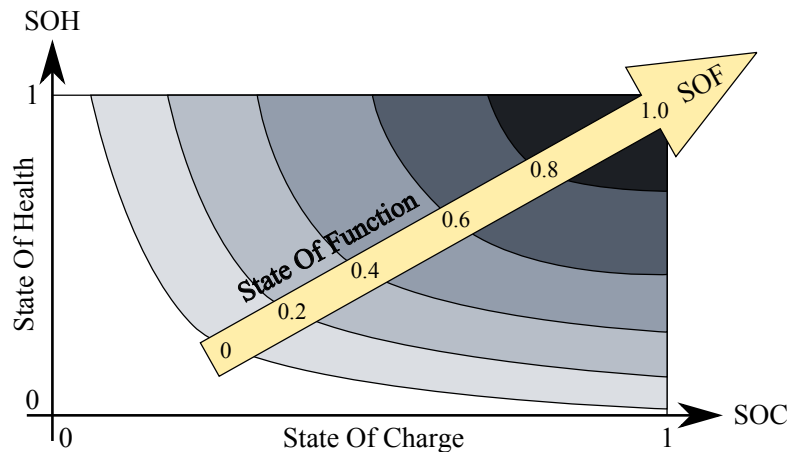


FIGURE 1.16 – Relation entre l'état de fonctionnement (SOF) pour la décharge en fonction de l'état de charge (SOC) et de l'état de santé (SOH) à une température constante [65]

Comme le terme de *vieillessement* n'est pas formellement défini, la notion de Remaining Useful Life (RUL) est fréquemment utilisée. Ce terme RUL correspond à la période restante avant le franchissement du critère de fin de vie (EOL). Le RUL quantifie une période de temps ou un nombre de cycles restants avant la fin de vie de la batterie dans son application. Notons que certaines études définissent le critère EOL comme l'instant correspondant à une valeur initiale de résistance doublée, sans prendre en compte la capacité [66].

1.3.2 Modélisation des phénomènes de vieillissement

Les phénomènes de vieillissement étant complexes à définir, de nombreuses et diverses études ont été effectuées. En dehors des mesures directes, différentes méthodes sont utilisées pour estimer ces notions de niveau de dégradation d'une batterie. Nous pouvons les diviser en trois parties distinctes :

- Modèles électrochimiques : les phénomènes se produisant au sein de la batterie sont détaillés et modélisés
- Modèles de circuit équivalent : la batterie est réduite à un circuit électrique équivalent
- Modèles semi-empiriques : le vieillissement est modélisé par des équations physiques
- Modèles analytiques : des paramètres de vieillissement sont fixés à partir des mesures
- Méthodes statistiques : approche ne considérant pas de connaissance a-priori, le vieillissement est estimé à partir des données d'expérimentations

1.3.2.1 Modèles électrochimiques

Approches phénoménologiques

Le développement des modèles électrochimiques a débuté depuis environ vingt ans

avec les travaux des équipes de Newman [67, 68], qui cherchait à estimer les performances d'une batterie à partir d'équations de Butler-Volmer, d'après la théorie des électrodes poreuses [69]. La première tentative de simulation d'un processus de vieillissement par un modèle physique a été faite par Darling et Newman [70], par la mise en œuvre d'une simple réaction d'oxydation solvant dans leur modèle. Ces résultats encourageants ont permis d'établir des conclusions concernant le choix du solvant. Sur la base de cette approche, Christensen et al. [71] explique la diminution de la capacité d'une batterie par l'augmentation de la résistance sur la surface de l'électrode négative ; un modèle de représentation de la SEI a alors été développé [72].

La plupart des mécanismes de vieillissement incorporés dans les modèles physiques sont liés à l'électrode négative. Ces modèles ont étudié l'impact des contraintes se produisant au cours du processus d'intercalation de lithium dans les batteries Lithium Manganèse Oxyde (LMO), réussissant à identifier les courants électriques, et des niveaux de SOC, causant l'apparition de fractures dans les particules de matière active [73].

Le développement de la SEI et sa relation avec la perte de capacité ont également été analysés avec un modèle de diffusion de solvants [74]. Plus récemment, des études ont estimé le coefficient de diffusion effective au sein de l'électrolyte avec le nombre de cycles [75, 76]. Safari et al. [77, 78] ont développé un modèle phénoménologique d'estimation des performances de cellules LFP/graphite commercialisées, tentant de quantifier l'impact du vieillissement d'une batterie sur ses performances en termes de capacité et d'impédance. Cette étude a permis d'évaluer le cyclage comme étant plus déterminant que le stockage dans le procédé de dégradation d'une batterie. D'autre part, la perte de matière active apparaît comme moins importante que la perte de lithium actif pour ce type de batteries, pour des températures entre 25°C et 45°C [79].

Approches atomistiques et moléculaires

Dans les approches proposées ci-dessus, les propriétés intrinsèques associées aux électrodes et aux électrolytes ne sont pas prises en compte avec une grande précision. Afin de comprendre à une échelle nanoscopique les phénomènes de vieillissement intervenant dans la batterie, quelques modèles ont été introduits utilisant des méthodes atomistiques et moléculaires [80].

La *Density Functional Theory* (DFT) est employée pour la description des phénomènes ayant lieu lors de la lithiation/délithiation de ces électrodes. Des fondements théoriques ont été développés pour approfondir cette phase de transition. Dalverny et al. [81] est ainsi parvenu à déterminer la morphologie des électrodes LiCoO_2 , à partir de calculs ab initio. Une autre étude a examiné cette phase de transition par le biais de particules LFP actives [82]. Les calculs ab initio contribuent à la compréhension des phénomènes de dégradation par l'estimation des énergies associées à la décomposition d'un solvant ou bien à la dissolution de sels lithium. Par ce procédé, Tasaki et al. [83, 84] a établi une relation entre la dissolution de sels lithium se produisant proche de la SEI de l'électrode négative, et la perte de capacité calendaire.

La dynamique moléculaire représente aussi une solution intéressante dans la compréhension des mécanismes de dégradation. Ainsi, une étude [85] a identifié la décomposition du solvant Ethylene Carbonate (EC) comme une conséquence de la croissance de la SEI. Cette propriété peut être une des clés de la compréhension des problèmes

de vieillissement reportés avec des électrodes graphite. De plus, l'évolution de la SEI a été examinée durant d'autres études [36], par l'usage de méthode *Kinetic Monte Carlo* (KMC), se focalisant sur la diminution de la surface active sur l'électrode par rapport à la formation de SEI lors des cycles.

Un des plus grands challenges de la modélisation physique reste la mise en relation entre les approches atomistiques (DFT, dynamiques moléculaires...) et les modèles macroscopiques. L'objectif de ce type d'approche est de parvenir à décrire de manière détaillée tous les phénomènes se produisant à l'échelle nanoscopique et donc de prédire directement leurs impacts sur les performances de la batterie. Cet objectif ambitieux semble accessible, mais les problématiques persistantes sont liées à l'association de modèles *top-down*, allant de l'observation macroscopique à la description nanoscopique de la batterie, avec des modèles *bottom-up* [86].

1.3.2.2 Modèles de circuit équivalent

D'autres méthodologies fournissent des estimations de vieillissement, par le biais de modèles de circuit équivalent. Cette méthode de modélisation tente de représenter la batterie, par un ensemble de composants électriques, dont le comportement est cohérent avec le fonctionnement réel, puis utilise différentes techniques d'identification pour estimer les paramètres de ce modèle [87, 88]. Il faut cependant noter que chaque étude définit et utilise un circuit équivalent qui lui est propre. Un exemple de ce type de circuit est présenté en figure 1.17.

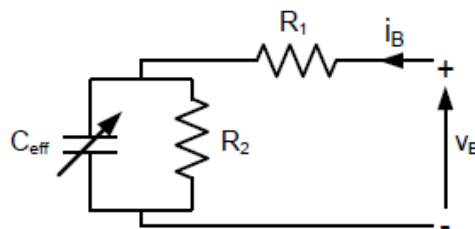


FIGURE 1.17 – Exemple de circuit équivalent proposé [89]

Les paramètres de ces modèles de vieillissement, sont issus directement de paramètres physiques de la batterie, ou bien par le biais d'approches plus complexes [90, 91]. De ce fait, ces études nécessitent l'identification de paramètres de circuits équivalents associés aux cyclages puis aux stockages, pouvant introduire du biais. Une telle méthodologie requiert l'acquisition d'une grande quantité de données diverses, résultant de longues expérimentations. Par exemple, une méthode d'apprentissage, *Relevance Vector Machine* (RVM), est fréquemment employée lors de l'estimation des paramètres de circuits équivalent. Cette méthode est une forme Bayésienne représentant une généralisation de l'expression linéaire du *Support Vector Machine* (SVM) introduit par Vapnik [92], actuellement l'état de l'art des algorithmes de régression et de classification [93, 94]. La méthode RVM estime différents paramètres du circuit à partir de très larges bases de données, afin de prédire le SOH d'une batterie. Saha et al. [95] ont mis en place cette approche d'apprentissage des paramètres d'un modèle de batterie estimant le SOH de la batterie étudiée. Ce type d'approche utilise fréquemment des *Filtre Particulaire* (PF) ou bien des *Rao-Blackwellized Particle Filter* (RBPF) pouvant être

appliqués pour l'obtention, et le suivi dans le temps, d'une distribution de prédiction de SOH ou bien de RUL [96].

1.3.2.3 Modèles semi-empiriques

Les modèles semi-empiriques utilisent, quant à eux, de simples corrélations entre les facteurs de dégradation et la diminution de capacité et/ou l'augmentation de résistance. Ces corrélations sont introduites à partir de tests de vieillissement effectués sous différentes conditions. L'objectif de cette approche est de quantifier directement l'impact des facteurs de vieillissement pour obtenir une expression descriptive du niveau de performance de la batterie, en fonction des conditions de vie qu'elle a subi. Le plus souvent, ces études traitent indépendamment le vieillissement calendaire et le vieillissement cyclique.

Vieillissement calendaire

Dans la modélisation du vieillissement calendaire, les principales variables sont le temps, la température et le SOC initial de stockage. Il est intéressant de constater que certaines études ne prennent cependant pas en compte cette dernière variable lors de la modélisation calendaire [97]. Or, comme cela a été détaillé dans la section 1.2.4, l'influence du SOC ne peut être négligée du fait de son impact important sous certaines conditions. Le vieillissement calendaire semble suivre une cinétique de type Arrhénius (équation 1.10), cette loi étant usuellement choisie pour prendre en compte l'influence des températures de stockage [30, 54, 98]. Cette approche considère la vitesse de dégradation d'une batterie comme principalement causée par un processus lié à la température et au temps [99].

$$v = A \cdot t \exp\left(\frac{-E_a}{RT}\right), \quad (1.10)$$

avec v la vitesse de dégradation à déterminer, t le temps, A une constante, R la constante des gaz parfait, T la température en K et E_a l'énergie d'activation.

Le SOC de stockage est pris en compte dans certaines études, par exemple dans l'équation de Tafel correspondant à une réaction électrochimique de surtension (équation 1.11) :

$$\Delta V = A \ln\left(\frac{i}{i_0}\right), \quad (1.11)$$

avec ΔV la surtension, A la pente de Tafel, i la densité de courant de la réaction électrochimique et i_0 la densité de courant échangé.

Dans la plupart des études, la perte de capacité due au stockage s'avère être linéairement dépendante de la racine carrée du temps [54, 100, 101]. À titre d'exemple, Ecker et al. [66] ont proposé un modèle d'estimation de durée de vie calendaire décrivant les évolutions de capacité et de résistance en fonction du temps, d'après l'équation 1.12 :

$$\frac{L(T, V, t)}{L(T_0, V_0, t_0)} = 1 + B(T, V) \cdot c_a \cdot t^{1/2} \quad (1.12)$$

avec c_a un coefficient du taux de dégradation (diminution de capacité ou augmentation de la résistance) sous des conditions références : T_0 , t_0 et V_0 . La variable $L(T, V, t)$ cor-

respond soit à la résistance soit à la capacité estimée au temps t , pour une température T et une tension V . De plus, l'impact de la température et de la tension de stockage est calculé suivant une relation de dépendance :

$$B(T, V) = c_T^{\frac{T-T_0}{\Delta T}} \cdot c_V^{\frac{V-V_0}{\Delta V}} \quad (1.13)$$

avec T_0 et V_0 des références de température et de tension. ΔT et ΔV sont arbitrairement définis comme étant respectivement 10°C et 0.1 V . Les paramètres c_T et c_V de loi proposée (1.13) sont déterminés à partir d'expérimentations de vieillissement calendaire accélérés.

D'après ces expérimentations, il existe une hypothétique contribution du nombre de cycles dans la part de vieillissement calendaire. Les résultats simulés sont, dans ce cas, en accord avec la diminution de capacité mesurée lors des cyclages, mais ont néanmoins tendance à surestimer substantiellement la résistance au fil du temps. Cependant, une approche basée sur des simulations ou sur des tests accélérés ne permet pas d'obtenir un modèle fiable sur les plages de fonctionnement non observées lors des expérimentations.

Vieillissement cyclique

Le vieillissement cyclique est, lui, plus complexe à estimer car il englobe plus de variables indépendantes telles que la température ou la tension mesurée. D'autre part, ces variables sont reliées aux conditions externes d'Utilisation de la batterie. Dans cette approche, les principaux facteurs considérés sont la température, le nombre de cycles, la valeur de ΔSOC , la tension. Notons que la majorité des études sont en accord sur le fait que le nombre de cycle est comparable à une notion de temps de vie. Ainsi, Wenzl et al. [102] évoque la possibilité d'une simple méthode de comptage, cependant ce choix ne semble pas pertinent par manque de prise en compte d'autres facteurs d'Utilisation.

Par exemple, une étude propose une dépendance entre le nombre total de cycles et le vieillissement cyclique [100], tandis qu'une autre estime que les paramètres d'un modèle suivent une dépendance en racine carrée du nombre de cycles [103]. De l'autre côté, Wang et al. [104] ont développé un modèle de vieillissement cyclique de cellules LiFePO_4 en considérant les variables : ΔSOC , température et C-rate. Cette dernière met en évidence une dépendance en $Ah^{0.55}$ de la perte de capacité, où Ah est la charge cumulée délivrée par la batterie tout au long de sa vie. D'après les auteurs, ce paramètre est directement proportionnel au temps de vieillissement et est corrélé aux dégradations pour différents C-rate.

Dans le cas de modèles calendaires et cycliques séparés, les deux approches de vieillissement sont regroupés pour obtenir une expression générale du vieillissement d'une batterie.

Vieillissement global

Une autre méthodologie consiste à estimer le niveau de vieillissement directement sans dissocier les approches calendaires et cycliques. Il peut ici s'agir du même type de modèle que ceux détaillés précédemment, en considérant le vieillissement d'une

manière globale [105, 106], ou sous la forme d'un modèle d'accumulation des stress (approche de fatigue des matériaux). Ces modèles de fatigue utilisent une corrélation entre les facteurs de stress et la perte de capacité dans le but d'identifier la perte de capacité dans le temps. Les modèles de type Palmgren-Miner sont un exemple de modèle d'accumulation des stress [107, 108].

Afin de mettre en œuvre ce type de méthodologie, un procédé de segmentation d'un profil est nécessaire pour décomposer un profil de courant en segments compatibles avec les corrélations empiriques choisies. Safari et al. [109] fournit une description détaillée du fonctionnement de cette méthodologie.

Le principal inconvénient des modèles semi-empiriques réside dans le fait que ceux-ci ne fournissent aucun renseignement sur le processus contribuant aux dégradations de la batterie. Par conséquent, l'impact de chaque facteur de stress doit être étudié de façon indépendante par le biais d'expérimentations exhaustives. En outre, l'extrapolation des résultats de tests accélérés peut conduire à d'importantes erreurs lors de l'application en conditions réelles, comme illustré par Takei et al. [110].

1.3.2.4 Modèles analytiques

Les méthodes analytiques sont appliquées uniquement sur les données, aussi nombreuses que possible, obtenues lors d'expérimentations afin d'évaluer ou de prédire la valeur d'indicateurs. Ces méthodes sont utilisées à la fois pour déterminer des paramètres de modèles, et pour estimer directement un indicateur de vieillissement. Les principales limites de cette méthodologie proviennent de la nécessité d'avoir une grande plage de données expérimentales et de possibles erreurs de mesures [111].

La méthode la plus utilisée est la *coulométrie* qui permet d'estimer le SOH d'une batterie par simple intégration du courant dans le temps [112]. Il est important de noter ici que ce procédé doit toujours être effectué dans les mêmes conditions expérimentales, telles que la température. Pour obtenir des estimations de SOH, la *coulométrie* requiert en outre une recalibration fréquente qui ne peut pas être faite en temps réel dans le contexte d'un usage VE [113].

La logique floue est également employée en admettant un faible niveau de bruit dans les données [114]. À partir des données, des règles de logique floue fixent une relation entrée-sortie basée sur des experts, l'estimation de paramètres de vieillissement est réalisée directement ou par le biais d'un modèle annexe⁴ [18, 115]. Par exemple, Salkind et al. [116] utilise cette méthode pour l'estimation du SOH de batteries NiMH. Ou encore, Singh et al. [117] se base sur des données d'*Electrochemical Impedance Spectroscopy (EIS)* pour identifier deux paramètres : l'ampleur de l'impédance et les angles entre phases. Cette méthode prédit le nombre de cycles restant avant la fin de vie de la batterie étudiée, mais demeure extrêmement complexe à mettre en œuvre, car nécessitant de nombreuses mesures de EIS qui nécessitent un protocole expérimental important.

Une autre méthode analytique est basée sur les états observés, modélisant le problème d'estimation du vieillissement par un système d'équations S (équation 1.14), avec une entrée u (vecteur d'état) et une sortie y (tension), dépendante de variables

4. Il est important de noter que le processus des méthodes de logique floue est susceptible d'ajouter des erreurs considérables à partir des hypothèses faites par les experts.

cachées x .

$$S = \begin{cases} \hat{x} = Ax + Bu \\ y = Cx \end{cases} \quad (1.14)$$

L'objectif est par conséquent d'ajuster un modèle à partir d'observations pour minimiser l'erreur entre \hat{x} et y via un gain K . La correction du gain K est fixée par un algorithme, le plus souvent un Filtre Particulaire (PF) [118], et plus précisément un Extended Kalman Filter (EKF) [119, 120]. La figure 1.18 illustre ce procédé. Le principal problème de cette approche provient de la possible divergence rapide sous certaines conditions non propices, en effet il est nécessaire de prédéfinir une matrice initiale pouvant induire ce type d'erreur [121].

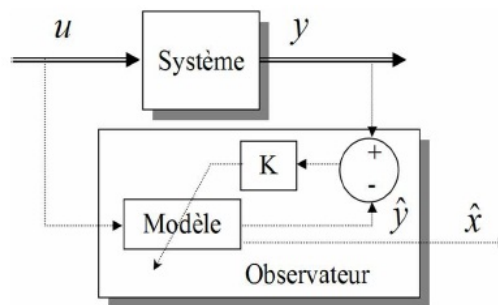


FIGURE 1.18 – Principe d'observation consistant à estimer les sorties \hat{y} d'un système à partir de son modèle [122]

1.3.2.5 Méthodes statistiques

Comme les modèles analytiques, les méthodes statistiques nécessitent de larges bases de données exhaustives pour être performantes. Ces méthodes ne demandent par contre aucune connaissance a-priori sur les mécanismes de vieillissement, elles ne font ainsi aucune hypothèse sur les facteurs de dégradation et ne considèrent pas les mécanismes physico-chimiques de la batterie.

Les Artificial Neural Networks (ANN) ou Neural Networks (NN) sont principalement utilisés dans le cadre de la prédiction du SOC, mais ces méthodes sont quelques fois employées dans l'estimation du SOH d'une batterie. L'avantage est ici la faible complexité des variables requises dans la mise en place de la méthodologie. Il est important de remarquer que les NN restent performants malgré les multi-dépendances présentes, et affirment ainsi leur capacité d'adaptation à toute technologie de batterie [123]. Cependant, comme tout mécanisme d'apprentissage, les NN ont besoin d'une grande variété de données pour être performants.

Une méthode simplement implémentable consiste en l'utilisation de procédés de séries temporelles, telle que l'Auto-Regressive Moving Average (ARMA) [124]. Cette méthodologie considère le niveau de vieillissement comme une série chronologique et la méthode ARMA permet d'en déduire la prochaine valeur [125]. Une telle méthode ne peut donc fonctionner que sur une unique batterie du fait de l'importance primordiale des conditions externes et des types d'usages effectués. De plus, la précision résultante est très largement dépendante des données, ce qui demande un grand

nombre de caractérisations de la batterie permettant d'obtenir des références de niveaux de vieillissement régulièrement. Ces inconvénients font de ce type d'approche, des issues très peu réalisables en pratique dans un contexte d'application en embarqué sur un véhicule électrique.

Une autre approche basée sur les données considère le critère de fin de vie d'une batterie comme une panne et modélise la probabilité de cette panne par une loi de Weibull [126, 127]. Cependant, cette méthode considère tous les types d'Usages et de conditions d'expérimentation comme unique, ce qui réduit considérablement la précision des résultats obtenues.

Les méthodes précédemment évoquées telles que NN, Fuzzy Logic et RVM peuvent également être mises en place par l'estimation directe d'un paramètre de vieillissement, sans passer par un modèle défini a-priori.

1.3.3 Bilan de l'estimation des phénomènes de vieillissement

Cette section discute des caractéristiques de chacune des méthodes d'estimation des phénomènes vieillissement, et présente un comparatif de leurs spécificités.

1.3.3.1 Évaluation des méthodes d'estimations

Depuis les mesures directes aux estimations statistiques, de nombreux domaines abordent l'estimation du niveau de vieillissement d'une batterie. Les mesures directes ne requièrent aucune hypothèse car il s'agit d'une mesure directement effectuée sur la batterie. De plus, cette méthodologie fonctionne pour tous types de batteries et pour toutes les utilisations. Cependant, la durée d'immobilisation et les coûts induits par ce procédé rendent impossible l'application dans une approche en temps réel sur un véhicule. Elle ne permet pas de prédire le comportement futur de la batterie. Il peut être envisagé de considérer l'application de telles méthodes au sein de garages spécifiques.

Les méthodes relatives aux modèles électrochimiques et aux circuits équivalents proposent des précisions intéressantes, mais elles ne peuvent pas être directement étendues à d'autres technologies de batterie (matériau, *design*), car dépendantes des caractéristiques propres aux batteries. Ce point correspond à l'inconvénient majeur de ces méthodes les technologies étant en constante évolution. Ces deux approches ne modélisent pas tous les mécanismes de dégradation intervenant au sein d'une batterie, mais seulement la tendance globale de vieillissement. Cependant, les modèles électrochimiques et physiques sont des outils puissants pour la compréhension des différentes et complexes interactions qui se produisent entre différents phénomènes internes. De la même manière, l'approche basée sur les circuits équivalents ne permet pas de modéliser tous les mécanismes de dégradation intervenant au sein d'une batterie, mais la tendance de vieillissement globale dans les meilleurs des cas.

Ces remarques sont également valables pour les modèles semi-empiriques. En effet le vieillissement est complexe et dépendant de nombreux phénomènes. Cette approche est aussi sensible à la technologie de batterie considérée. Les modèles étant bâtis sur des tests simulés sur banc et sur seulement quelques cellules, les études spécifiques présentent généralement de bons résultats, mais les algorithmes ne sont que très rarement appliqués sur des batteries en usage réel sur un véhicule. Dans ce cas l'intervention des conditions externes induirait de grossières erreurs d'estimation.

Les méthodes statistiques sont, elles, très facilement adaptables à d'autres technologies de batteries, et permettent un diagnostic d'état de santé en temps réel. Mais ces approches nécessitent une très large variété de données pour être précises. Cette collecte de données est une tâche difficile et fastidieuse car la durée de vie des batteries en usages réels est longue (> 5 ans). L'enregistrement de données représentatives de batteries lors d'usages réels est un problème économique et temporel. Dans le cas idéal, un grand nombre de batteries utilisées sur des véhicules, avec des usages réels divers, permettrait la mise en place de modèles statistiques fiables. Pour résumer, l'inconvénient primordial de cette approche vient du processus de collecte des données.

Comme nous l'avons détaillé, aucune des méthodes actuelles n'est suffisamment performante pour obtenir une estimation du vieillissement capable d'être mise en place dans le cadre de diagnostics réels embarqués sur un véhicule en usage réel. La figure 1.19 représente les performances des méthodes décrites, permettant l'estimation du niveau de vieillissement d'une batterie, selon différents critères tels que la capacité de prédiction, le fonctionnement en temps réel ou bien la précision. Les différentes méthodes s'avèrent complémentaires, et du fait de la nature complexe des phénomènes de dégradation, il semble nécessaire d'établir une méthode utilisant plusieurs de ces domaines. La figure 1.19 illustre le compromis actuel auquel se confronte la problématique d'estimation du niveau de vieillissement d'une batterie.

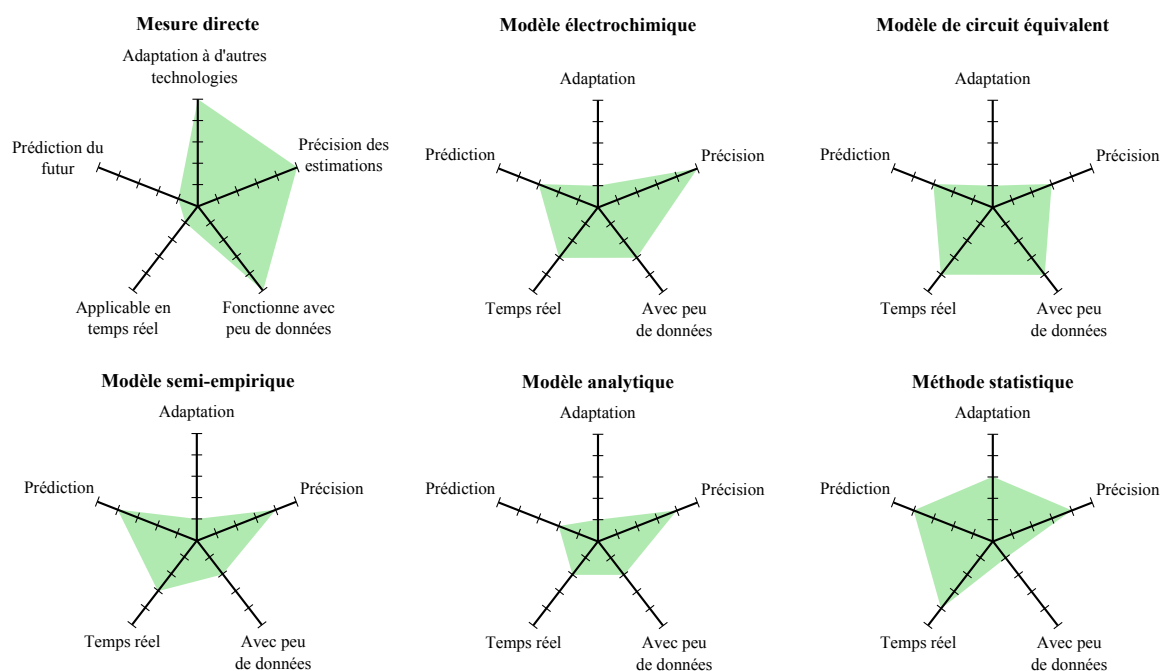


FIGURE 1.19 – Comparaison des spécificités des principales méthodes utilisées dans l'estimation du niveau de vieillissement d'une batterie [35]

1.3.3.2 Bilan des méthodes décrites

Le développement actuel du marché des VE implique de hautes attentes concernant les futures performances de ceux-ci, et par voie de conséquence sur le coût des batteries. L'évolution désirée de ces performances nécessite une durée de vie plus im-

portante des batteries. Ces objectifs requièrent une bonne compréhension et estimation des phénomènes de dégradation, ce qui reste à ce jour un défi complexe.

Le vieillissement des batteries est difficile à identifier et à quantifier à cause de la diversité et la complexité des phénomènes ayant lieu au sein de la batterie durant toute sa vie. Les deux principaux effets globaux du vieillissement, la perte de capacité et l'augmentation de la résistance, sont distincts car provoqués par des mécanismes et des causes différentes, mais couplés.

De nombreuses variables sont donc impliquées dans le processus de vieillissement, ayant chacune un rôle différent. Les facteurs les plus récurrents dans le vieillissement calendaire sont : la température, le temps et le SOC de stockage. Le vieillissement cyclique concerne quant à lui les variables suivantes : température, ΔSOC , nombre de cycles, tension de charge/décharge, et d'autres facteurs issus des conditions d'Usages qui ne sont pas encore clairement identifiés, tels que les pics de courant. Ces variables sont issues des conditions internes et externes, et interagissent entre elles. Toutes ces interactions sont les raisons principales de la difficulté de la compréhension et d'estimation du vieillissement.

Les études basées sur des expérimentations sur banc, ce qui est le cas de la majorité des études, ne permettent pas de prendre en compte l'impact de toutes les conditions d'usage. Par exemple, Bögel et al. [53] a tenté de reproduire un usage VE réel sur un banc de test, mais les faibles résultats obtenus illustrent parfaitement les limites de cette approche. Ces simulations amènent parfois les résultats à être linéairement dépendants de quelques variables. De telles observations permettent généralement de construire différents modèles. Il est donc important de considérer le biais initialement introduit par les expérimentations faites sur banc de test.

1.4 Conclusion des problématiques liées à l'estimation du vieillissement d'une batterie

1.4.1 Point de vue utilisateur sur le vieillissement

L'étude présentée dans ce chapitre a mis en évidence l'importance du diagnostic des phénomènes de vieillissement d'une batterie en usage réel sur un véhicule. Cela se traduit concrètement par l'obtention d'un indicateur fiable de l'état de santé de la batterie, mis à jour en temps réel, sans contrainte particulière. Or, nous avons vu qu'il n'existait à ce jour aucune méthodologie permettant de satisfaire tous ces critères.

Il est important de considérer les attentes précises des utilisateurs de VE en terme de vieillissement de batterie. En effet, comme nous l'avons évoqué, les consommateurs guident le marché et les constructeurs axent leurs développements afin de satisfaire ces attentes. Le développement d'outils et de méthodologies doivent donc chercher à répondre à des besoins précis d'utilisateurs pour être pertinents et utiles à l'expansion du marché des VE.

Le marché des utilisateurs des VE est divisé en deux parties distinctes : les propriétaires individuels de véhicules, et les gestionnaires de flottes de véhicules. Dans les deux cas certaines besoins sont communs, cependant les moyens d'actions diffèrent. Nous analysons ici les besoins spécifiques, non répondus à ce jour, de ces deux groupes de consommateurs, concernant le vieillissement de batterie en usage réel sur un VE.

Cette identification des attentes permettra alors d'axer, par la suite, les développements autour de ces besoins spécifiques.

1.4.1.1 Utilisation d'un VE personnel

La majorité des propriétaires de véhicules (électriques ou non) sont des utilisateurs qui achètent et qui utilisent leur véhicule dans un cadre privé. Cette partie du marché possède des caractéristiques particulières. En effet, dans ce cas, les véhicules sont stockés la très grande majorité du temps (~ 92 % de sa vie). Ces véhicules sont prioritairement destinés à des petits trajets quotidiens, le plus souvent des parcours domicile - lieu de travail. Les distances parcourues par jour sont faibles (~ 46 km) et les performances de ces VE doivent répondre à ces besoins quotidiens (section 1.1.2).

Cette part de marché ne devrait en théorie pas être affectée fortement par l'autonomie d'un VE, qui satisfait la majorité des besoins actuels. Or, comme nous l'avons décrit dans la section 1.1.2, le facteur de *range anxiety* est très présent chez les consommateurs. Il apparaît donc primordial d'obtenir un indicateur fiable du niveau de l'autonomie maximale d'une batterie. Cet indicateur doit fournir à un conducteur une information précise sur sa capacité à rejoindre la destination souhaitée, prenant en compte la topologie de la route impliquant des consommations d'énergie différentes. Du point de vue du vieillissement de la batterie, il est donc nécessaire d'obtenir un indicateur du SOH de la batterie en temps réel, précis, lors des usages du VE étudié. Celui-ci permettra par voie de conséquence de fiabiliser les indicateurs de SOC.

L'autre considération importante pour les utilisateurs privés, réside dans la compréhension des phénomènes de vieillissement. En effet, les consommateurs actuels ne sont pas ou très peu sensibilisés aux contraintes provoquant des dégradations accélérées de leurs batteries. Ce manque de sensibilisation des utilisateurs se vérifie par les usages des batteries de téléphones portables. Une identification des conditions d'utilisation provoquant des vieillissements et la quantification de leurs impacts est donc nécessaire. Cela peut éviter certains usages contraignants pour la batterie et donc de la préserver de dégradations précoces.

1.4.1.2 Gestionnaire d'une flotte de VE

Les propriétaires de flottes, principalement des entreprises ou des loueurs, possèdent une flotte de véhicules utilisés par ses employés. Il peut également s'agir de services de locations de véhicules (électriques ou non). Cette clientèle nécessite le développement d'un outil de diagnostic de l'état de santé d'une batterie fiable et applicable en ligne. En effet, les conducteurs de ces véhicules sont soumis eux aussi à une inquiétude liée à l'autonomie du véhicule qu'ils utilisent, d'autant plus qu'ils ne connaissent pas ce véhicule et le biais associé aux indicateurs, contrairement aux utilisations personnelles.

Les gestionnaires de flottes sont sensibles à l'optimisation d'utilisation de celles-ci, et aux questions de maintenance. L'obtention d'un indicateur précis de vieillissement d'une batterie implique la question de l'optimisation de la flotte afin de maximiser la durée de vie des VE la composant. Cette gestion optimale d'une flotte n'est donc possible qu'avec la connaissance précise des facteurs ayant provoqués les dégradations estimées des batteries. La connaissance détaillée des variables contraignantes pour chacun des véhicules peut permettre à un gestionnaire de gérer sa flotte au mieux,

proposant un outil pour l'optimisation de l'utilisation des VE et des batteries, engendrant une réduction des coûts d'exploitation.

1.4.2 Bilan des solutions existantes

Les batteries lithium-ion sont la solution privilégiée actuellement dans le secteur des transports. Des effets de vieillissement se produisent à chaque instant de la vie de ces batteries, ce qui en fait l'un des critères les plus contraignants de cette technologie en particulier dans un objectif de réduction du coût des batteries. Dans le contexte des VE, la dégradation d'une batterie engendre une perte de ses performances qui se répercute directement sur l'utilisateur. Ces dégradations peuvent se caractériser à la fois par une diminution de l'autonomie du véhicule et par une perte de puissance. Cette réduction des performances des batteries est une entrave majeure au développement du marché des VE, et est de plus couplée à la faible confiance des utilisateurs vis à vis des indicateurs de SOH.

Les mécanismes internes du processus de dégradation sont identifiés, pour la plupart de manière précise, cependant l'estimation de l'état de santé d'une batterie demeure une question ouverte et complexe. En effet, la multitude des facteurs impliqués dans ces phénomènes et leurs interactions nécessitent des investigations importantes. Dans ce chapitre les différentes méthodes existantes ont été parcourues. Toutes les méthodes présentées proposent une solution partielle au problème de l'estimation du vieillissement. En effet, chacune d'entre elles comportent des désagréments importants à leur implémentation en embarqué sur un VE en usage réel. Les études récentes ne permettent pas d'estimer précisément le niveau de santé d'une batterie, sans protocole complexe à mettre en œuvre, lorsqu'elle est utilisée dans un VE. Toutes les méthodologies évoquées proposent de bons résultats sous certaines conditions d'expérimentations ou de tests.

L'adaptation d'une méthodologie dans le contexte des VE en usage réel requiert une méthode flexible, nécessitant peu de mesures à obtenir en temps réel. Une méthode idéale serait capable de fournir une estimation de SOH rapidement à partir des mesures obtenues par le Battery Management System (BMS) (tension, courant) et le VE (température, vitesse). Considérant les contraintes des calculs en temps réel, seulement deux des méthodes présentées semblent acceptables : les modèles de circuit équivalent avec filtres de Kalman et les méthodes statistiques.

De plus, le niveau de précision est primordial pour les applications automobiles. Comme nous l'avons décrit précédemment, l'estimation de l'état de performance d'une batterie est principalement motivée par le manque de confiance des utilisateurs en l'autonomie de leur VE. Les performances des méthodes d'estimation actuelles ne permettant pas de répondre à ce compromis entre application en temps réel et précision, il apparaît clairement la nécessité d'établir un compromis entre les qualités des méthodologies développées, ou bien le développement d'une autre solution répondant à ces contraintes particulières.

1.4.3 Orientation des travaux

Ce chapitre a mis en évidence le besoin de méthodes génériques d'analyse de l'état de santé d'une batterie, indépendantes de sa technologie. L'étude présentée ici porte

sur l'utilisation des données issues d'usages réels, permettant la construction de méthodologies d'analyse et d'estimation du vieillissement d'une batterie en usage sur un VE. Une telle approche, basée sur un volume de données important, offre une bonne représentativité des différents Usages possibles et par conséquent des modèles robustes aux diverses batteries et conditions d'expérimentation possibles. L'analyse des données mesurées lors des Usages possède l'avantage de ne pas produire de biais lié aux hypothèses physiques ou aux conditions de tests engendrées par certains modèles (section 1.3.2), ce qui s'avère être important dans le cadre d'une analyse en usage réel.

Les méthodologies décrites dans cette étude sont fondées sur l'analyse des données mesurées et ont pour but répondre aux problématiques identifiées en section 1.4.1. Les outils développés correspondent à quatre axes de travaux relatifs au vieillissement des batteries en usages réel sur des VE :

1. Compréhension des facteurs provoquant des phénomènes de dégradation, et quantification de leurs impacts.
2. Estimation hors ligne de l'état de santé d'une batterie, sans protocole expérimental dédié.
3. Estimation en ligne de l'état de santé d'une batterie lors d'usages sur un VE.
4. Identification des variables contraignantes lors d'usages réels d'une batterie.

Les diverses méthodologies employées lors de la résolution de ces problématiques sont illustrées par la figure 1.20, décrivant un procédé d'analyse hors ligne des historiques de données enregistrées (en bleu), et une approche de traitement des mesures en ligne (en rouge).

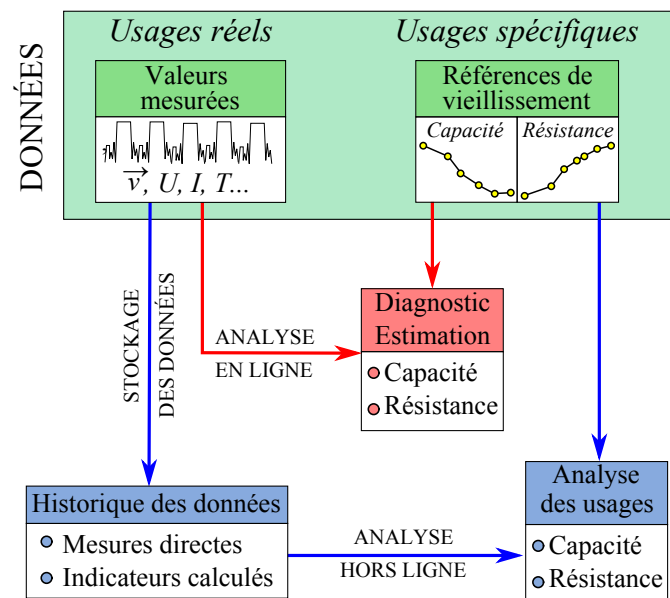


FIGURE 1.20 – Illustration des deux approches globales employées, une analyse **hors ligne** des historiques de données, et une étude **en ligne** des valeurs mesurées

Chacune des problématiques identifiées comportant des caractéristiques propres, il est nécessaire de les traiter de manière dissociée. Les deux premiers points évoqués correspondent à un travail d'analyse des données enregistrées lors des usages réels.

Cette étude hors ligne est étudiée dans le chapitre 2. La troisième problématique nécessite le développement d'une méthodologie d'estimation de l'état de santé d'une batterie, basée sur l'exploitation des mesures obtenues et applicable en ligne. Ce procédé est détaillé dans le chapitre 3. Enfin, le dernier point doit être abordé spécifiquement dans le cadre d'une gestion de flotte. Cette problématique, liée au vieillissement des batteries d'une flotte de VE, est abordée dans le chapitre 4.

Chapitre 2

Analyse hors ligne des données de batterie

2.1	Objectifs et problématiques	50
2.2	Présentation des données	50
2.2.1	Expérimentation ForeWheel	50
2.2.1.1	Description du véhicule	50
2.2.1.2	Description de l'expérimentation ForeWheel	52
2.2.2	Expérimentation AX - SAXO	53
2.2.2.1	Description des véhicules	53
2.2.2.2	Description de l'expérimentation	54
2.2.3	Synthèse des expérimentations	55
2.2.4	Données descriptives	57
2.3	Analyse des causes de vieillissement	59
2.3.1	Méthodologie d'analyse des données	60
2.3.1.1	Sélection des variables	60
2.3.1.2	Analyse en Composantes Principales	62
2.3.2	Application aux données ForeWheel	66
2.3.2.1	Interaction entre les variables mesurées	67
2.3.2.2	Identification des variables influençant le vieillissement	69
2.3.3	Conclusion de l'analyse des données	71
2.4	Vers un estimateur du niveau de dégradation	71
2.4.1	Modèle statistique	71
2.4.1.1	Régression <i>ridge</i>	72
2.4.1.2	<i>Last absolute shrinkage and selection operator</i>	72
2.4.1.3	<i>Elastic net</i>	72
2.4.1.4	Sélection <i>forward</i>	72
2.4.1.5	<i>Least angle regression</i>	73
2.4.2	Résultats et interprétations	73
2.5	Déduction du temps de vie restant d'une batterie	77
2.5.1	Méthodologie	77
2.5.2	Application et résultats	78
2.6	Conclusions - Discussions	79

2.1 Objectifs et problématiques

Ce chapitre traite des analyses hors ligne des données accumulées lors des différentes expérimentations. Dans un premier temps nous explicitons les roulages effectués ayant fourni les données étudiées. Par la suite, l'objectif est d'utiliser les informations contenues dans ces données afin d'identifier et de quantifier les interactions des variables avec les phénomènes de vieillissement d'une batterie. En effet, l'ensemble des données accumulées forme une quantité importante d'informations nécessitant une analyse approfondie.

Dans ce chapitre nous étudions ces données en les traitant hors ligne. La première étape consiste à sélectionner les variables mesurées les plus pertinentes, ayant un impact sur les phénomènes de dégradation d'une batterie. À partir de ces variables, l'objectif est de parvenir à identifier les conditions, d'Usage, responsables des phénomènes de vieillissement d'une batterie. Cette étape permet de relier les conditions d'Usage d'une batterie avec les phénomènes de vieillissement.

Utilisant ces variables détectées comme influençant les mécanismes de dégradations d'une batterie, l'objectif est de construire un modèle d'estimation hors ligne du niveau de SOH d'une batterie en usage. Cette méthodologie produit une estimation du niveau de dégradation subit par une batterie, à partir de variables mesurées simplement lors de ces Utilisations (charges, décharges ou stockages). Une telle approche peut ensuite être étendue pour prédire, à un instant donné, le temps d'utilisation restant d'une batterie, par l'apprentissage d'un modèle à partir de l'historique des données mesurées depuis la mise en service de cette batterie.

2.2 Présentation des données

Nous considérons ici des données issues de trois packs batterie de trois véhicules différents. Cette section présente les différentes batteries et véhicules utilisés ainsi que les détails des procédures d'acquisition des données considérées dans cette étude.

2.2.1 Expérimentation ForeWheel

2.2.1.1 Description du véhicule

Cette expérimentation a été effectuée avec une batterie LiFeO_4 (LFP) instrumentée. Cette batterie utilise 10 modules assemblés à partir de cellules LiFeBatt^1 , ayant les mêmes caractéristiques initiales : une tension nominale de 3.2 V et une capacité maximale de 15 Ah. Le pack batterie Bat_1 possède une tension nominale de 400 V, avec une capacité totale de 75 Ah et une densité énergétique de 80 Wh.kg^{-1} ².

Ces essais ont été effectués sur 23 mois comprenant à la fois des périodes alternées d'utilisation et de stockage. Le processus global est composé de trois différents modes d'utilisation. Le premier mode comprend les périodes de roulage, utilisant un véhicule électrique : la WILL développée par Michelin (figure 2.1), constituée d'une chaîne de traction électrique. La WILL repose sur un châssis d'Opel Agila et intègre deux

1. cf. annexe B.

2. Les données présentées correspondent au packaging complet de la batterie, BMS compris.

moteurs-roues électriques dans les roues avant, générant une vitesse maximale de 140 km.h^{-1} .

Ces roulages ont été effectués sur un circuit privé Michelin (figure 2.2) [128]. Le second mode d'Utilisation de la batterie correspond à de longues périodes de stockage, sans aucune charge/décharge. Le troisième type d'Utilisation appliqué consiste à appliquer des profils de puissance, représentatifs de roulages, sur un banc de test.



FIGURE 2.1 – Le véhicule électrique WILL utilisé durant les phases de roulage de la batterie Bat_1



FIGURE 2.2 – Le circuit de test Michelin employé lors des expérimentations de la batterie Bat_1

L'ensemble du processus mis en œuvre est décomposé en huit sessions distinctes, contenant chacune un des trois types d'Utilisation décrits suivi d'une caractérisation³ complète de la batterie (figure 2.3).

Ces caractérisations suivent toujours le même protocole de test. Ces tests sont donc réalisés dans une pièce maintenue à 25°C . La batterie est initialement laissée au repos pendant une heure, puis elle est chargée à un courant constant de $C/9$ et entièrement équilibrée, précédant un nouveau stockage d'une heure. Après cette période de charge,

3. Une caractérisation de la batterie permet d'obtenir les valeurs références du niveau de performance via un procédé spécifique.

la batterie est complètement déchargée suivant trois profils de puissance. Le premier est une décharge à $C/2$, et les deux autres sont des profils de puissance choisis afin de caractériser la batterie sous conditions contrôlées. Ces profils sont définis conformément aux cycles *ARTEMIS Urbain* [129] et *New European Driving Cycle (NEDC)*, fournissant données nécessaires à l'identification de certaines grandeurs relatives à l'état de santé de la batterie. Les valeurs de capacité globale sont obtenues par simple intégration du courant, et les valeurs de résistance par impulsions de courant.

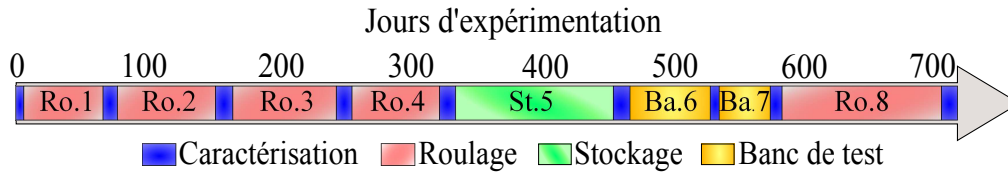


FIGURE 2.3 – Illustration du déroulement global des expérimentations, de la batterie Bat_1 , au fil du temps

2.2.1.2 Description de l'expérimentation ForeWheel

Charge : Toutes les charges sont effectuées sous courant constant, jusqu'à ce que la tension maximale de 3.2 V soit atteinte lors des charges complètes (SOC = 100 %). Dans ce cas, une phase d'équilibrage des cellules est également effectuée. Ceci permet, pour chaque branche parallèle, d'avoir la même tension pour éviter au pack batterie la perte de capacité globale.

Roulage : Les périodes de roulages sont composées d'un profil de vitesse appliqué de manière répétitive sur le circuit privé. L'emploi d'un circuit privé est primordial afin de pouvoir parfaitement contrôler le cycle de roulage. Le même profil de vitesse est réalisé par un unique conducteur professionnel durant l'ensemble des expérimentations. Le profil de vitesse choisi est présenté en figure 2.4. Ce profil contient toutes les phases typiques des usages réels, tels que des arrêts, des démarrages, des parties à basse vitesse, des accélérations, des situations à vitesse constante etc. Le profil de vitesse choisi est représentatif d'une grande partie des sollicitations que la batterie subit en usage véhicule électrique. Notons que ce profil a été modifié lors des premiers essais car la température de la batterie atteignait des valeurs trop élevées. De ce fait, la vitesse maximale initialement fixée à 140 km.h^{-1} a été réduite à 100 km.h^{-1} .

Stockage : La batterie a été stockée environ quatre mois lors de la session *St.5*, dans une pièce dont la température était contrôlée à 25°C . Cette partie de l'expérimentation permet de réduire la proportion de roulages de la batterie, ce qui occulterait la partie de vieillissement calendaire. En effet, lors d'un usage en véhicule électrique, une batterie est stockée une large majorité de sa vie ($> 90 \%$ du temps), et les phénomènes de vieillissement d'une batterie liés au vieillissement calendaire peuvent représenter une grande partie du vieillissement de la batterie. Dans cette expérimentation, les périodes de stockages représentent environ 80 % de l'ensemble de la période considérée, une part importante des stockages ayant lieu entre différents roulages.

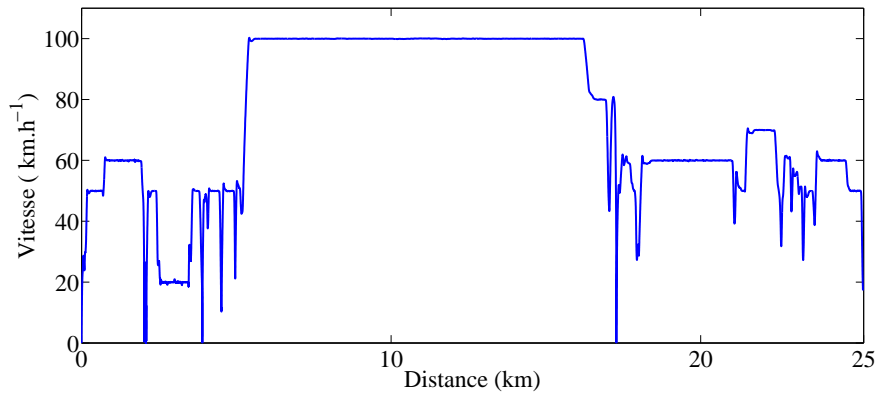


FIGURE 2.4 – Profil de vitesse appliqué à la batterie Bat_1 lors des périodes de roulage, par tour de circuit

Banc de test : Les expérimentations effectuées sur banc de tests reproduisent le même profil de puissance que celui appliqué lors des roulages, sous une température ambiante constante à 25°C. Ce profil de puissance est appliqué sur le banc de tests afin de simuler des roulages, cette fois-ci dans des conditions extérieures contrôlées (absence de vent, pluie, biais du conducteur...). Cette période implique une non continuité de certaines variables issues des roulages, telles que la distance ou la vitesse.

2.2.2 Expérimentation AX - SAXO

2.2.2.1 Description des véhicules

Cette expérimentation utilise deux packs batteries différents, tous les deux composés de cellules Li-ion identiques A123 Systems⁴, ayant une tension nominale de 3.3 V. Le premier pack batterie produit une tension nominale de 134.4 V, une densité énergétique de 72 Wh.kg⁻¹ et une capacité initiale de 119 Ah. Cette batterie Bat_2 est mise en service sur une AX électrique (figure 2.5). De même, une seconde batterie Bat_3 est utilisée sur une SAXO électrique, et fournit une tension nominale de 128 V, une densité énergétique de 75 Wh.kg⁻¹ et une capacité initiale de 95 Ah. L'ensemble des caractéristiques de chacune des trois batteries étudiées sont énumérées en table 2.1.

	Véhicule	Chimie	Cellules	Tension nominale (V)	Densité énergétique (Wh.kg ⁻¹)	Capacité de la batterie (Ah)
Bat_1	WILL	LFP	LifeBatt	400	80	75
Bat_2	AX	LFP	A123	134.4	72	119
Bat_3	SAXO	LFP	A123	128	75	95

TABLE 2.1 – Fiches techniques des trois packs batterie utilisés

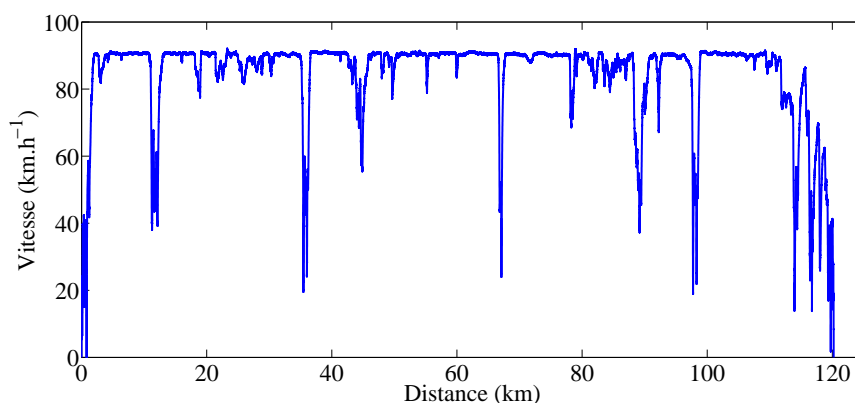
4. cf. annexe A.



FIGURE 2.5 – AX utilisée lors des expérimentations

2.2.2.2 Description de l'expérimentation

Ces deux véhicules sont exploités sous forme d'auto-partage, permettant à différents conducteurs d'utiliser l'une des voitures lors de trajets professionnels ou privés. Les données enregistrées sont représentatives d'utilisation quotidienne d'un véhicule électrique. Ainsi, ces deux véhicules sont employés de manière régulière et non intensive par différents usagers. Cela permet d'obtenir des données issues de divers styles de conduite et d'une large variété de conditions environnementales. Un exemple de profil de vitesse issu de ces expérimentation est présenté en figure 2.6. Ces deux batteries ont été soumises à des caractérisations afin d'obtenir les valeurs références sur leurs niveaux de performance (capacité totale et résistance globale).

FIGURE 2.6 – Exemple de profil de vitesse appliqué à la batterie Bat_2 lors d'un roulage rocade

2.2.3 Synthèse des expérimentations

Chacune des batteries étudiées a été exploitée durant de nombreux mois d'expérimentation (entre 22 et 36 mois). Lors de ces tests, un grand nombre de variables ont été enregistrées lors des charges et des décharges (entre 20 et 50 selon les essais). Une liste non exhaustive des principales variables mesurées lors des tests est présentée table 2.2.

Système mesuré	Description des variables
Véhicule	Distance parcourue (<i>Distance</i>), Temps (<i>Temps</i>) Kilomètres cumulés parcourus (<i>Kilometres</i>) Vitesse du véhicule (<i>Vitesse</i>)
GPS	Position du véhicule (<i>Longitude, Latitude, Altitude</i>) Vitesse (<i>Vit_{GPS}</i>)
Batterie	Nombre de charges (<i>Nb_{cycles}</i>), SOC (<i>SOC</i>), Température moyenne (<i>T_{bat}</i>), Courant et tension (<i>I_{bat}, U_{bat}</i>), Énergie délivrée, fournie (<i>E_{bat,entre}, E_{bat,sorti}, E_{bat,conso}</i>)
Cellules	Températures extrêmes (<i>T_{cell,min}, T_{cell,max}</i>), Tensions (<i>U_{cell,min}, U_{cell,max}</i>)
Autres	Température ambiante (<i>T_{amb}</i>) Roues informations (<i>Roues_{couple}</i>) Tension accessoires (<i>U_{acc}</i>) Courant accessoires (<i>I_{acc}</i>), courant traction (<i>I_{traction}</i>) ⁵

TABLE 2.2 – Noms et descriptions des variables mesurées durant les expérimentations

Toutes ces variables sont mesurées pour chacune des batteries à une fréquence d'échantillonnage de 10 Hz, produisant une base de données globale d'environ 144 Go. Ces variables fournissent un suivi des conditions d'Usages vécus par chacune des batteries durant ces expérimentations. Un descriptif des conditions d'expérimentations des différentes batteries est détaillé table 2.3.

	Lieu	Durée (mois)	Taille des données (Go)	Nombre de trajets	Distance (km)	Énergie délivrée (kWh)	Capacité finale (%)
<i>Bat₁</i>	Ladoux	23	27	1 687	44 038	13 077	87
<i>Bat₂</i>	Grenoble	22	29	1 204	69 449	4 132	87
<i>Bat₃</i>	Grenoble	36	88	1 597	14 283	1 428	91

TABLE 2.3 – Résumé des conditions d'expérimentation pour chacune des trois batteries

Néanmoins, les conditions vécues par les batteries diffèrent et sont donc difficilement interprétables. À titre d'exemple, les répartitions des températures moyennes mesurées sur les batteries (resp. des distances parcourues à chaque roulage) sont pré-

sentées en figure 2.7 (resp. figure 2.8), illustrant la diversité des conditions d'expérimentation.

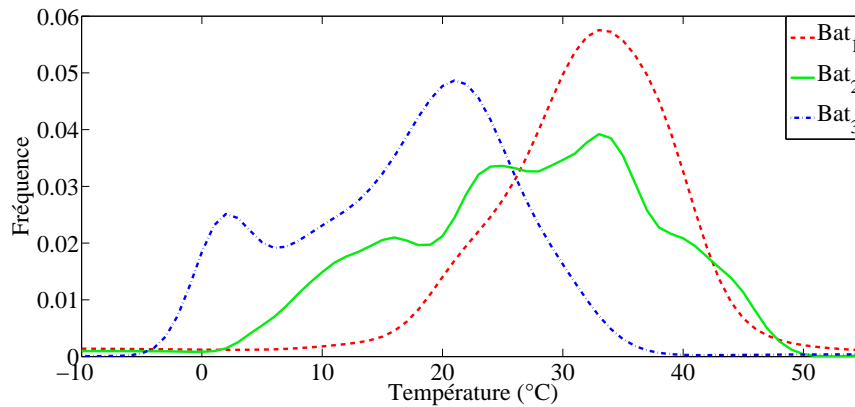


FIGURE 2.7 – Répartition des températures mesurées sur chacune des trois batteries lors des roulages

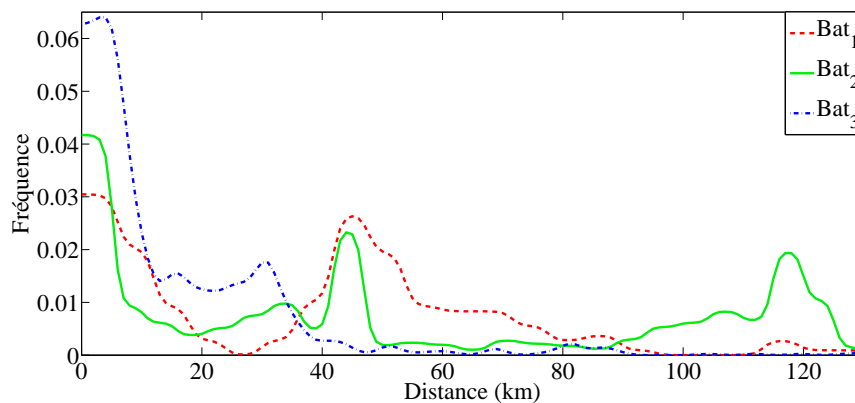


FIGURE 2.8 – Répartition des distances effectuées par roulage, pour chacune des trois batteries

Les caractérisations effectuées sur les batteries permettent d'obtenir des valeurs de capacité et de résistance globale, servant de référence pour la suite de l'étude. Un comparatif des valeurs de capacité et de résistance pour chacune des batteries est présenté en figure 2.9 et figure 2.10. Toutes ces mesures démontrent la difficulté d'analyser les données et le vieillissement obtenu. Les conditions d'utilisation des batteries sont variées et complexes, et résultent de nombreux facteurs interdépendants ce qui induit des analyses directes de variables peu fiables. En effet, les phénomènes de dégradations des batteries diffèrent d'une période à l'autre et les causes de ces phénomènes sont difficiles à mettre en évidence.

Les évolutions du niveau de performance sont différents entre chaque caractérisation, et ce pour chacune des batteries. La problématique est de parvenir à identifier les usages responsables de ces phénomènes de dégradation et d'interpréter les tendances obtenues dans les figures 2.9 et 2.10. Or, les analyses directes des variables (par exemple sur à partir des figures 2.7 et 2.8) ne proposent pas d'interprétation suscep-

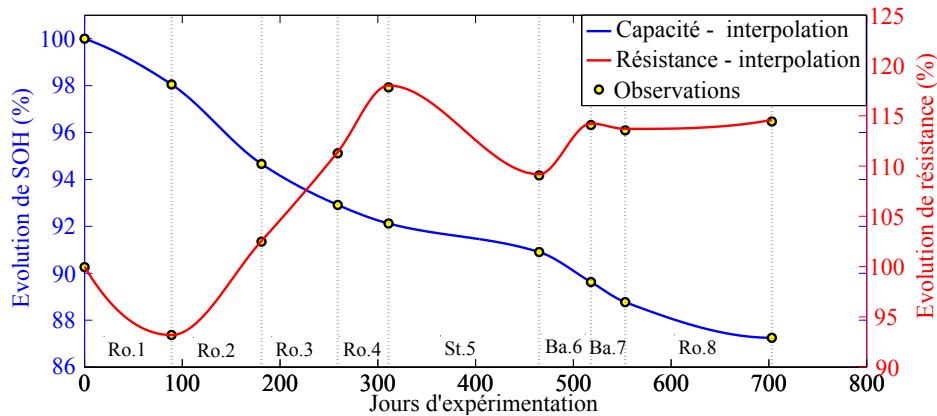


FIGURE 2.9 – Évolutions de capacité et de résistance déterminées par caractérisations sur la batterie Bat_1

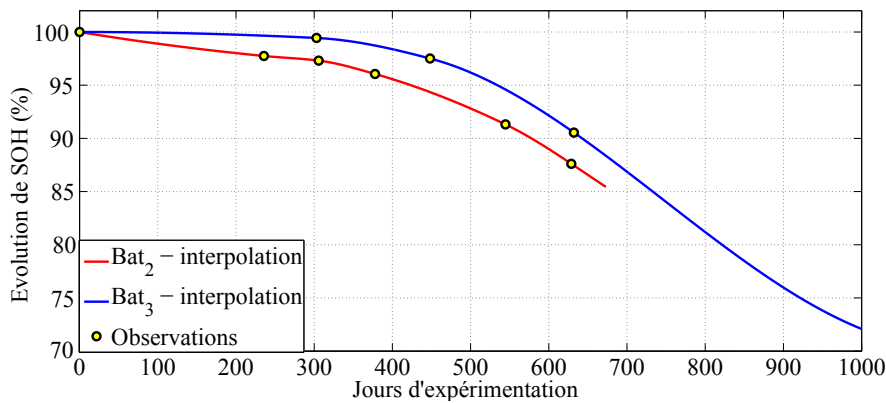


FIGURE 2.10 – Évolution de SOH déterminé par caractérisations sur les batteries Bat_2 et Bat_3

table d'expliquer les modifications de dégradations mesurées pour chacune des batteries.

La complexité de l'analyse directe des mesures amène une problématique concernant la compréhension et l'identification des causes des deux phénomènes de vieillissement mesurés sur une batterie. La résolution de cette problématique nécessite donc une investigation plus approfondie des données, qui sera détaillée dans les prochaines sections.

2.2.4 Données descriptives

Nous avons jusqu'ici analysé les variables mesurées directement sur les batteries. Néanmoins, des informations décrivant les conditions d'Utilisation de la batterie ne sont pas directement mesurées. Il semble possible d'extraire des informations plus détaillées directement à partir de ces mesures. Ainsi, nous définissons ici de nouvelles valeurs de haut niveau susceptibles d'explicitier plus précisément les conditions d'Utilisation de la batterie.

En effet, aucune mesure ne concerne directement les conditions de stockage, qui correspondent pour une batterie employée sur un véhicule électrique particulier à plus de 90 % du temps. Nous utilisons les mesures enregistrées pour obtenir les informations correspondant à ces périodes de stockages d'une batterie. Ces nouvelles variables sont les durées (t_{cal}) et l'état de charge (SOC_{cal}). L'autre variable intervenant dans les conditions de stockage d'une batterie est la température. Pour obtenir cette information, nous employons la dernière valeur de température de la batterie mesurée avant le stockage (T_{av}) et la première valeur mesurée lors de l'usage suivant ce stockage (T_{ap}). À partir de ces deux valeurs, nous estimons une valeur moyenne de température de stockage (T_{cal}), par le biais d'un modèle simplifié de refroidissement thermique. Des exemples illustrent ce modèle en figure 2.11, présentant la température de stockage (T_{cal}) estimée en fonction du temps de stockage t_{cal} , et des valeurs de températures observées (T_{av} et T_{ap})⁶.

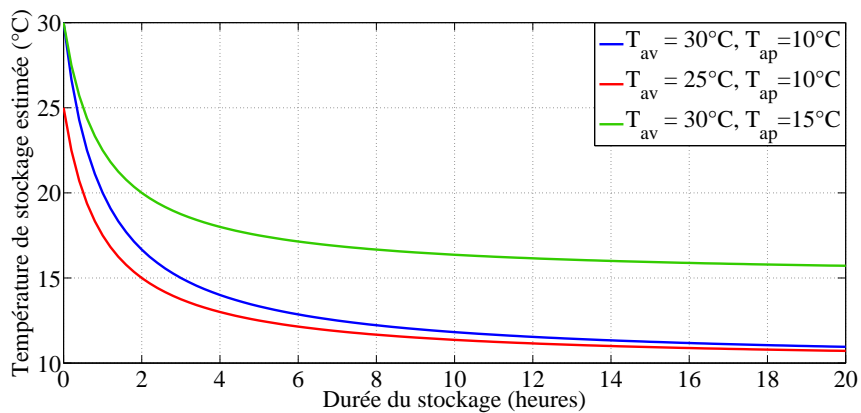


FIGURE 2.11 – Évolution des températures de stockages estimées pour trois différents cas, en fonction de la durée de stockage, avec T_{av} la température avant le stockage et T_{ap} la température après le stockage

Nous introduisons également un comptage du nombre de changements de signe du courant (Nb_{alt}), identifiant les contraintes imposées à la batterie lors de nombreuses phases de freinage électromagnétique. De même, nous calculons la profondeur de décharge ou de charge (ΔSOC) associée à chaque charge ou décharge, correspondant à l'écart de SOC entre le début et la fin d'un cyclage de la batterie, en valeur absolue.

Pour caractériser partiellement le comportement de conduite, le Positive Kinetic Energy (PKE) est déterminé à partir de la vitesse \vec{v} et de la distance d , et est défini selon l'équation 2.1. Le PKE est une valeur positive représentant le comportement de conduite, dont la valeur est proportionnelle à l'agressivité" [130].

$$PKE(t) = \frac{\sum_t (\vec{v}(t)^2 - \vec{v}(t-1)^2)}{d(t) - d(t-1)}, \text{ avec } \vec{v}(t) > \vec{v}(t-1) \quad (2.1)$$

Ces variables calculées à partir des mesures faites directement lors des usages des batteries sont énumérées en table 2.4. D'autre part, la table 2.5 présente un comparatif de ces données descriptives sur l'ensemble des batteries étudiées. Ce comparatif

6. Avec pour modèle : $T_{cal} = T_{ap} + \frac{T_{av} - T_{ap}}{t_{cal} + 1}$.

démontre les différences de conditions subies par ces batteries. Par exemple, la batterie Bat_3 a connue des températures de stockage considérablement plus faibles que les deux autres batteries. La batterie Bat_3 a été utilisée de manière plus contraignante par les utilisateurs : ceci est clairement visible notamment par le nombre très important de changements de signe de courant enregistrés. Notons que cette valeur est également en partie expliquée par certains tests spécifiques effectués sur banc lors des caractérisations. Nous pouvons enfin remarquer la valeur moyenne du PKE calculée pour la batterie Bat_3 , plus importante que celle des deux autres batteries, ce qui s'explique de la même façon par des conditions d'Usage plus contraignantes (figure 2.12).

Origine	Description des variables	Nom
Véhicule	Positive Kinetic Energy	(PKE)
	Température de stockage	(T_{cal})
Batterie mesures	Durée de stockage	(t_{cal})
	SOC initial de stockage	(SOC_{cal})
	Nombre de changements de signe de courant	(Nb_{altI})
	Profondeur de charge/décharge	(ΔSOC)

TABLE 2.4 – Noms et descriptions des variables calculées à partir des mesures faites lors des expérimentations

	Temps de stockage (jours)	Température de stockage (°C)	SOC de stockage (%)	Nombre alt. I (par km)	ΔSOC (%)	PKE (moyen)
Bat_1	389	28	27	38.0	31	0.45
Bat_2	599	21	57	13.3	44	0.42
Bat_3	551	9	71	459.9	25	0.61

TABLE 2.5 – Résumé et moyenne des données descriptives issues de chacune des batteries

De manière générale, les conditions d'Usage de la batterie sont donc très diverses selon la batterie considérée. L'objectif de l'étude réside dans l'analyse de tous ces essais et conditions d'expérimentation complexes et variées, pour extraire des informations sur les causes globales de vieillissement des batteries.

2.3 Analyse des causes de vieillissement

L'objectif de cette section est d'analyser en détails les données obtenues lors des expérimentations, afin d'identifier les variables influentes dans le processus de dégradation d'une batterie. Le but de cette analyse est de parvenir à des identifications et interprétations sur les conditions d'usages induisant différents niveaux de dégradation. Dans un premier temps la méthodologie globale d'analyse des données est explicitée. Puis, dans un second temps, elle est mise en œuvre sur les données réelles obtenues sur la batterie Bat_1 .

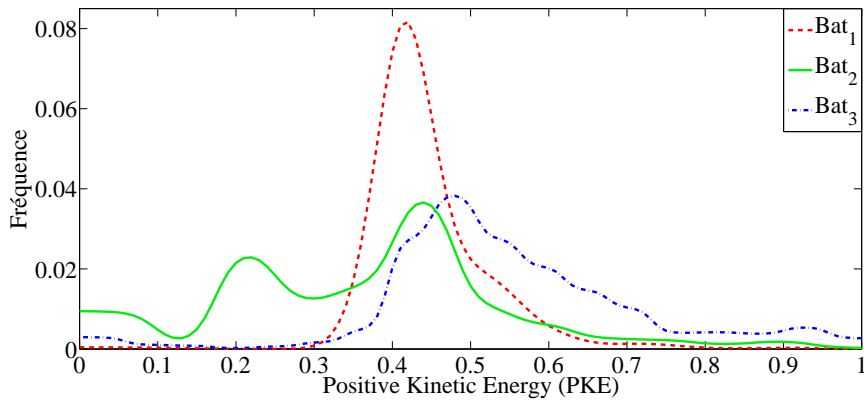


FIGURE 2.12 – Représentation des PKE pour chacune des batteries étudiées

2.3.1 Méthodologie d'analyse des données

L'analyse des données mise en place dans cette étude se divise en deux étapes successives. Dans un premier temps l'approche consiste en une analyse de la pertinence de chacune des variables. Ceci permet d'écarter dans la suite de l'étude les variables redondantes et/ou non pertinentes pour expliquer le phénomène cible. À partir de ces variables sélectionnées, la suite de l'étude consiste en une quantification de leurs impacts respectifs sur ces phénomènes de vieillissement.

2.3.1.1 Sélection des variables

Considérons dans un premier temps l'ensemble des variables mesurées pour conserver toutes les sources possibles d'information. La base de données formée par la totalité de ces mesures étant d'une taille importante, il est nécessaire de procéder à une sélection préliminaire de variables. L'objectif est de réduire la complexité des données analysées et d'éviter des redondances d'information susceptibles de perturber certains algorithmes de modélisation statistique. Cette sélection des variables est également nécessaire pour l'identification des facteurs impliqués dans les phénomènes de vieillissement, validée par les connaissances physiques.

Pour ce faire, l'approche consiste en une étude des informations contenues par chacune des mesures effectuées. Un calcul des coefficients de corrélations entre chacune des variables permet la détection de dépendances. Le calcul de ce coefficient de corrélation de Pearson ($\text{corr}(X, Y)$) entre deux variables X et Y est effectué selon l'équation 2.2, avec σ_X et σ_Y les écarts types, μ_X et μ_Y les espérances et \mathbf{E} l'opérateur de l'espérance [131].

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mathbf{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (2.2)$$

La valeur obtenue renseigne sur la dépendance linéaire entre ces deux variables. Ce coefficient de Pearson est égal à 1 (resp. -1) lorsque l'une des variables est fonction affine croissante (resp. décroissante) de l'autre variable. Plus le coefficient est proche de ces valeurs extrêmes -1 et 1, plus la corrélation entre les variables est forte. Au contraire, une corrélation égale à 0 signifie que les variables ne sont pas corrélées li-

néairement entre elles, mais ne renseigne pas si d'autres relations existent (quadratique par exemple). Une indication de l'interprétation usuellement du coefficient de corrélation est indiquée table 2.6.

Corrélation	Négative		Positive	
	Nulle	-0.1	0	0
Faible	-0.3	-0.1	0.1	0.3
Moyenne	-0.5	-0.3	0.3	0.5
Forte	-1	-0.5	0.5	1

TABLE 2.6 – Interprétation possible du coefficient de corrélation de Pearson

Notons que le coefficient de corrélation n'est ni sensible aux unités des variables étudiées ni aux différences de plages de valeurs. Il est possible, par exemple, de calculer la corrélation entre un courant et une tension. Le calcul de ce coefficient pour chaque couple de variables possible permet de détecter des mesures étant très fortement corrélées à d'autres. Si deux variables sont fortement corrélées, nous sommes dans le cas où elles contiennent la même information. C'est à dire qu'il est suffisant de ne considérer qu'une seule de ces deux variables dans l'étude globale, sans perte d'information, et de synthétiser les données. L'étude des corrélations quantifie donc la force de la relation linéaire entre des paires de variables.

Nous étudions à présent la notion d'*information mutuelle*, qui nécessite la définition préalable de l'*entropie de Shannon*. L'étude de l'*entropie de Shannon* mesure la quantité d'information contenue par les variables (en *bits*) [132]. L'objectif est ici de sélectionner les variables pertinentes et d'améliorer la précision de l'information conservée. Formellement, pour une variable aléatoire X composée de valeurs dans l'ensemble fini \mathcal{X} , l'entropie de Shannon $H(X)$ est définie par :

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log_2 p(x), \quad (2.3)$$

où $p(x)$ est la distribution de probabilité de X .

Cette valeur $H(X)$ est positive ou nulle et quantifie l'incertitude de la distribution de probabilité p . En particulier, $H(X) = 0$, i.e. une incertitude nulle, correspond au cas d'une variable aléatoire X constante. L'entropie de Shannon se définit également pour des variables aléatoires multivariées, mais cela requiert la connaissance de leurs distributions jointes. Pour deux variables aléatoires X et Y , prenant leurs valeurs dans deux ensembles \mathcal{X} et \mathcal{Y} , on a :

$$H(X, Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2 p(x, y), \quad (2.4)$$

où $p(x, y)$ est la distribution de probabilité jointe de X et Y .

À partir de ces définitions, on obtient l'*entropie conditionnelle* $H(X|Y)$, qui correspond à la moyenne des entropies des distributions de probabilités conditionnelles $p(X|Y = y)$:

$$H(X|Y) \equiv H(X, Y) - H(Y) = \sum_{y \in \mathcal{Y}} p(y) \left[- \sum_{x \in \mathcal{X}} p(x|y) \log_2 p(x|y) \right] \quad (2.5)$$

On a donc $H(X|Y) \leq H(X) \leq H(X, Y)$, en particulier, $H(X|Y) \leq H(X)$ démontre le fait que l'incertitude sur X n'est pas diminuée par la connaissance de Y . Notons que dans le cas de conditionnement multiples on a $H(X|Y, Z) \leq H(X|Y) \leq H(X)$.

L'information mutuelle $I(X; Y)$ mesure la dépendance totale entre X et Y :

$$\begin{aligned} I(X; Y) &= H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= \sum_{x \in X} \sum_{y \in Y} p(x, y) \cdot \log \frac{p(x, y)}{p(x)p(y)} \end{aligned} \quad (2.6)$$

L'information mutuelle mesure une distance entre les distributions de X et Y , la valeur obtenue est donc positive, non bornée, et symétrique : $I(X; Y) \geq 0$, $I(X; Y) = I(Y; X)$ [133, 134]. Cette valeur mesure l'information commune entre X et Y . Si X et Y sont indépendants alors l'information mutuelle calculée est nulle. Cela se traduit par le fait qu'une de ces deux variables n'apporte aucune information sur l'autre. Au contraire, si X et Y sont identiques alors l'information contenue dans X est suffisante pour obtenir la variable Y . La figure 2.13 illustre toutes ces notions. L'interprétation de l'information mutuelle est donc différente de celle de la corrélation linéaire.

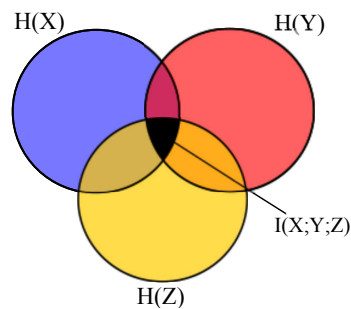


FIGURE 2.13 – Illustration de l'entropie H et de l'information mutuelle I

L'objectif du calcul des entropies et informations mutuelles entre les variables est de démontrer les interactions entre certaines variables. Ceci fournit de plus une quantification des informations partagées par les variables. Nous pouvons ainsi, pour deux variables X et Y , calculer le ratio (2.7) représentatif de la proportion d'information apportée par Y sur X :

$$\frac{H(X) - H(X|Y)}{H(X)} = \frac{I(X; Y)}{H(X)} \quad (2.7)$$

Cette méthode identifie les variables partageant la même information et celles n'apportant aucune information pertinente supplémentaire.

2.3.1.2 Analyse en Composantes Principales

En utilisant désormais un nombre réduit de variables, le but est de les analyser pour identifier certaines relations et modèles au sein de ces données. La description des relations entre deux variables par des techniques statistiques bidimensionnelles, telles que les méthodes explicitées en partie 2.3.1.1, conduit à la question d'une représentation simultanée de données de grande dimension. En effet, les interprétations peuvent



FIGURE 2.14 – Illustration du choix de la représentation des données (J.P. Fenelon ©)

être très différentes selon la manière dont les données sont représentées. L'importance du choix de la représentation choisie est illustrée par la figure 2.14.

L'objectif est d'aborder la structure de corrélation présente dans les données, entre au moins deux variables. L'outil utilisé ici est l'Analyse en Composantes Principales (ACP) [135]. L'avantage de cette méthode est qu'elle permet de décrire des données et de mettre en évidence des structures au sein des données. Le but de l'ACP est de revenir à un espace de dimension réduite en déformant le moins possible de la réalité. Il s'agit donc d'obtenir le résumé le plus pertinent des données initiales. D'un point de vue mathématique, cela consiste en un simple changement de base, c'est à dire de modifier l'espace de représentation dans un nouveau repère. Autrement dit, il s'agit de passer d'une représentation dans la base canonique des variables initiales (dans \mathbb{R}^p) à une représentation dans la base des facteurs définis par les vecteurs propres des corrélations.

Formalisons cette notion, on s'intéresse à l'étude d'un tableau de données stockées sous la forme d'une matrice \mathbf{M} de taille $n \times p$. Cette matrice \mathbf{M} contient sur chaque ligne i , les valeurs prises par l'individu i sur les p variables : (m_i^1, \dots, m_i^p) ; et sur une colonne j , les valeurs de la variable j pour les n individus : $(m_1^j, \dots, m_n^j)^t$. Pour simplifier les notations nous identifions l'individu i par $m_i = (m_i^1, \dots, m_i^p)$, et la variable j par $m^j = (m_1^j, \dots, m_n^j)^t$.

Notons $d(i, j)$ la distance euclidienne entre deux individus m_i et m_j , selon l'équation suivante :

$$d(i, j) = \left(\sum_{k=1}^p (m_i^k - m_j^k)^2 \right)^{1/2} \quad (2.8)$$

Le centre de gravité \mathbf{g} du nuage de points est défini par l'équation 2.9, il est également appelé point moyen.

$$\mathbf{g} = (\bar{m}^1, \dots, \bar{m}^p)^t \quad (2.9)$$

où $\bar{m}^j = \frac{1}{n} \sum_{i=1}^n m_i^j$ est la moyenne⁷ des valeurs prises par la j -ème variable.

7. Nous supposons ici que tous les individus ont le même poids.

Afin d'éviter qu'une variable ne prenne trop d'importance par rapport aux autres simplement parce qu'elle varie dans de plus grandes proportions, on procède à une normalisation des données en centrant et en divisant chaque variable par son écart-type. Par abus de langage, \mathbf{M} désignera par la suite cette matrice centrée-réduite.

L'inertie est une notion fondamentale car elle mesure la dispersion du nuage de points autour de son centre de gravité \mathbf{g} . L'inertie totale $I_{\mathbf{g}}$ du nuage autour de son centre de gravité \mathbf{g} est :

$$I_{\mathbf{g}} = \frac{1}{n} \sum_{i=1}^n d^2(\mathbf{g}, m_i) = \frac{1}{n} \sum_{i=1}^n (m_i - \mathbf{g})^t \mathbf{M} (m_i - \mathbf{g}) \quad (2.10)$$

Plus l'inertie est grande, plus le nuage est dispersé, et au contraire plus elle est petite, plus le nuage est concentré autour de son centre de gravité \mathbf{g} .

On cherche à présent une représentation des n individus dans un sous-espace F_k de \mathbb{R}^p de dimension k (avec k petit). Comme la projection diminue nécessairement les distances, on cherche le sous-espace F_k qui maximise la moyenne des carrés des distances entre individus, formalisé en équation 2.11. Il faut donc que l'inertie du nuage projeté soit maximale.

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d^2(i, j) = 2I_{\mathbf{g}} \quad (2.11)$$

Cet espace factoriel F_k devra être *ajusté* le mieux possible au nuage de points, étant défini comme le sous-espace tel que le nuage projeté ait une inertie $I_{\mathbf{g}}$ maximale. Autrement dit, on cherche à définir k nouvelles variables (A_1, \dots, A_k) comme des combinaisons linéaires des p variables initiales, qui feront perdre le moins d'informations possible. Ces variables A_i seront appelées *composantes principales*, les axes qu'elles déterminent *axes principaux* et les formes linéaires associées *facteurs principaux*. La figure 2.15 illustre cette modification d'axes.

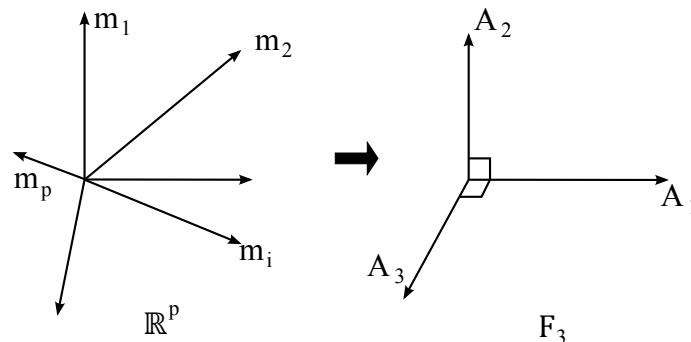


FIGURE 2.15 – Illustration de la modification d'axes, de \mathbb{R}^p vers F_3

Pour déterminer le sous-espace F_k , on définit le premier axe (premier axe factoriel A_1) sur lequel le nuage de points, formé par les données, se déforme le moins possible en projection. Géométriquement, A_1 détermine une nouvelle direction dans le nuage de points qui suit l'axe d'allongement maximal du nuage, et passe par le centre de gravité \mathbf{g} du nuage (figure 2.16). Concrètement, le premier axe factoriel A_1 sera celui sur lequel le nuage de points projeté est d'inertie maximale.

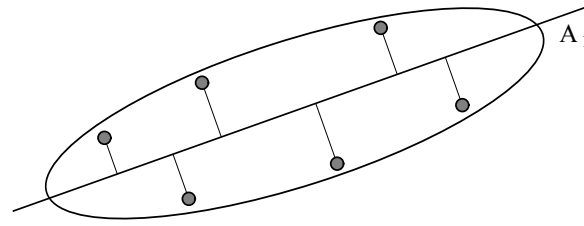


FIGURE 2.16 – Illustration de la formation du premier axe factoriel A_1 , à partir d'un nuage de points

On cherche ensuite un second axe A_2 sur lequel le nuage se déforme le moins en projection, après le premier axe, tout en étant orthogonal au premier, au point \mathbf{g} . Le second A_2 sera celui qui, après le premier axe, est tel que le nuage projeté est d'inertie maximal, tout en étant orthogonal au premier ; ainsi de suite jusqu'à l'obtention des k axes souhaités. Ainsi, la capacité des axes à résumer le nuage se détériore au fur et à mesure que l'on observe des axes de rang élevé.

On démontre que pour une dimension k donnée, le sous-espace maximisant l'inertie du nuage projeté est engendré par les k vecteurs propres de la matrice de covariance V_M associés aux k premières valeurs propres $(\lambda_1, \dots, \lambda_k)$, rangées par ordre décroissant. La matrice V_M étant définie par l'équation 2.12.

$$V_M = \frac{1}{n} \mathbf{M}^t \mathbf{M} \quad (2.12)$$

Nous avons donc obtenu k composantes principales (A_1, \dots, A_k) centrées, non corrélées entre elles. L'inertie expliquée par l'axe A_i est la valeur propre λ_i . L'inertie $I_{F_k^\perp}$ expliquée par le sous-espace factoriel F_k engendré par les axes factoriels est définie par l'équation 2.13.

$$I_{F_k^\perp} = \lambda_1 + \dots + \lambda_k \quad (2.13)$$

Le pourcentage d'inertie expliqué par ce sous-espace est donc $\frac{\lambda_1 + \dots + \lambda_k}{p}$.

La sélection du nombre d'axes à retenir est une étape importante de l'ACP. Les critères utilisés sont principalement empiriques : la méthode *elbow* qui consiste à détecter un coude sur le diagramme des valeurs propres ; ou le critère de Kaiser qui consiste à ne retenir uniquement les valeurs propres supérieures à la moyenne de celles-ci [136]. Ainsi sur l'exemple présenté en figure 2.17, le choix des trois premières valeurs propres permet d'atteindre 80 % de l'inertie initiale. Le choix des deux premiers axes factoriels offre une représentation de 70 % de l'information initiale dans un plan. Ce dernier choix permet de visualiser les données dans ce plan composé des deux premières composantes principales, ce qui est particulièrement intéressant pour les interprétations, en perdant les informations contenues sur les axes manquants.

La qualité de représentation d'une variable j sur le k -ème axe factoriel est exprimée par le coefficient de corrélation linéaire $r(A_i, m^j)$ entre la composante principale A_i et la variable m^j . On montre que cette valeur est définie par l'équation 2.14 :

$$r(A_i, m^j) = \sqrt{\lambda_i} u_i^j \quad (2.14)$$

où u_i^j est la j -ème coordonnée du vecteur propre u_i associé à λ_i .

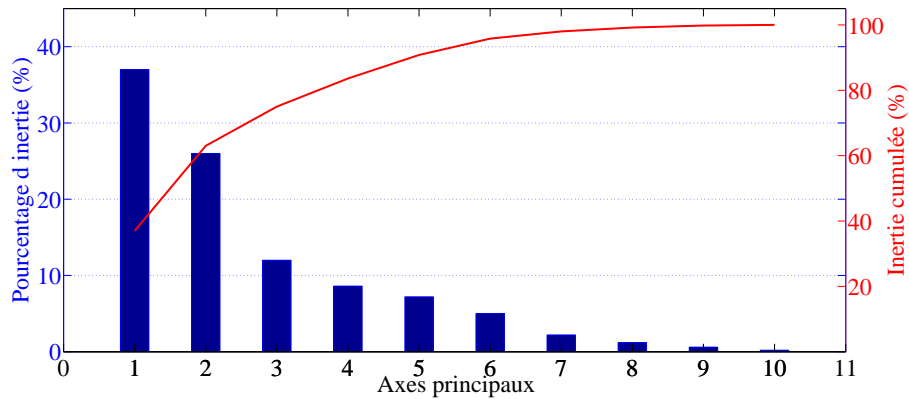
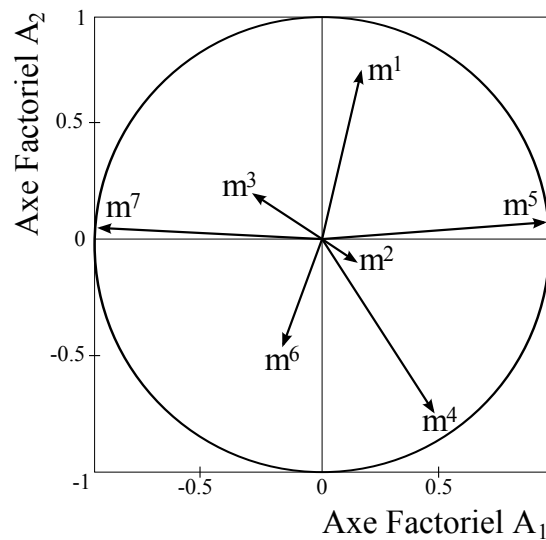


FIGURE 2.17 – Exemple de répartition de l’inertie associée aux axes principaux

La méthode la plus naturelle pour interpréter les composantes principales est d'utiliser ces corrélations. Usuellement, pour un couple de composantes principales A_1 et A_2 , on représente les corrélations dans une figure nommée cercle des corrélations (figure 2.18), où chaque variable m^j est représentée par un point ayant pour coordonnées $(r(A_1, m^j), r(A_2, m^j))$.

FIGURE 2.18 – Exemple de projection de variables sur un plan factoriel (A_1, A_2)

Cette représentation permet d'identifier les variables reliées avec les axes factoriels (positivement ou négativement), et d'en déduire des relations entre variables initiales. L'étude peut se poursuivre si nécessaire par une visualisation des individus pour la détection de possibles groupes d'individus.

2.3.2 Application aux données ForeWheel

Le but est à présent d'utiliser l'ensemble des méthodologies d'analyse des données, décrites dans la section 2.3.1, avec les données collectées lors des expérimentations.

Les méthodes explicitées sont appliquées à la base de donnée, avec pour objectif de comprendre les phénomènes et interactions ayant un impact sur les phénomènes de vieillissement d'une batterie.

Notons que les résultats présentés sont uniquement issus des tests effectués par la batterie Bat_1 . Ce choix est justifié car le protocole d'expérimentation mis en œuvre sur cette batterie Bat_1 est le plus adapté pour ce type d'étude. Ce protocole de test contient en effet de nombreuses caractérisations, fournissant un grand nombre d'information. De plus, la batterie Bat_1 est la seule présentant des valeurs références fiables de résistance au cours du temps.

2.3.2.1 Interaction entre les variables mesurées

Dans un premier temps, nous considérons l'ensemble des variables mesurées lors des expérimentations. La liste des principales valeurs prises en compte est énumérée en table 2.2 et en table 2.4. Comme explicité dans la section 2.3.1.1, il est nécessaire de trier ces variables et de ne considérer pour la suite de l'étude uniquement celle contenant une information différente des autres.

En effet, de nombreuses variables mesurées ne présentent aucun lien avec les phénomènes de vieillissement d'une batterie, ou même avec la batterie de manière générale. Certaines variables sont identifiées comme totalement indépendantes de notre étude. Nous pouvons déduire intuitivement que certaines mesures décrites en table 2.2 n'apportent aucune information sur la durée de vie de la batterie, telles que les données Global Positioning System (GPS), les informations relatives aux roues ou bien aux accessoires.

Considérons ici les variables restantes. Afin de voir si certaines variables sont redondantes nous calculons le coefficient de corrélation de ces mesures (équation 2.2), lors de charges puis lors de décharges. Les résultats de ces coefficients de corrélation obtenus pour les mesures obtenues à 10 Hz pour un ensemble de charges (resp. de décharges) sont présentés en figure 2.19a (resp. figure 2.19b). Selon la période considérée, les variables étudiées diffèrent, certaines n'étant pas mesurées lors de chaque cyclage de la batterie. Par exemple, les variables *Vitesse*, *Distance* et *Kilometres* ne sont pas mesurés lors des charges. Notons que ces résultats ne concernent donc que les périodes de cyclage de la batterie et non les stockages.

À partir de ces résultats, on peut clairement conclure à quelques fortes dépendances linéaires entre variables. Ainsi, les variables de températures mesurées au niveau cellule contiennent la même information ($|\text{corr}(T_{bat}, T_{cell, min})| > 0.9$) que celle mesurée au niveau batterie, ce qui est interprétable du fait de la relation directe liant ces variables. Dans ce cas, nous conservons pour la suite de l'étude uniquement la valeur T_{bat} correspondant à la température mesurée du pack batterie ; les variables $T_{cell, min}$ et $T_{cell, max}$ n'apportant pas d'information supplémentaire par rapport à T_{bat} , sont supprimées de la suite de l'étude.

De même, les mesures de tension faites au niveau cellule ne présentent pas d'apport comparé aux mesures obtenues au niveau batterie. Les variables $U_{cell, min}$ et $U_{cell, max}$ sont donc également supprimées de l'étude, au profit de U_{bat} . Notons que cette remarque est valable ici car nous étudions le vieillissement global de la batterie, et non un suivi détaillé du SOH des cellules et de leur homogénéité.

Les calculs de corrélations démontrent également la relation linéaire entre les vitesses du véhicule mesurées par le GPS et par le véhicule directement (figure 2.19b).

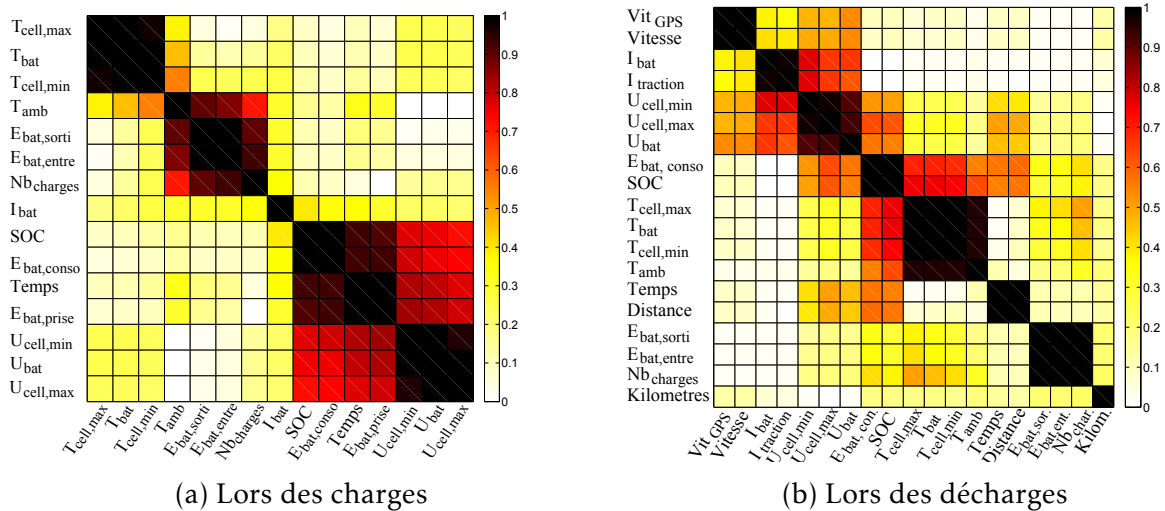


FIGURE 2.19 – Coefficients de corrélation calculés entre les variables mesurées lors des charges puis des décharges - Des variables totalement linéairement liées ont un coefficient de 1

L'une de ces mesures est superflue et érudée, nous conservons uniquement la vitesse enregistrée par le véhicule (*Vitesse*).

Il apparaît une forte corrélation entre les énergies mesurées en entrée et en sortie de la batterie. Ceci est la conséquence du fait que la batterie délivre l'énergie qu'elle a accumulée soit lors des charges, soit lors des moments de récupération d'énergie dans les roulages. Il sera par la suite seulement pris en compte l'énergie délivrée par la batterie ($E_{bat,sorti}$).

Les figures 2.19 permettent d'identifier la forte corrélation absolue ($corr > 0.8$) entre d'autres variables moins prévisibles. Ainsi, l'énergie consommée ($E_{bat,conso}$) apparaît comme liée linéairement au niveau de charge (SOC), ce qui peut être expliqué par la technologie de la batterie.

D'autre part, il est important de prendre du recul sur certains résultats dus au protocole expérimental et non à la relation linéaire liant ces variables. C'est ici le cas de la forte corrélation absolue obtenue ($corr > 0.9$) entre la distance ($Distance$) et le temps ($Temps$), qui est la conséquence de l'expérimentation mise en œuvre lors des tests de la batterie Bat_1 . En effet, le véhicule parcourait tout au long de la journée le même profil de vitesse et les mêmes plages de charges, ce qui explique cette forte relation linéaire. Dans ce cas, les deux variables sont bien évidemment prises en compte dans l'étude. Cette observation est faite de la même manière sur la relation entre le nombre de charge (Nb_{cycles}) et les énergies entrées et sorties de la batterie ($E_{bat,entre}$, $E_{bat,sorti}$). Cette particularité est obtenue du fait de la régularité du profil de vitesse et de la distance parcourue entre chaque charge.

L'étude des coefficients de corrélation entre variables a permis de réduire considérablement la liste des mesures étudiées de plus de 30 variables à désormais 8 mesures ($Vitesse$, SOC , $E_{bat,sorti}$, T_{bat} , $Distance$, Nb_{cycles} , T_{amb} , I_{max}) sans perte d'information. Afin de compléter cette étude des paramètres, nous ajoutons les variables descriptives présentées en table 2.4, ainsi que les valeurs références de capacité et de résistance. Nous procédons au calcul des entropies des variables prises en compte, dans le but

de détecter d'éventuelles variables ayant la même information, ou bien une variable constante. Le calcul mis en œuvre consiste ici en l'application de l'équation 2.7 sur chaque couple de variable (X, Y) . Le ratio obtenu par ce calcul est présenté en figure 2.20 et illustre la proportion d'information apportée par une variable Y sur une autre X .

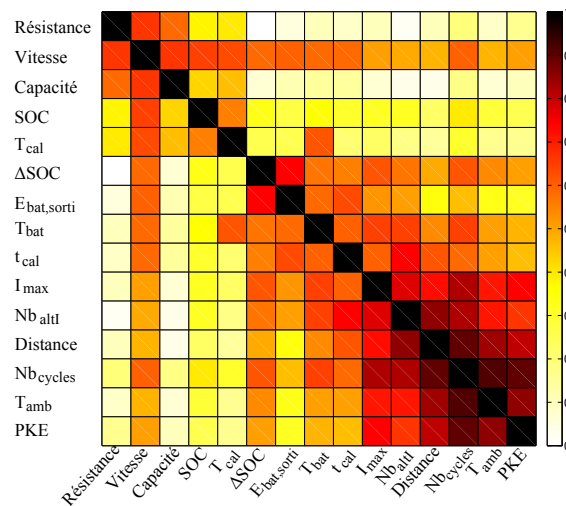


FIGURE 2.20 – Ratio de l'entropie conditionnelle $H(X|Y)$ par rapport à l'entropie de $H(X)$, défini selon l'équation 2.7 - Des variables contenant la même information ont un ratio de 1

Notons que les résultats ne sont pas symétriques. En effet nous observons :

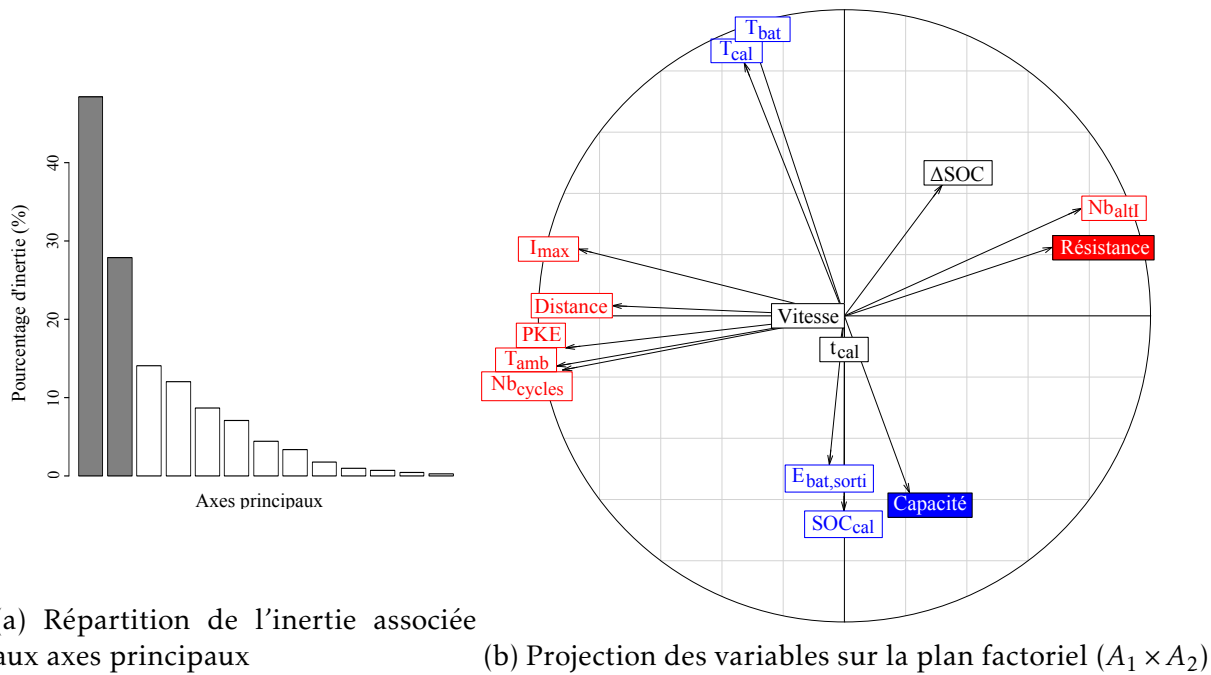
$$\begin{aligned} H(X) &\neq H(Y) \\ H(X|Y) &\neq H(Y|X) \end{aligned} \quad (2.15)$$

Cette figure 2.20 permet de conclure à une non redondance des informations apportées par les variables prises en compte. En effet, peu des ratios calculé n'illustre une large proportion d'information commune entre deux variables (ratio > 0.8). Il est de plus important de mettre en évidence le fait que les mesures de capacité et de résistance ne partagent pas d'information directe avec la majorité des autres variables. Cela illustre parfaitement les nombreuses interactions responsables des phénomènes de vieillissement, leurs identifications nécessitent donc des investigations plus approfondies.

2.3.2.2 Identification des variables influençant le vieillissement

Nous avons obtenu un nombre de variables réduit, simplifiant leur étude. Le but est à présent d'analyser ces données afin d'identifier les causes du vieillissement, par la composition de groupes de variables. Notons, comme précédemment, que l'étude porte ici sur les données de la batterie Bat_1 .

Pour ce faire nous appliquons une Analyse en Composantes Principales (ACP) sur l'ensemble des données, d'après la méthodologie détaillée en section 2.3.1.2. Le choix

FIGURE 2.21 – Résultat de l'ACP effectuée sur les mesures de la batterie Bat_1

se porte sur une projection sur les deux premiers axes factoriels (figure 2.21a). La projection des variables sur ces deux axes factoriels est présentée en figure 2.21b.

Les résultats de cette ACP mettent en évidence deux groupes de variables différents. On peut relier chacun des axes factoriels à un groupe de variables différent. Cette analyse identifie les deux phénomènes principaux de vieillissement (capacité et résistance) comme associés à des axes factoriels différents, ce qui revient à considérer leurs comportements comme disjoints. Autrement dit, nous pouvons ici distinguer les évolutions de capacité et de résistance car elles ne sont pas corrélées naturellement.

Les variables principalement reliées aux axes factoriels sont représentées avec une couleur **rouge** (resp. **bleue**) pour l'axe factoriel A_1 (resp. A_2). Les groupes de variables sont :

- **Groupe 1** : *Résistance, Nb_{altI} , I_{max} , Distance, PKE, T_{amb} , Nb_{cycles}*
- **Groupe 2** : *Capacité, SOC_{cal} , T_{cal} , T_{bat} , $E_{bat,sorti}$*

Les variables non associées à ces groupes (ΔSOC , *Vitesse*, t_{cal}) ne sont pas détectées comme directement corrélées aux comportements de la capacité et de la résistance. Il est important de noter que cette remarque ne signifie pas que ces variables sont indépendantes du vieillissement d'une batterie.

Il est possible de procéder à une interprétation des axes et donc des relations entre les variables de chacun de ces groupes. Ainsi, le groupe 1, lié à la résistance, est composé de variables issues du comportement de l'utilisateur (*PKE, I_{max} , Nb_{altI}*), de mesures relatives à la durée et aux conditions des essais (*Distance, T_{amb} , Nb_{cycles}*). Ceci permet de mettre en évidence la relation entre l'évolution de la résistance de la batterie avec les conditions d'Usage de la batterie. Il apparaît une dépendance importante entre les conditions d'évolution de la résistance d'une batterie et de ses conditions d'Usage, qu'elles soient contrôlées (style de conduite...) ou subies (température ambiante...).

Le groupe 2, lié à la capacité, correspond à des variables provenant des conditions de stockage de la batterie (T_{cal} , SOC_{cal}) et à des températures vécues par la batterie (T_{bat}). Les mesures de capacité ont également une part importante de la composition du second axe factoriel. Ces analyses permettent de conclure à une relation entre l'évolution de capacité d'une batterie et les conditions de stockage subies.

2.3.3 Conclusion de l'analyse des données

L'ensemble des analyses effectuées sur les données obtenues par les expérimentations ont permis, dans un premier de temps, de trier les variables redondantes ou non pertinentes. Ainsi, une liste réduite de variables est considérée pour la suite de l'étude (section 2.3.2.1).

D'autre part, une ACP effectuée sur ces données réduites ont permis l'identification de deux groupes de variables (section 2.3.2.2), chacun lié à un phénomène global de vieillissement distinct. L'évolution de capacité d'une batterie est détectée comme rattachée aux conditions de stockage ainsi qu'aux niveaux de température qu'elle atteint lors de ces Usages.

L'évolution de résistance totale d'une batterie est, quant à elle, identifiée comme étant reliée aux conditions d'usages de celle ci. Ainsi, les variables illustrant directement le mode d'utilisation, telles que PKE ou I_{max} , sont détectées comme dépendant de l'évolution de la résistance mesurée.

Notons que ces conclusions confirment les résultats de l'analyse bibliographique du chapitre 1 sur les différents travaux portant sur les mécanismes de dégradation des batteries.

2.4 Vers un estimateur du niveau de dégradation

Sur la base des observations faites jusqu'à présent, l'extension naturelle réside dans la construction d'un modèle d'estimation du niveau de vieillissement d'une batterie. Un tel modèle peut être formé à partir des variables détectées comme influençant le mode de dégradation d'une batterie (section 2.3.1).

La méthodologie privilégiée est l'utilisation d'un modèle simple d'estimation linéaire. Ce choix est justifié car le vieillissement d'une batterie est fréquemment modélisé par des équations physiques linéaires. Ceci peut être expliqué par la nature des tendances des dégradations observées.

2.4.1 Modèle statistique

Nous voulons établir une fonction f , permettant d'estimer une sortie réelle Y à partir d'un vecteur d'entrées $X = (X_1, \dots, X_n)$, suivant l'équation 2.16 :

$$f(X) = \beta_0 + \sum_{j=1}^p X_j \beta_j \quad (2.16)$$

L'équation 2.16 fournit une estimation $\hat{Y} = f(X)$ à partir d'un vecteur d'observations X de dimension p . L'objectif est d'estimer les paramètres β de ce modèle linéaire.

Pour ce faire, différentes résolutions sont considérées et les résultats de leurs applications respectives sont comparées. Le fonctionnement de chacune de ces méthodes est explicité dans les prochains paragraphes.

2.4.1.1 Régression *ridge*

La régression *ridge* restreint les coefficients de la régression en imposant la complexité du modèle [139]. Le paramètre optimal $\hat{\beta}^{ridge}$ minimise une somme résiduelle des carrés, avec une pénalisation L_2 contrôlée par un paramètre de régularisation $\lambda_{ridge} \geq 0$ (équation 2.17).

$$\hat{\beta}^{ridge} = \underset{\beta_0 \in \mathbb{R}, \beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda_{ridge} \sum_{j=1}^p \beta_j^2, \quad (2.17)$$

avec n le nombre d'observations et p la dimension.

2.4.1.2 *Last absolute shrinkage and selection operator*

Le *Last Absolute Shrinkage and Selection Operator* (LASSO) est une technique similaire à la régression *ridge* permettant d'estimer les paramètres β_i et d'éliminer les variables X_i en ramenant leurs coefficients à zéro grâce à sa parcimonie [140]. Les variables non pertinentes possèdent un coefficient de régression laissé nul par l'algorithme. Pour cela, on définit l'estimateur $\hat{\beta}^{lasso}$ par l'équation 2.18, avec $\lambda_{lasso} \geq 0$ un paramètre de régularisation L_1 .

$$\hat{\beta}^{lasso} = \underset{\beta_0 \in \mathbb{R}, \beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda_{lasso} \sum_{j=1}^p |\beta_j|, \quad (2.18)$$

avec n le nombre d'observations et p la dimension.

2.4.1.3 *Elastic net*

La méthode *elastic net* suit un processus automatique de sélection de variables. Elle *elastic net* surmonte les limites des méthodes précédentes dues à leurs pénalisations [141]. L'estimation des paramètres $\hat{\beta}^{elasticnet}$ se fait avec l'utilisation de deux termes de régularisation λ_1 et λ_2 , selon l'équation 2.19.

$$\hat{\beta}^{elasticnet} = \underset{\beta_0 \in \mathbb{R}, \beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2, \quad (2.19)$$

avec n le nombre d'observations et p la dimension.

2.4.1.4 *Sélection forward*

La sélection *forward* est une méthode de régression pas-à-pas, s'initialisant sans aucune variable dans le modèle, puis en les ajoutant une à une itérativement, jusqu'à ce qu'un critère de précision sur l'erreur soit atteint. À chaque pas, la variable ajoutée

au modèle est celle ayant la plus grande corrélation absolue avec le vecteur résiduel [137].

2.4.1.5 *Least angle regression*

Least Angle Regression (LARS) est une méthode de régression offrant une version adoucie de la sélection *forward* [138]. L'algorithme LARS démarre avec un ensemble vide de variables actives. De la même façon que la sélection *forward*, la corrélation entre chacune des variables et le résidu correspondant sont calculés, la variable induisant la plus grande corrélation est ajoutée au modèle. La première direction est définie par la solution des moindres carrés considérant la seule variable active à la première étape. Allant dans cette direction, une autre variable obtient la même corrélation que le vecteur résiduel en tant que variable active. La marche est interrompue et la nouvelle variable est ajoutée à l'ensemble des variables actives. La nouvelle direction de la marche est alors définie par la solution des moindres carrés des deux variables actives, et ainsi de suite. Après un nombre d'itérations à définir, la solution des moindres carrés complète est atteinte suivant un chemin de régularisation.

2.4.2 Résultats et interprétations

Chacune de ces méthodes est appliquée dans le contexte de l'estimation du niveau de capacité et de résistance d'une batterie. Dans ce domaine spécifique de l'étude, les observations X coïncident à une matrice ($S \times n$) composée des n variables pertinentes mesurées pendant S semaines. Ainsi, pour une semaine donnée chacune de ces n variables correspondent à des moyennes, des cumuls ou des maxima de la totalité des valeurs enregistrées durant une semaine. La fonction f définie par l'équation 2.16 fournit, pour cette entrée X , une estimation d'évolution d'un phénomène de vieillissement \widehat{Y} .

Dans le contexte du vieillissement de batteries, comme détaillé dans les sections précédentes, les deux principaux phénomènes de vieillissement sont les évolutions de capacité et de résistance. Ces phénomènes ont des dynamiques différentes et nécessitent deux approches séparées. Ainsi, deux types de modèles : f_{cap} et f_{res} , sont construits à partir de chacune des méthodes explicitées en partie précédente, un pour estimer l'évolution de capacité et un concernant l'évolution de la résistance de la batterie étudiée. Ces modèles sont établis à partir des données d'apprentissages constituées d'un ensemble d'observations X^{app} , associé à des références de capacité Y_{cap}^{app} et de résistance Y_{res}^{app} extrapolées pour chaque semaine à partir des caractérisations.

Cette méthodologie est appliquée séparément sur chacune des trois batteries étudiées. En effet, celles-ci possèdent des caractéristiques différentes et un modèle appris sur une technologie de batterie ne peut être testé sur une autre technologie. La seule batterie ayant des valeurs de résistance régulières est la batterie Bat_1 . Nous présentons ici uniquement les résultats obtenus sur cette batterie.

D'après la méthodologie explicitée en section 2.4.1, les différents modèles d'estimation de capacité f_{cap} (resp. de résistance f_{res}) sont bâtis à partir de l'ensemble d'apprentissage $\{Y_{cap}^{app}, X^{app}\}$ (resp. $\{Y_{res}^{app}, X^{app}\}$). Dans cette étude, différents taux d'apprentissage sont explorés. Les modèles sont construits en définissant l'ensemble d'apprentissage comme étant composé de 10, 25 ou 50 % (resp. 25, 50 ou 75 %) de la totalité

des données disponibles.

La partie restante des données est utilisée afin de former un ensemble de test $\{Y_{cap}^{test}, X^{test}\}$ (resp. $\{Y_{res}^{test}, X^{test}\}$). Ce dernier calcule une erreur d'estimation moyenne pour chacun des modèles. L'erreur $err(f)$ entre les estimations $\widehat{Y}^{test} = \{\widehat{Y}_1^{test}, \dots, \widehat{Y}_m^{test}\}$, obtenues avec un modèle f , et des données références $Y^{test} = \{Y_1^{test}, \dots, Y_m^{test}\}$, est calculée selon l'équation 2.20 :

$$err(f) = \frac{100}{m} \sum_{j=1}^m \frac{|\widehat{Y}_j^{test} - Y_j^{test}|}{Y_j^{test}} \quad (2.20)$$

Les erreurs obtenues à partir de chacun des modèles évoqués sont présentées en table 2.7. Notons que les résultats obtenus pour l'évolution de capacité sont nettement meilleurs que ceux acquis pour l'évolution de résistance. Ceci est expliqué par la tendance monotone de la capacité qui décroît constamment, avec des variations relativement faibles. Toutes les méthodes étudiées permettent d'obtenir de très bons résultats sur l'estimation de la capacité d'une batterie ($err < 1\%$), à partir d'un taux d'apprentissage de 25 %. Au contraire, l'évolution de résistance est plus complexe et les modèles nécessitent plus de données d'apprentissage pour obtenir une précision acceptable. Ceci était prévisible considérant la nature complexe de l'évolution de résistance réelle (figure 2.9). De ce fait, les estimations du niveau de résistance de la batterie faites à partir d'un taux d'apprentissage de 75 % permettent d'obtenir une erreur d'estimation d'environ $err = 3\%$.

Méthode	Erreur d'estimation (%)						
	Capacité			Résistance			
	Taux app. (%)	10	25	50	25	50	75
Forward		2.19	0.79	0.69	5.14	3.53	3.47
LARS		2.54	0.81	0.73	5.16	3.46	3.37
Ridge		0.95	0.91	0.88	4.38	4.22	4.29
LASSO		2.26	0.74	0.74	4.91	3.49	3.41
Elastic Net		0.99	0.74	0.72	4.00	3.58	3.52

TABLE 2.7 – Erreur d'estimation des modèles d'estimation du niveau de dégradation d'une batterie, pour différents taux d'apprentissage - Les meilleurs résultats sont affichés en gras ($err < 0.75\%$ pour la capacité, et $err < 3.50\%$ pour la résistance)

Notons que les méthodes étudiées présentent des performances différentes. Ainsi, pour l'estimation du niveau de capacité, les méthodes *ridge* et *elastic net* fournissent les résultats les plus réguliers, quelque soit le taux d'apprentissage considéré, mais peu performant par rapport aux autres méthodes pour le *ridge* avec un taux d'apprentissage de 50 %. L'estimation de la résistance met, elle, en évidence les méthodes LASSO et LARS avec un haut taux d'apprentissage, et la méthode *elastic net* en lors de la présence d'un faible taux d'apprentissage. L'obtention des données étant complexe, il est préférable de privilégier les méthodes proposant des résultats satisfaisants avec peu d'apprentissage. Ainsi, la méthode *elastic net* semble présenter le meilleur compromis pour l'estimation de la capacité, et le LASSO pour l'estimation de la résistance.

D'autre part, les méthodes explicitées n'utilisent pas toutes les variables lors de la construction d'un modèle, ce point est ici particulièrement intéressant dans le contexte du vieillissement des batteries. Cela permet d'identifier les variables liées aux phénomènes de vieillissement. La table 2.8 liste les variables sélectionnées par quelques modèles : *elastic net* et LASSO pour l'estimation de capacité et de résistance.

Méthode	Capacité				Résistance			
	Elastic Net		LASSO		Elastic Net		LASSO	
	25	50	25	50	25	50	25	50
Taux d'app. (%)								
Variables								
SOC_{cal}	✓	✓	✓	✓	✓	✓	✓	✓
T_{cal}	✓	✓	✓	✓	✓	✓	✓	✓
$Distance$	✓	✓	✓		✓	✓	✓	✓
$E_{bat,sorti}$	✓	✓	✓			✓	✓	✓
T_{bat}		✓	✓	✓		✓	✓	✓
ΔSOC		✓	✓	✓		✓	✓	✓
t_{cal}		✓	✓	✓	✓		✓	✓
Nb_{altI}		✓		✓	✓	✓	✓	✓
T_{amb}		✓					✓	✓
Nb_{cycles}		✓				✓	✓	
PKE					✓	✓	✓	✓
I_{max}							✓	✓
$Vitesse$								

TABLE 2.8 – Liste des variables sélectionnées par les modèles d'estimation

Le point intéressant est le faible nombre de variables utilisées dans le cadre de l'estimation du niveau de capacité, en particulier pour la méthode *elastic net*. Les variables conservées proviennent des stockages et des températures vécues par la batterie. Au contraire, l'estimation de la résistance est plus complexe et nécessite plus d'information. Les variables utilisées par les modèles d'estimation de la résistance sont issues de toutes les conditions de vie de la batterie (mode d'Utilisation, environnement).

Il est important ici de noter que ces modèles sont bâtis à partir de données correspondant à une semaine, ce qui explique l'absence de la notion de durée d'expérimentation dans ces résultats. Toutes ces interprétations confirment les points démontrés en section 2.3.2.2 concernant les variables impactant les phénomènes de vieillissement.

À titre d'exemple, le modèle normalisé *elastic net* basé sur un taux d'apprentissage de 25 % s'écrit :

$$\hat{Y}_{cap}^{elasticnet} = -1.553 \cdot T_{cal} + 1.157 \cdot SOC_{cal} + 0.221 \cdot Distance - 0.035 \cdot E_{bat,sorti} \quad (2.21)$$

Cette équation 2.21 fournit une estimation $\hat{Y}_{cap}^{elasticnet}$ de l'évolution de la capacité (en pourcentage) de la batterie lors d'une semaine d'expérimentation, à partir de quatre valeurs normalisées : T_{cal} , SOC_{cal} , $Distance$ et $E_{bat,sorti}$. Les coefficients associés à ces variables permettent de quantifier l'impact de celles-ci sur l'évolution de la capacité déterminée par ce modèle. Cet exemple propose ainsi une prédominance des conditions de stockages T_{cal} et SOC_{cal} dans l'évolution estimée de capacité.

Un résumé des principaux paramètres détectés par cette étude, comme influençant les dégradations d'une batterie, en accord avec les conclusions faites précédemment, sont :

- Évolution de capacité : stockage (SOC_{cal}, T_{cal}), température de la batterie (T_{bat})
- Évolution de résistance : mode d'Utilisation de la batterie ($PKE, I_{max...}$), conditions extérieures (T_{amb}), ...

La capacité semble donc principalement liée aux conditions de stockage et la résistance aux mode d'Usage de la batterie. Notons que dans les deux cas les conditions environnementales interviennent.

L'extension de cette étude consiste à mettre en place cette méthodologie pour procéder aux estimations sur de nouvelles données. Considérons une phase de test postérieure à la phase d'apprentissage. Dans un tel cas, les données d'apprentissage proviennent d'une première partie de la vie de batterie, et sont employées pour la construction d'un modèle d'estimation. Par la suite, ce modèle est évalué sur les données tests. Cette expérimentation est semblable à ce qu'un constructeur pourrait proposer en construisant un modèle sur une phase expérimentale préalable à la mise en service d'une batterie. Il peut également être envisageable de construire un modèle sur une batterie puis d'appliquer ce modèle sur une autre batterie ayant les mêmes propriétés.

Nous appliquons cette approche sur les données de la batterie Bat_1 . Les modèles d'estimation de capacité obtenus après 54 semaines d'apprentissage, et après 66 semaines pour l'évolution de résistance sont illustrés en figure 2.22 et figure 2.23.

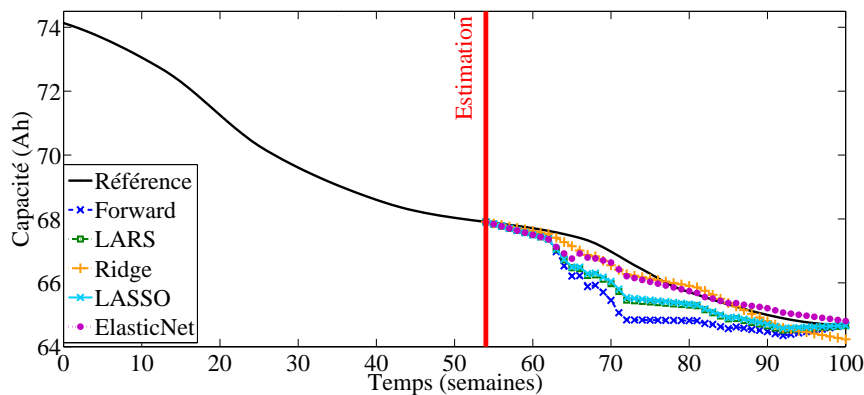
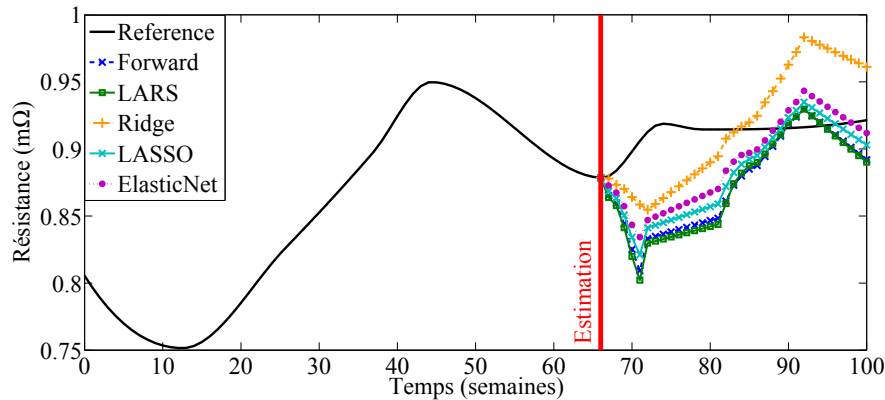


FIGURE 2.22 – Estimation de l'évolution de capacité de la batterie Bat_1

Ces résultats confirment les bonnes performances fournies par les modèles lors de l'estimation du niveau de capacité d'une batterie. En effet, les estimations ne divergent pas et restent très proches de la valeur référence, produisant pour les méthodes *ridge* et *elastic net* une erreur d'estimation moyenne inférieure à $err = 1\%$. L'estimation de la résistance d'une batterie est, quant à elle, plus difficile à calculer et les résultats obtenus à partir d'une phase préalable d'apprentissage (figure 2.23) démontrent un besoin plus important en données d'apprentissage pour parvenir à des résultats similaires à l'estimation de capacité. L'amélioration des estimations de résistance requiert la prise en compte de plus diverses données d'apprentissage couvrant l'ensemble des conditions impactant la résistance. En effet, l'évolution de la résistance d'une batterie

FIGURE 2.23 – Estimation de l'évolution de résistance de la batterie Bat_1

est un phénomène complexe et dépendant de conditions d'Utilisation plus variées que l'évolution de la capacité.

2.5 Déduction du temps de vie restant d'une batterie

2.5.1 Méthodologie

À partir des estimations d'évolution de capacité, il est utile d'un point de vue utilisateur de déduire le temps de vie restant (RUL) avant d'atteindre le critère de fin de vie EOL. Ce dernier est ici considéré comme étant le seuil défini par 80 % de capacité nominale de la batterie. Le temps de vie restant d'une batterie lors des Usages est important, car il s'agit du critère utilisé par les constructeurs pour mettre en place les garanties des batteries, et d'un indicateur référence d'autonomie du véhicule pour les utilisateurs. Il est alors intéressant de pouvoir obtenir, à un instant donné une prédiction \widehat{RUL} , du temps de vie restant d'une batterie.

Pour ce faire, à un instant t , nous procédons par le biais d'un *bootstrap* sur les données d'apprentissage X^{app} , qui sont dans notre cas les données précédant l'instant t . Le *bootstrap* consiste ici en un tirage avec remise des individus d'apprentissage.

Cette étape permet de générer B ensembles d'apprentissage X_B^{app} , résultant du *bootstrap*. Chacun de ces B ensembles fournit une estimation des paramètres β du modèle 2.16, pour une méthode d'estimation donnée, produisant un intervalle de confiance IC associé. La moyenne de ces paramètres $\{\beta_1, \dots, \beta_B\}$ génère à un instant t , un intervalle de confiance moyen IC_{moy} autour des estimations de capacité obtenues par les méthodes précédemment citées.

Les bornes supérieures et inférieures de l'intervalle de confiance IC_{moy} sont donc connues jusqu'à l'instant t de la prédiction. Ces deux bornes sont extrapolées après l'instant t linéairement. En détectant l'instant où ces extrapolations franchissent le seuil de fin de vie EOL, on obtient une prédiction par le biais une densité de probabilité du temps de vie restant. Cette étape est illustrée par la figure 2.24.

Par extension, nous pouvons envisager cette méthodologie à partir de données d'apprentissage provenant d'une autre batterie de mêmes caractéristiques, permettant de mettre en place des modèles produisant l'intervalle de confiance.

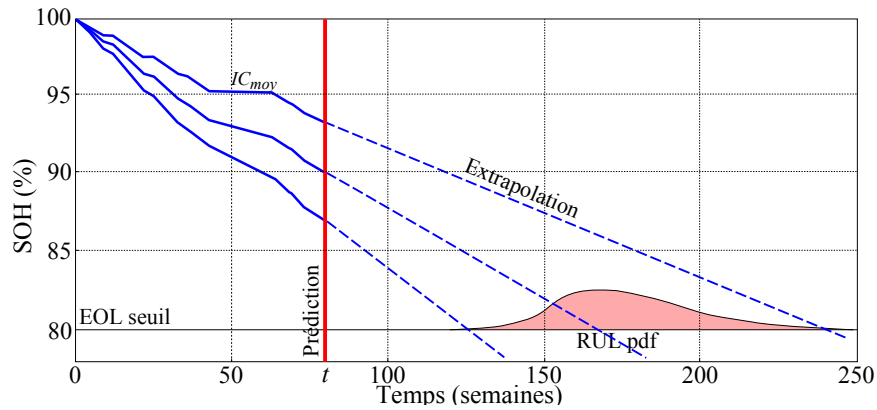


FIGURE 2.24 – Illustration de la construction de la Probability Density Function (*pdf*) du RUL, à partir de l'extrapolation de l'intervalle de confiance moyen IC_{moy} du SOH, à un instant t

2.5.2 Application et résultats

Cette méthodologie est mise en application sur les données de la batterie Bat_1 . La figure 2.25 illustre les fonctions de densité du RUL obtenues à trois instants différents (après 45, 70 et 100 semaines), sur cette batterie Bat_1 .

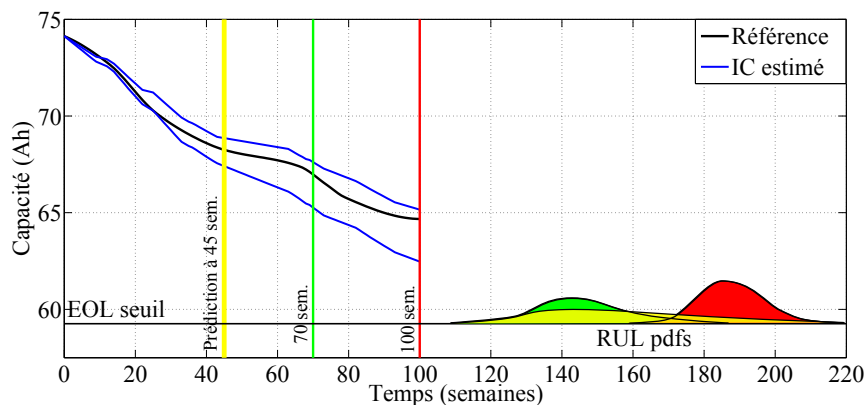


FIGURE 2.25 – Prédiction de la fonction de densité *pdf* du temps de vie restant RUL pour une batterie en usage réel, à trois instants différents avec la batterie Bat_1

Ces résultats soulignent la sensibilité des densités de fin de vie prédites, en fonction des variations de l'intervalle de confiance construit. En effet, plus l'instant de la prédiction est proche de la fin de vie, plus la prédiction du RUL est précise, ce qui est visible dans la figure 2.25 par l'évolution de la forme de chacun des trois *pdfs*. Nous observons la réduction de la plage d'incertitude lorsque l'instant de la prédiction augmente, ce qui est expliqué par la plus grande proximité de la fin de vie, et une meilleure connaissance de la batterie.

D'autre part, la prédiction du RUL est dépendante de la tendance de l'intervalle de confiance. La figure 2.26 représente l'évolution des prédictions de fin de vie faites en fonction de l'instant considéré. Cette figure 2.26 met en évidence ce phénomène

en présentant les différentes densités obtenues pour des estimations faites entre les semaines 60 et 100, illustrant en particulier les différences de densité de probabilité en fonction des dernières tendances de capacité observées qui sont propagées lors de l'extrapolation.

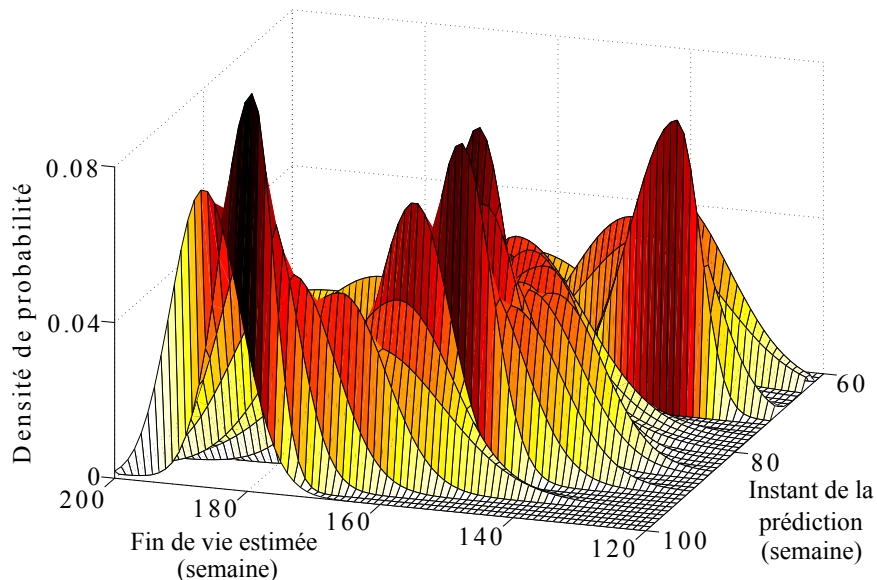


FIGURE 2.26 – Densité de probabilité du RUL pour différents instants de prédiction pour la batterie Bat_1

Tous ces résultats représentent la modification des prédictions de RUL au fil des semaines considérées. Les prédictions faites entre les semaines 70 et 90 induisent une fin de vie aux alentours de la semaine 160, alors que les dernières estimations faites (semaine 100) fournissent une prédiction de fin de vie vers la semaine 180. Cette modification de quelques semaines est néanmoins légère si on considère le temps restant d'utilisation de la batterie, les dernières prédictions équivalant à une période d'utilisation restante d'environ un an et demi.

2.6 Conclusions - Discussions

L'objectif de ce chapitre est d'utiliser les informations contenues dans les données mesurées lors de la vie d'une batterie pour identifier et quantifier l'impact de variables sur les phénomènes de vieillissement.

La première partie de ce chapitre consiste en une présentation détaillée du matériel et des expérimentations mises en places ayant permis l'obtention des données traitées tout au long de cette étude.

Par la suite, l'approche utilisée dans ce chapitre a identifié les variables mesurées étant susceptibles de contenir de l'information relative aux phénomènes de vieillissement, tout en évitant les redondances d'information. Cette étude a dans un premier temps validé une suppression de nombreuses mesures faites, qui ne procureraient aucun apport d'information vis à vis de notre problématique du vieillissement des bat-

teries. D'autre part, il est primordial de considérer d'autres valeurs non directement mesurées, telles que les conditions et durées de stockage ou bien des indicateurs inhérents au mode d'utilisation de la batterie ou au style de conduite appliqué. Des variables correspondantes à ces critères sont calculées et ajoutées à l'étude, offrant un maximum d'informations liées à la durée de vie de la batterie.

À partir de cette sélection préliminaire de variables, nous avons procédé à une analyse approfondie de ces données. Cette analyse a mis en évidence les relations entre les variables et les phénomènes de vieillissement d'une batterie. Comme décrit précédemment, le vieillissement d'une batterie se caractérise à la fois par une diminution de la capacité et une augmentation de la résistance. L'étude présentée dans ce chapitre identifie les causes de chacune de ces mécanismes globaux de vieillissement. L'évolution de capacité étant principalement reliée aux conditions de stockage et l'évolution de résistance associée aux modes d'utilisation, chacun dépendant des conditions environnementales.

Les facteurs reliés à l'évolution de la capacité d'une batterie sont les variables provenant des conditions de stockage et de température vécues par la batterie. Ces résultats permettent de déduire les usages responsables d'une diminution importante de la capacité d'une batterie ; à savoir des conditions de stockages contraignantes ou des températures élevées ($> 40^{\circ}\text{C}$) mesurées sur la batterie lors des cyclages.

D'autre part, l'évolution de résistance d'une batterie est identifiée comme dépendante des conditions d'Usage de la batterie. Cela revient à dire que l'augmentation de la résistance est causée par des Usages intensifs ou contraignants de la batterie. Ce type d'Usage peut par exemple être un type de conduite *agressif* (PKE, I_{max}, Nb_{altI}), ou bien par des cyclages longs et fréquents ($Distance, Nb_{cycles}$).

Ces conclusions sont en accord avec les propriétés des phénomènes de vieillissement décrits dans le chapitre 1, et sont donc validées par les approches physiques et électrochimiques, démontrant la pertinence d'une approche uniquement basée sur les données. Notons que ces interprétations des causes du vieillissement d'une batterie proposent une aide aux préconisations d'utilisations, évitant des dégradations prématurées et les conditions d'usage contraignantes. De plus la quantification des impacts respectifs des variables sur les phénomènes de vieillissement offre un outil d'interprétation intéressant susceptible d'être utilisé par des experts du domaine.

L'étude des données issues des expérimentations réelles à également permis de construire des modèles d'estimation linéaire à la fois du SOH et du niveau de résistance de la batterie. Les constructions de ces modèles parcimonieux confirment les remarques faites précédemment sur les relations entre des facteurs responsables des deux phénomènes de vieillissement.

L'estimation du niveau de résistance est difficile du fait des dynamiques mesurées plus complexes et du nombre plus important de variables intervenant dans ce phénomène de vieillissement. Les erreurs d'estimations sont néanmoins relativement faibles ($err \sim 3.5\%$) considérant le caractère peu limitant de la valeur de résistance pour une batterie en usage sur un véhicule électrique.

Les résultats d'estimation du niveau de capacité démontrent un niveau d'erreur faible ($err < 1\%$), permettant l'application d'une telle méthodologie dans un contexte d'estimation de capacité sur une batterie instrumentée en usage réel, seulement à partir d'un modèle préalablement conçu. Ce niveau de précision fournit une estimation de capacité fiable sans nécessiter de protocole particulier et lourd à mettre en œuvre.

En effet, les modèles évoqués n'utilisent que des données descriptives des usages, hors ligne, proposant un suivi global, ce qui démontre la compatibilité de ces outils avec les enregistrements limités des données.

Cette méthodologie d'estimation de capacité présentant des résultats intéressants, une extension statistique a été implémentée prédisant la durée de vie restante (RUL) d'une batterie. Une telle approche présente une bonne confiance dans la prédiction de la fin de vie d'une batterie, la valeur moyenne du RUL estimée ne diffère que de quelques semaines en fonction de l'instant de la prédiction. Ce résultat permet d'obtenir une information importante pour un constructeur sur la durée restante d'utilisation d'une batterie, en fonction de son mode d'Utilisation. En effet, le critère de diminution de capacité restante est très restrictif pour l'emploi des véhicules électriques. Cet outil accorde une possibilité de gestion des usages, en appliquant des préconisations, afin de maximiser la durée de vie d'une batterie.

Ce chapitre a démontré l'intérêt d'une approche basée sur l'analyse des données, détectant les causes de vieillissement et allant jusqu'à obtenir des estimations des phénomènes de vieillissement d'une batterie sans protocole expérimental complexe.

Chapitre 3

Analyse en ligne des données d'une batterie

3.1	Problématique de l'analyse en ligne	84
3.2	Étude de signaux issus de mesures embarquées	84
3.2.1	Analyse des signaux	84
3.2.2	Extraction de motifs	87
3.2.3	Interprétation visuelle	90
3.3	Approche de classification de l'état de santé d'une batterie	92
3.3.1	Choix d'une méthodologie de comparaison des signaux adaptée	93
3.3.1.1	Distance euclidienne	93
3.3.1.2	Recalage des signaux	94
3.3.1.3	Dynamic Time Warping	96
3.3.1.4	Contraintes de Sakoe & Chiba	98
3.3.1.5	Global Alignment	98
3.3.1.6	Application à l'étude des motifs	99
3.3.2	Méthodes de classification	99
3.3.2.1	k-Nearest Neighbors	100
3.3.2.2	k-Means	100
3.3.2.3	Classification Ascendante Hiérarchique	101
3.3.3	Résultats des méthodes de classification	102
3.3.3.1	Approche sur la batterie Bat_1	102
3.3.3.2	Application aux accélérations de 20 à 40 km.h ⁻¹	103
3.3.3.3	Application aux accélérations de 10 à 60 km.h ⁻¹	104
3.3.3.4	Influence des critères d'extraction	105
3.3.3.5	Extension à 8 classes	106
3.4	Vers une estimation continue de l'état de santé d'une batterie	107
3.4.1	Méthodologie	107
3.4.1.1	Relevance Vector Machines	107
3.4.1.2	Extension du RVM aux noyaux multiples	108
3.4.1.3	Noyaux utilisés	108
3.4.2	Résultats	109
3.5	Déduction du temps de vie restant d'une batterie	112
3.5.1	Méthodologie	112
3.5.2	Application et résultats	114
3.6	Conclusion - Discussion	117

3.1 Problématique de l'analyse en ligne

Dans le chapitre 2, l'analyse hors ligne des données collectées sur un ensemble de véhicules a démontré son habileté dans la mise en évidence des causes des phénomènes de vieillissement d'une batterie, avec le minimum d'a-priori possible sur les causes de vieillissement. De plus, les modèles d'estimations de capacité bâtis avec cette approche ont montré une bonne faculté d'estimation en considérant simplement un historique des données d'Utilisation. Cette approche permet d'obtenir une estimation à des intervalles de temps réguliers (chaque semaine dans les exemples du chapitre 2), mais ne propose pas d'estimation en ligne. Une autre limite de cette méthodologie réside dans la difficile amélioration des précisions obtenues. En effet, les performances de cette catégorie de modèles ne peuvent être améliorées que par l'ajout des données d'apprentissage, ce qui requiert de nouvelles expérimentations en usage réel. Cet inconvénient induit une augmentation importante des coûts d'expérimentation pour obtenir une estimation précise de l'état de santé d'une batterie.

Ces inconvénients d'une analyse hors ligne des données nous amènent à envisager une approche complémentaire à la méthodologie développée au chapitre 2. Ainsi, nous nous plaçons ici dans un contexte d'une analyse en ligne ou en embarqué des données mesurées par une batterie en usage sur un véhicule électrique. La problématique considérée à présent consiste à étudier les mesures en temps réel, sans prendre en compte un quelconque historique. L'objectif est toujours d'obtenir des informations sur le niveau de dégradation d'une batterie utilisée par un véhicule électrique.

Cette approche uniquement basée sur les données enregistrées directement sur la batterie, sans prise en compte d'un modèle physique au contraire de certaines approches [96, 120], présente l'avantage d'être implémentable en temps réel et en embarqué sur un véhicule, avec un faible stockage des données.

Le seul phénomène de vieillissement ici pris en compte est la capacité de la batterie. En effet, nous avons évoqué précédemment les caractéristiques de vieillissement des technologies de batteries utilisées sur les véhicules électriques, or l'augmentation de la résistance ne présente qu'un faible impact sur l'usage lors d'une batterie sur un véhicule électrique (section 1.1.2). Autrement dit, la principale dégradation limitante lors des Usages d'une batterie sur un véhicule, tel que ceux actuellement utilisés, est la perte de capacité d'une batterie.

L'étude présentée dans ce chapitre se concentre donc sur l'analyse des signaux mesurés en temps réel et l'extraction automatique des informations qu'ils contiennent, relatives à l'état de santé de la batterie.

3.2 Étude de signaux issus de mesures embarquées

Considérons ici les signaux issus des mesures acquises lors des Usages de la batterie. Cette section détaille les réflexions et analyses menant à la méthodologie présentée dans la suite de ce chapitre.

3.2.1 Analyse des signaux

Chacune des variables est mesurée en temps réel lors des Usages, à une fréquence de 10 Hz, ce qui permet l'obtention de nombreux détails sur leurs comportements et

évolutions. En effet, cette fréquence d'échantillonnage est plus rapide que les évolutions des phénomènes physiques sous-jacents. Afin d'illustrer la forme des signaux pouvant être utilisés dans cette approche, la figure 3.1 expose un exemple de signaux obtenus lors d'un roulage de la batterie *Bat*₂.

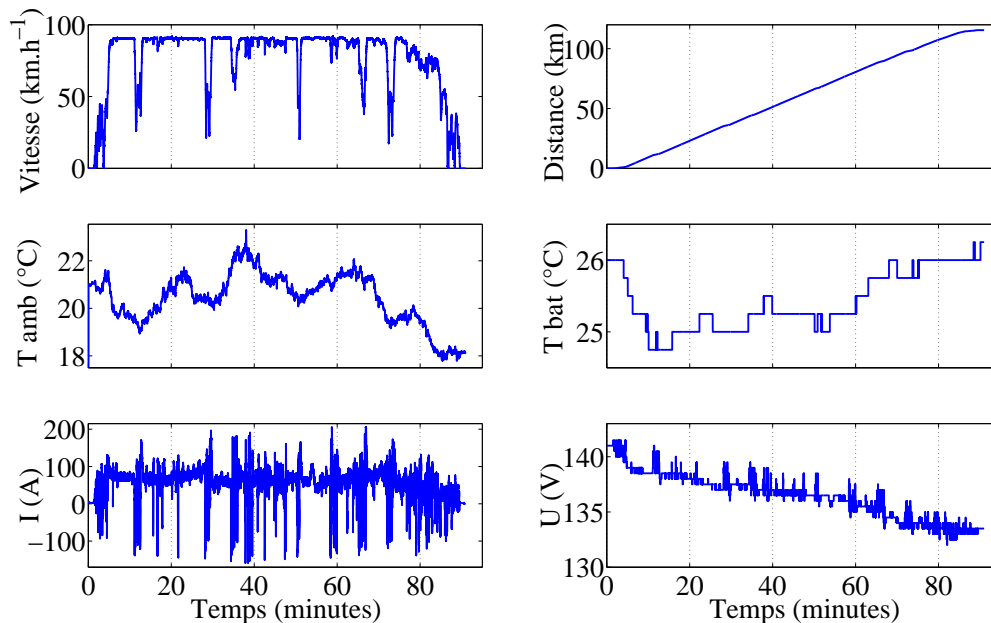


FIGURE 3.1 – Exemple des mesures réelles obtenues lors d'un roulage de la batterie *Bat*₂, concernant le véhicule (vitesse, distance), les conditions climatiques (température ambiante, température batterie), et le cyclage de la batterie (courant, tension)

La figure 3.1 illustre clairement les différentes dynamiques de chacune des variables mesurées en temps réel lors des expérimentations. Ainsi, les changements de température ambiante sont relativement lents et le comportement de la température de la batterie est relativement stable comparativement à la fréquence d'échantillonnage. Les signaux de vitesse, de courant et de tension enregistrés sont, eux, plus irréguliers et donc plus complexes à interpréter.

Les phénomènes de vieillissement, et en particulier l'évolution de capacité, impliquent une modification électrochimique du comportement d'une batterie. Les signaux représentatifs du comportement d'une batterie et de sa capacité à fournir une énergie demandée sont les plus pertinents dans cette analyse. Nous concentrons donc la suite de l'étude sur l'investigation des informations contenues dans les signaux de courant et de tension mesurées directement au niveau de la batterie. Ces deux variables sont en effet les plus représentatives du comportement de la batterie.

Ces deux variables de courant et tension sont associées à des signaux contenant de nombreuses variations (figure 3.1), laissant envisager des perspectives d'estimation et de diagnostic de l'état de santé d'une batterie, basées sur l'analyse du comportement de ces deux signaux.

L'objectif est d'identifier les informations relatives à l'état de santé d'une batterie, uniquement à partir de ces signaux de courant et tension mesurés en sortie de batterie. Pour ce faire, il est nécessaire de définir un critère de comparaison des signaux au cours du temps, et donc en fonction de l'état de santé de la batterie.

Les conséquences d'une modification de l'état de santé d'une batterie sur son comportement, ont été décrites dans le chapitre 1. Ainsi, il apparaît que d'éventuelles modifications dans les signaux sont susceptibles d'être détectées lors d'une irrégularité ou d'un changement brusque de sollicitation. Ces conditions correspondent en effet à une sollicitation particulière de la batterie, qui peut permettre d'identifier une altération de comportement de cette batterie.

La problématique réside désormais dans la définition précise d'un critère de comparaison des signaux, et de la méthodologie choisie pour permettre d'obtenir des informations sur l'état de santé de la batterie. L'objectif est de comparer les réactions de la batterie pour une même demande, en fonction de son état de santé. Nous considérons donc l'hypothèse du vieillissement impactant directement l'énergie sortant de la batterie, pour une même sollicitation.

L'expérimentation de la batterie Bat_1 est particulièrement adaptée pour cette analyse du fait des profils de vitesses récurrents, car faite sur un circuit (section 2.2.1). La récurrence de tels profils durant la vie de la batterie permet de valider l'hypothèse de modification du comportement d'une batterie. La figure 3.2 présente un comparatif des signaux de vitesse, courant et de tension mesurés lors de trois roulages espacés dans le temps, correspondant chacun à un niveau de santé différent de la batterie Bat_1 .

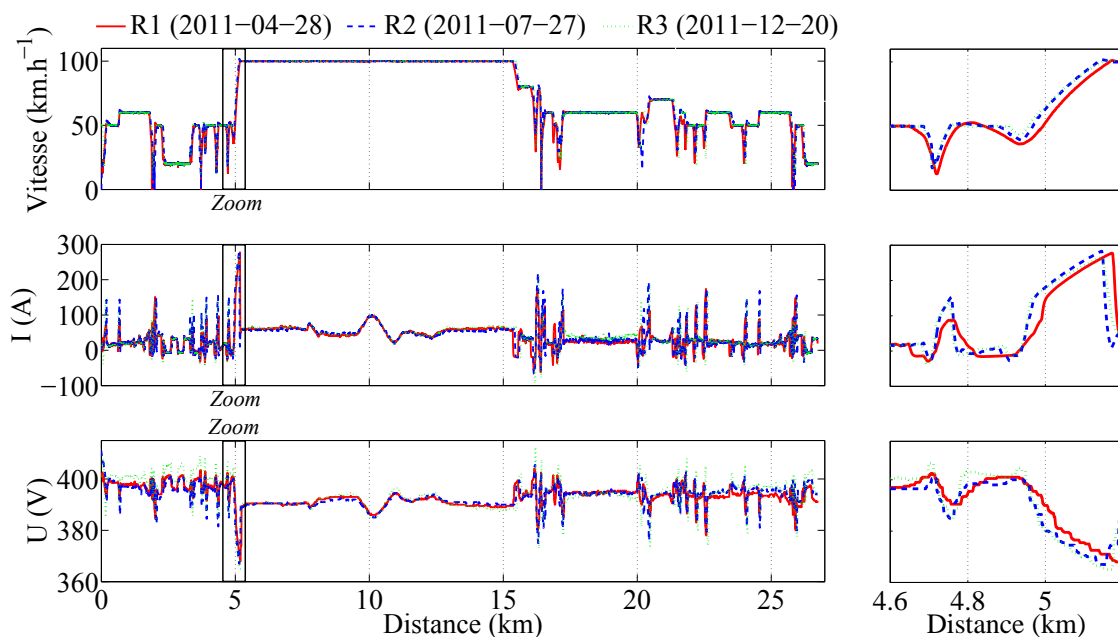


FIGURE 3.2 – Signaux obtenus pour trois roulages similaires (R1, R2, R3) avec la batterie Bat_1 , pour des états de santé différents, accompagné d'un zoom sur une courte période

L'étude de ces signaux démontre la constance des profils vitesses et donc des sollicitations appliquées à la batterie Bat_1 , avec des conditions environnementales extérieures similaires du fait de l'utilisation d'un circuit privé. Ces expérimentations sont idéales pour permettre la comparaison des signaux mesurés sur la batterie. Il appa-

rait des comportements différents pour chacun des signaux provenant de la batterie. En se focalisant sur les profils de courant et de tension, ils paraissent en effet moins semblables entre eux que les signaux de vitesse. Néanmoins, il n'est pas possible visuellement de mettre en évidence et d'interpréter une modification des dynamiques de courant ou de tension à partir du signal entier.

Il est important de noter ici que les études électrochimiques ne permettent pas de justifier totalement ce type de modification de dynamique. Dès lors, la suite de cette approche nécessite alors une investigation plus approfondie de ces signaux pour la détection d'information sur l'état de santé de la batterie utilisée et la justification de la modification de dynamique observée dans les signaux courant et tension.

3.2.2 Extraction de motifs

Dans un premier temps, il est nécessaire d'explorer les signaux de courant et de tension, enregistrés en temps réel, dans leur globalité. Ces signaux sont complexes et contiennent notamment des dynamiques variées, ce qui requiert une investigation précise pour détecter des modifications de comportement.

Les études des signaux dans leur globalité ne permettent pas de mettre en évidence des différences entre les signaux d'un point de vue global, et les modifications de dynamique demeurent imperceptibles. En effet, aucune distinction n'a pu être mise en évidence par ce type d'approche. En considérant les propriétés électrochimiques des batteries Li-ion, les phénomènes de dégradation se traduisent par une modification du comportement de celles-ci. Afin de détecter de telles modifications, nous concentrons la suite de l'étude sur des dynamiques spécifiques de la batterie. Ces dynamiques correspondent alors à une sollicitation particulière et précise de la batterie pouvant contenir des informations sur son état de santé.

Le choix des dynamiques étudiées est primordial dans la détection d'une altération du comportement d'une batterie. Il peut être envisagé de choisir comme critère les changements de signe de courant, ce qui correspond aux instants où la batterie passe d'une décharge à une récupération d'énergie (ou inversement). Ce phénomène est donc susceptible de contenir des informations relatives au niveau de santé de la batterie, de part la dynamique du courant perçue à ce moment précis. D'un point de vue électrochimique celui-ci se traduit alors par la dynamique associée au changement de direction des éléments chargés, au sein de la batterie. Les signaux de courant, correspondant à ce choix de dynamique, obtenus lors du changement de signe de courant de la batterie Bat_1 sont présentés, pour trois niveaux d'état de santé, en figure 3.3. Notons que ces motifs sont extraits à partir des signaux initiaux, lors de trois journées de tests différentes, permettant la comparaison des dynamiques pour trois états de santé.

Cette figure 3.3 illustre une très large diversité des motifs extraits et des critères de comparaison complexes, ce qui ne permet pas d'interprétation spécifiquement adaptée à l'investigation de l'état de santé d'une batterie.

Une autre période de transition est alors envisagée, correspondant aux instants d'accélération ou de freinage maximal du véhicule électrique. Cette situation revient à analyser les périodes de sollicitations maximales de la batterie. Or, les dégradations d'une batterie sont susceptibles de provoquer des limitations de comportement pour ce type de sollicitations. Nous avons donc extrait des motifs de courant associés à des accélérations et des freinages importants provenant de la batterie Bat_1 (figure 3.4).

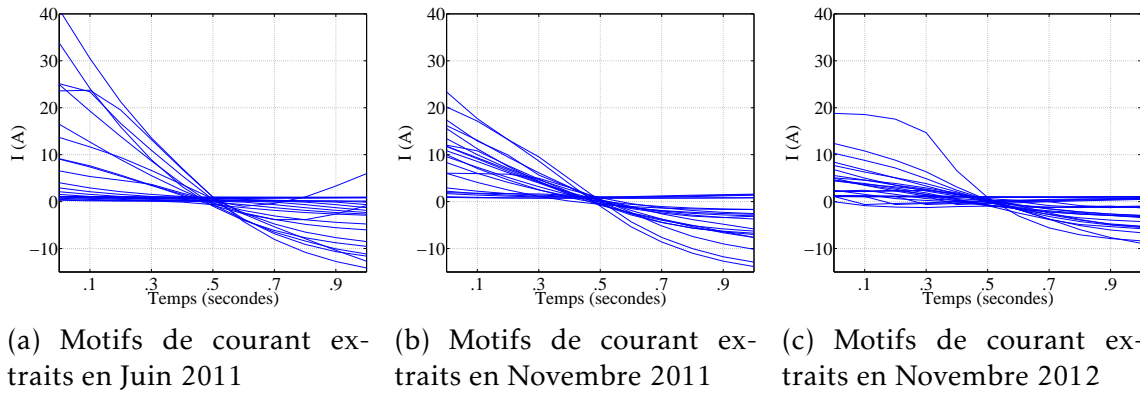


FIGURE 3.3 – Illustration des motifs de courant extraits lors changement de signe associé au passage d'une décharge à une phase de récupération d'énergie - lors de trois journées différentes avec la batterie Bat_1

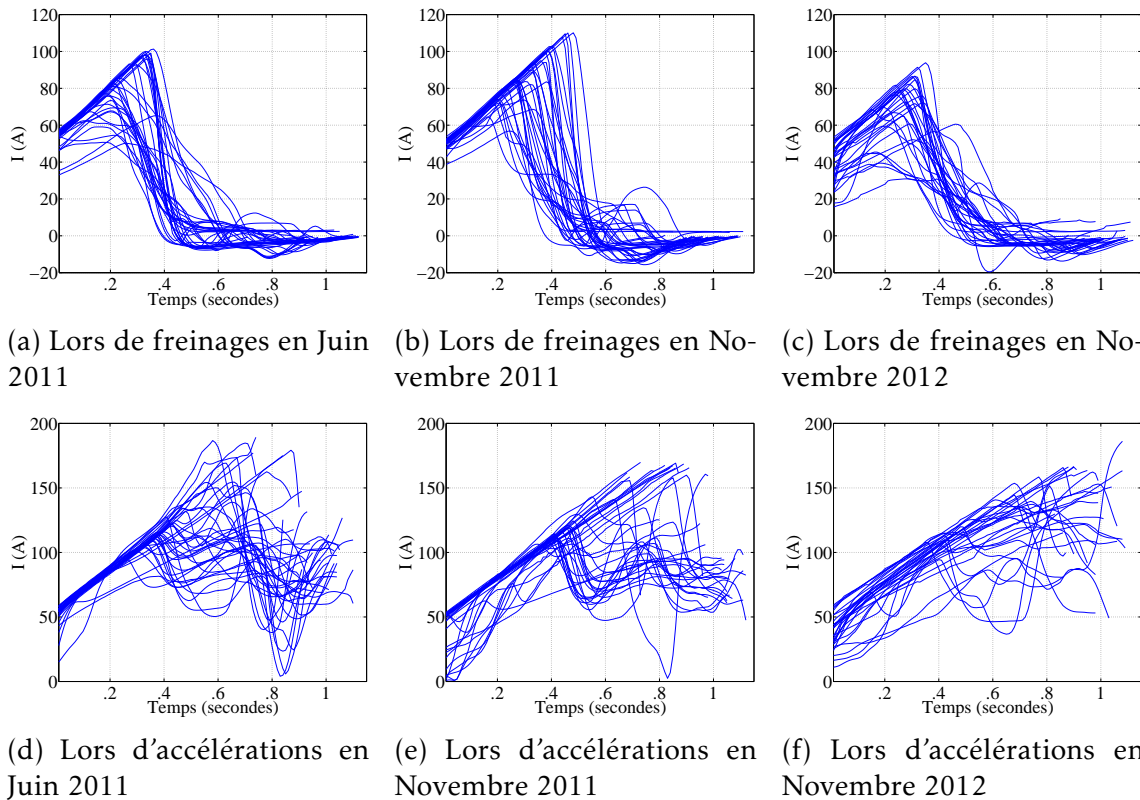


FIGURE 3.4 – Motifs de courant extraits correspondant selon les mêmes critères d'accélération et de freinages, à trois dates différentes avec la batterie Bat_1

Cette figure 3.4 met en évidence une large variété de dynamique entre les motifs de courant obtenus pour chacune des dates étudiées, pour une même évolution de vitesse. Cependant ces motifs extraits coïncident avec une même tendance globale avec des modifications de comportement selon les motifs. Ainsi, cette exploration suggère des modifications de comportement des signaux issus de la batterie lors des usages réels. La suite de l'étude consiste alors à utiliser ces motifs afin d'explorer les informations qu'ils contiennent pour obtenir des renseignements sur l'état de santé de la batterie

analysée.

Ces motifs sont extraits aux mêmes instants du profil de roulage spécifique appliqué à la batterie Bat_1 . Or, il est nécessaire de définir précisément des critères d'extraction de ces motifs pour appliquer cette approche à tout type de roulage. Ainsi, nous définissons des seuils sur la durée $S_t = (t_{min}, t_{max})$ et la vitesse $S_v = (v_{init}, v_{term})$, formalisant l'extraction des motifs souhaités. Cela revient à ne conserver que les phases d'accélération (ou de freinage) pour lesquelles la vitesse passe de v_{init} à v_{term} dans une durée comprise entre t_{min} et t_{max} . La figure 3.5 illustre la définition de ces seuils à la base cette méthodologie d'extraction.

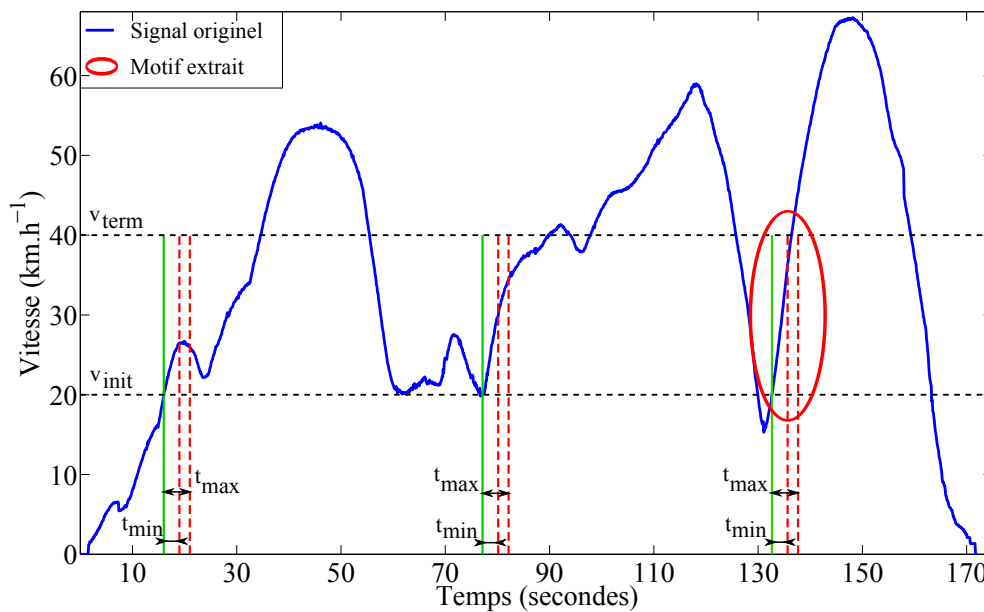


FIGURE 3.5 – Méthodologie d'extraction d'un motif d'accélération répondant aux critères de vitesse $(v_{init}, v_{term}) = (20, 40) \text{ km.h}^{-1}$, et de durée $(t_{min}, t_{max}) = (5, 7)$ secondes

Les seuils définissant les critères d'extraction sont définis selon les caractéristiques de la batterie et du véhicule pour correspondre à une accélération (ou un freinage) maximal, selon les performances de la batterie étudiée. Ces caractéristiques sont propres à chaque batterie, en fonction de ses performances optimales. De ce fait, le choix de ces critères requiert une définition externe préalable. Nous discuterons dans les prochaines sections l'influence des choix de ces seuils. Se basant sur les motifs de vitesse extraits, il est primordial pour l'étude des réactions de la batterie en ces instants d'extraire également les motifs de courant et de tension associés à ces intervalles. Ainsi, nous obtenons une méthodologie d'extraction de motifs courant et tension, correspondant à des profils de vitesses similaires. La figure 3.6 illustre les résultats d'une extraction faite sur la batterie Bat_3 , de motifs répondant à des accélérations de 20 à 40 km.h^{-1} lors d'une période comprise entre 3 et 5 secondes.

Après l'investigation des signaux obtenus à partir des usages réels de batteries, nous avons choisi d'étudier les motifs correspondant aux accélérations (ou aux freinages) maximales. L'extraction de tels motifs, au sein des signaux mesurés en temps réel, se fait d'après la méthodologie décrite dans cette section. Celle-ci dépend de deux

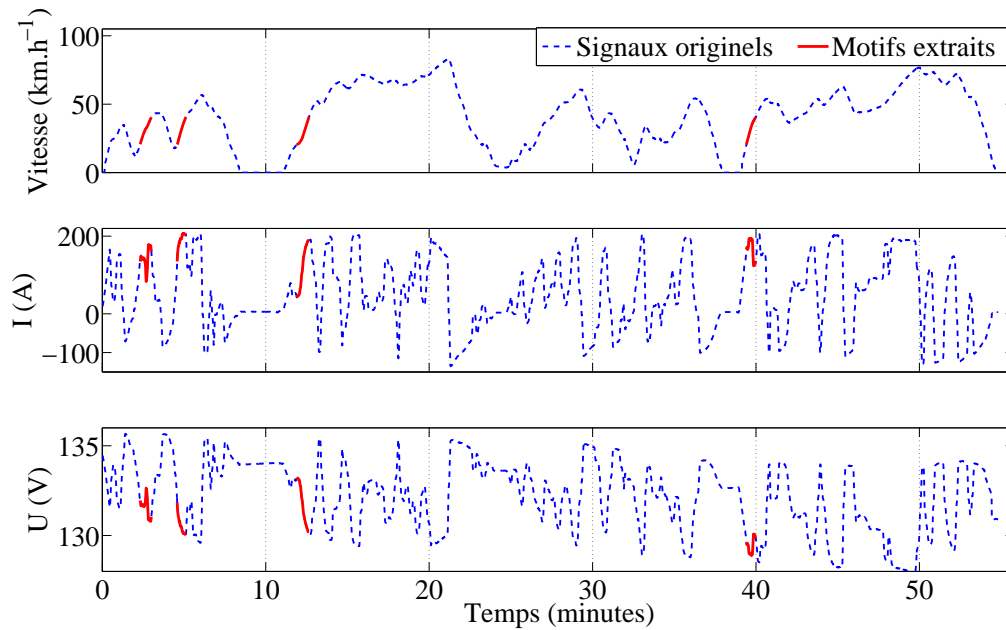


FIGURE 3.6 – Exemple d'extraction de motifs courant et tension à partir du signal de vitesse de la batterie Bat_3 - en considérant les accélérations de 20 à 40 km.h^{-1} entre 3 et 5 secondes

seuils prédéfinis (S_t et S_v) sur le signal de vitesse. Les motifs de courant et de tension associés à ces instants sont également extraits, fournissant une indication sur le comportement de la batterie pour cette sollicitation précise. L'objectif est désormais de pouvoir interpréter ces motifs et ainsi de permettre d'identifier d'éventuelles relations entre ces motifs et l'état de santé de la batterie à l'étude.

Notons que les motifs présentés dans les résultats sont issus des tests de la batterie Bat_1 , où les mêmes sollicitations sont répétées à chaque roulage. Cette expérimentation permet d'éviter de potentiels effets dus aux conditions extérieures (pente, type de route...). Les seules variables non contrôlées sont alors dans ce cas la température ambiante et le niveau de charge de la batterie. Or le même type d'observation sur la modification des motifs peut être menée sur les autres batteries. La suite de l'étude étend cette méthodologie aux données provenant des autres batteries.

3.2.3 Interprétation visuelle

Nous étudions ici des motifs provenant d'extractions ayant eu lieu tout au long de la vie de la batterie. Cette méthodologie fournit un grand nombre de motifs qui doivent ici être analysés et interprétés. Le but est de pouvoir visualiser les évolutions de profil de ces motifs, en fonction de l'état de santé de la batterie. Nous avons introduits dans ce but l'algorithme Filtered Shape Averaging (FSA) permettant de résumer un ensemble de motifs par un *motif moyen*, formant un prototype représentatif. Le fonctionnement de cet algorithme est détaillé en annexe D.

Nous appliquons la FSA aux motifs extraits pour différents niveaux d'état de santé de la batterie, il en résulte alors un prototype pour chacun de ces états de santé. Des

classes d'état de santé sont choisies comme étant trois périodes distinctes de la vie de la batterie Bat_1 . La figure 3.7 détaille la sélection de trois classes d'état de santé de la batterie Bat_1 . Le classe 1 et la classe 3 représentent ici respectivement le début et la fin de vie de la batterie Bat_1 . De plus, chacune de ces trois classes comporte le même nombre de motifs extraits.

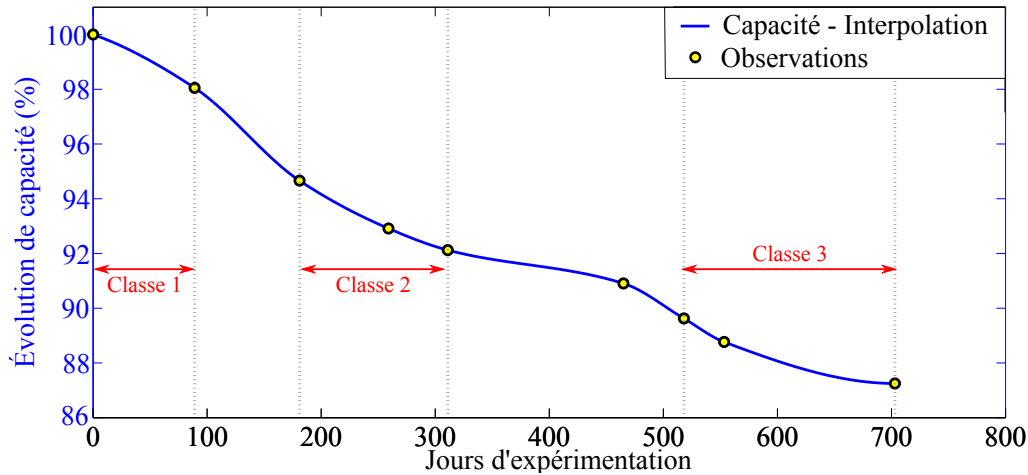


FIGURE 3.7 – Illustration de la création de 3 classes distinctes d'état de santé de la batterie Bat_1

Les classes disjointes de motifs extraits des Usages de la batterie Bat_1 forment une base de comparaison permettant l'identification des dynamiques de la batterie propres à son état de santé. La représentation des motifs de chaque classe, définis par le prototype moyen correspondant, est décrite par la figure 3.8.

Nous pouvons analyser les résultats de cette figure 3.8, comme une illustration de modification comportementale de la batterie pour une même vitesse mesurée. En effet, les motifs moyens de vitesses extraits sont similaires sur l'ensemble des trois classes constituées, ce qui n'est pas le cas des motifs courant et tension associés à ces accélérations de 10 à 60 km.h^{-1} . Ils présentent en effet des différences visibles de dynamique. On peut par exemple identifier le fait que le motif moyen de courant de la classe 2 possède une dynamique différente de celles des deux autres classes, du fait de la présence d'un pallier clairement visible. Cette variété des motifs est reproduite, à un degré moindre, dans les puissances mesurées.

De telles différences de comportement des signaux provenant de la batterie peuvent être interprétées comme une modification des réactions internes de la batterie en fonction de son état de santé. Les approches électrochimiques ne permettent cependant pas d'expliquer entièrement ce niveau de modification comportemental au sein des signaux. En effet, il est communément reporté dans la littérature une diminution de la puissance disponible associée aux phénomènes de vieillissement d'une batterie, mais il n'a pas été trouvé d'étude portant sur l'exploration des signaux issus des batteries avec ce niveau de détail.

La figure 3.8 présente donc des motifs de vitesse, courant, tension et de puissance. Ces derniers motifs s'avèrent différents selon les classes observées, pour une même vitesse mesurée. La suite de l'étude se place dans l'optique de l'investigation plus approfondie des motifs provenant de la batterie (courant, tension et puissance), afin de

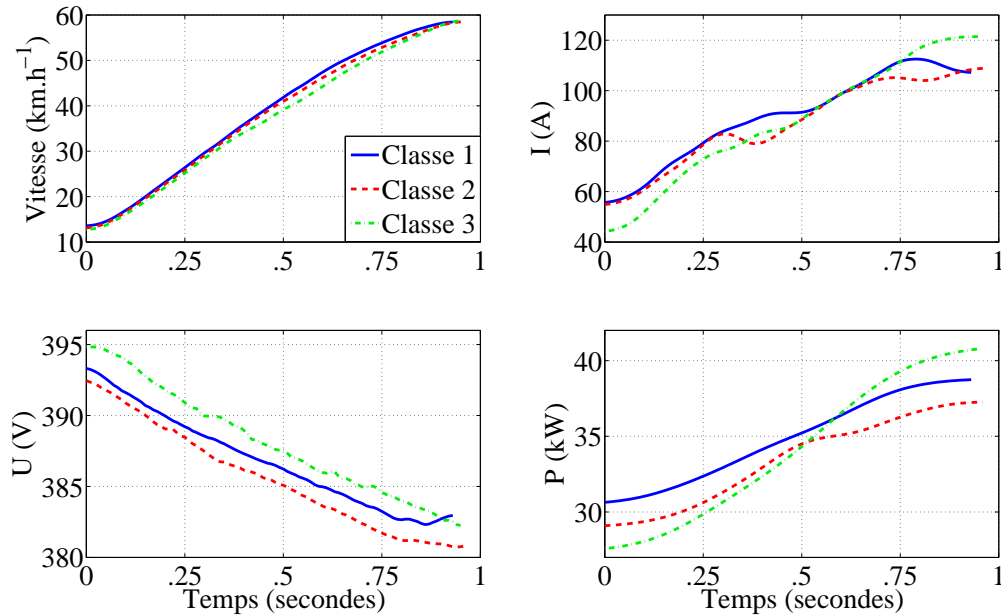


FIGURE 3.8 – Illustration des motifs moyens extraits pour chacune des trois classes d'état de santé - Pour les variables de vitesse, courant, tension et puissance

caractériser l'état de santé d'une batterie pour ces motifs.

Notons qu'il est intuitif que plus les phases d'accélération maximales sont longues, plus elles peuvent traduire le comportement de la batterie. Cela revient alors à dire qu'une accélération maximale de 20 à 40 km.h⁻¹ contient moins d'informations qu'une accélération permettant de passer de 10 à 60 km.h⁻¹. En effet, la batterie atteint plus rapidement ses limites de fonctionnement et de performance pour de fortes accélérations, générant une sollicitation plus importante que les courtes accélérations.

Dans le cadre d'une utilisation réelle d'un véhicule électrique, il est plus rare de détecter une longue et forte accélération qu'une courte accélération, qui peut se produire lors d'un feu rouge ou d'un ralentissement. Le choix des seuils de vitesses (S_v) est important, ce qui fera l'objet d'une analyse spécifique par la suite.

Cette section a donc révélé des modifications de comportement de la batterie, en fonction de son état de santé, pour un même profil de vitesse référence. Le but de cette approche est donc désormais de mettre en place une méthodologie utilisant ceci pour déduire l'état de santé d'une batterie, seulement à partir de motifs extraits lors d'une accélération (ou freinage) spécifique.

3.3 Approche de classification de l'état de santé d'une batterie

À partir des motifs de tension et de courant extraits en réponse à des critères de vitesse (S_v) et de durée (S_t), cette section explicite une méthodologie de diagnostic du niveau de santé d'une batterie en usage réel sur un véhicule électrique. La mise en place d'un tel algorithme permettant la classification de l'état de santé d'une batterie,

requiert dans un premier temps la définition d'une méthode de comparaison entre ces motifs. Cette quantification des similarités entre les motifs est fondamentale est au cœur de l'application des algorithmes de classification.

3.3.1 Choix d'une méthodologie de comparaison des signaux adaptée

L'objectif de cette section est le choix d'outils nécessaires à la comparaison de plusieurs ensembles de signaux. Lorsque l'on étudie un phénomène physique, uniquement à partir des signaux, on peut remarquer des courbes présentant de fortes similarités. Le problème consiste à quantifier ces similarités entre les signaux observés.

3.3.1.1 Distance euclidienne

Considérons ici la distance euclidienne permettant de calculer une distance entre deux séquences $X_1 := (x_1^1, \dots, x_N^1)$ et $X_2 := (x_1^2, \dots, x_N^2)$. Ainsi, la distance euclidienne $d_{ED}(X_1, X_2)$ entre ces deux séquences est calculée par l'équation 3.1.

$$d_{ED}(X_1, X_2) = \sqrt{\sum_{i=1}^n (x_i^1 - x_i^2)^2} \quad (3.1)$$

Cette métrique cumule les écarts point à point entre les deux signaux étudiés, comme l'illustre la figure 3.9.

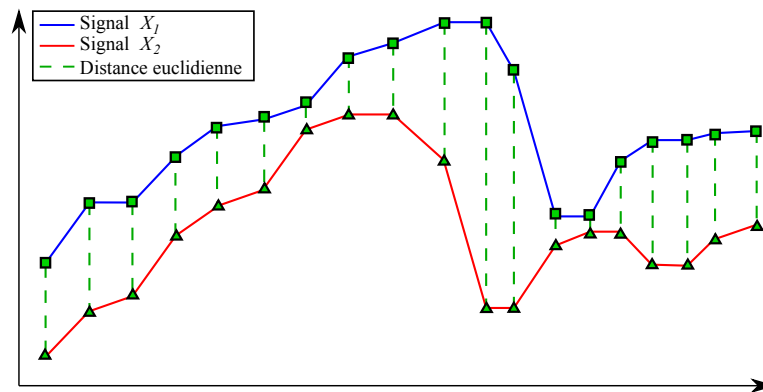


FIGURE 3.9 – Illustration du calcul de la distance euclidienne d_{ED} entre deux signaux X_1 et X_2

Cette propriété devient peu adaptée dans notre cadre de la mesure de similarité entre signaux. En effet, il se peut que deux signaux possèdent une forme fortement similaire mais légèrement décalée dans le temps. Dans ce cas, la distance euclidienne entre ces deux séquences est élevée et n'est pas représentative de cette similarité. D'autre part, la distance euclidienne ne permet pas, par sa définition, de mesurer une similarité entre deux séquences de longueurs différentes. Cette autre propriété est donc rédhibitoire quant à l'usage de cette métrique dans le contexte de l'analyse de motifs issus de la batterie, ceux-ci étant de longueurs très variées.

3.3.1.2 Recalage des signaux

Considérant les limitations évoquées à propos de la distance euclidienne, il est nécessaire de trouver une alternative permettant de comparer des signaux tels que ceux issus d'une batterie. En particulier, les signaux réels sont le plus souvent observés sur des intervalles de temps différents. Ainsi, nous étudions ici une approche de recalage des signaux. Cette méthodologie consiste alors à *synchroniser* ou *recaler* des séquences deux-à-deux, permettant d'éviter de potentiels décalages temporels des dynamiques, tout en obtenant de nouveaux signaux recalés de même longueur.

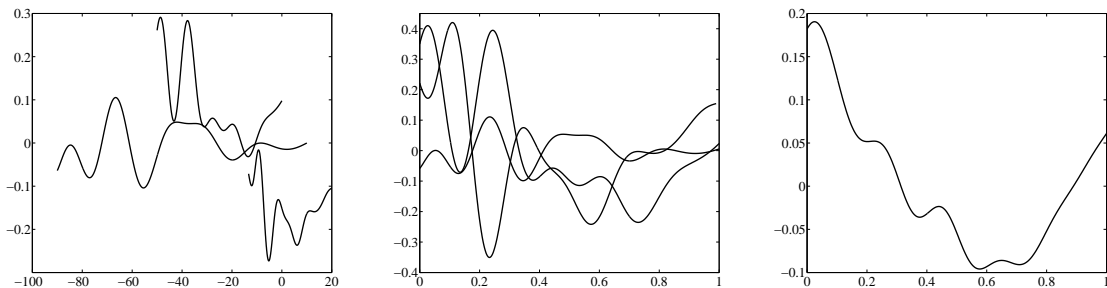
La première méthode, appelée *recalage linéaire*, consiste à représenter les séquences étudiées sur un support commun. Considérons N signaux X_1, \dots, X_N observés sur N intervalles différents $[a_1, b_1], \dots, [a_N, b_N]$. L'objectif étant d'obtenir un signal moyen \widehat{X} , la méthodologie revient à représenter toutes ces séquences sur un même intervalle, typiquement $[0, 1]$. Les signaux X_i sont dilatés pour devenir des signaux \tilde{X}_i , définis par l'équation 3.2.

$$\tilde{X}_i(t) = X_i(a_i + t(b_i - a_i)), \text{ pour } t \in [0, 1], \text{ et } i = 1, \dots, N \quad (3.2)$$

Cette définition 3.2 des signaux recalés \tilde{X}_i induit simplement une forme caractéristique du signal moyen $\widehat{X}(t)$ de cet ensemble :

$$\widehat{X}(t) = \frac{1}{N} \sum_{i=1}^N \tilde{X}_i(t), \text{ pour } t \in [0, 1] \quad (3.3)$$

Cette approche de recalage linéaire est illustrée en figure 3.10.



(a) 3 courbes ayant des caractéristiques similaires (b) Recalage linéaire des courbes normalisées (c) Moyenne des courbes après recalage linéaire

FIGURE 3.10 – Estimation de la forme moyenne d'un ensemble de courbes par recalage linéaire

Cette méthode s'avère particulièrement adaptée dans le cas d'une comparaison de signaux identiques mais dilatés. Or, si les signaux sont tronqués, ou possédant un ajout d'information, le recalage linéaire ne permet pas d'obtenir de résultat satisfaisant. Ces limitations du recalage linéaire ont engendré le développement d'autres méthodologies de recalage.

Formellement, le problème de la synchronisation de deux séquences X_1 et X_2 consiste à déterminer une transformation u , suffisamment régulière afin que les caractéristiques du signal original X_2 soient conservées après déformation, telle que :

$$X_1 \approx X_2 \circ u \quad (3.4)$$

Si l'on souhaite recaler un ensemble de S signaux X_1, \dots, X_S , la méthodologie revient à trouver les S transformations u_1, \dots, u_S suffisamment régulières, telles que les signaux déformés $X_1 \circ u_1, \dots, X_S \circ u_S$ soient alignés de manière optimale au sens d'une certaine mesure de similarité entre signaux.

Les approches de recalage de signaux peuvent être décomposées en deux classes qui consistent soit en un alignement global soit en un alignement basé sur des points caractéristiques de la forme d'un signal appelés *landmarks* [142]. La première approche vise à trouver des transformations alignant l'ensemble des points des deux signaux, au sens d'une certaine métrique entre deux courbes.

Au contraire, la seconde approche correspond à un alignement à partir de *landmarks*. Ces derniers peuvent par exemple être, selon le type de signal étudié, des points d'inflexion ou bien des singularités. Cette seconde approche vise à mettre en correspondance, les paires de landmarks issus de deux signaux devant être recalés. Ainsi, les landmarks des signaux se trouvent à la même position après déformation temporelle (figure 3.11).

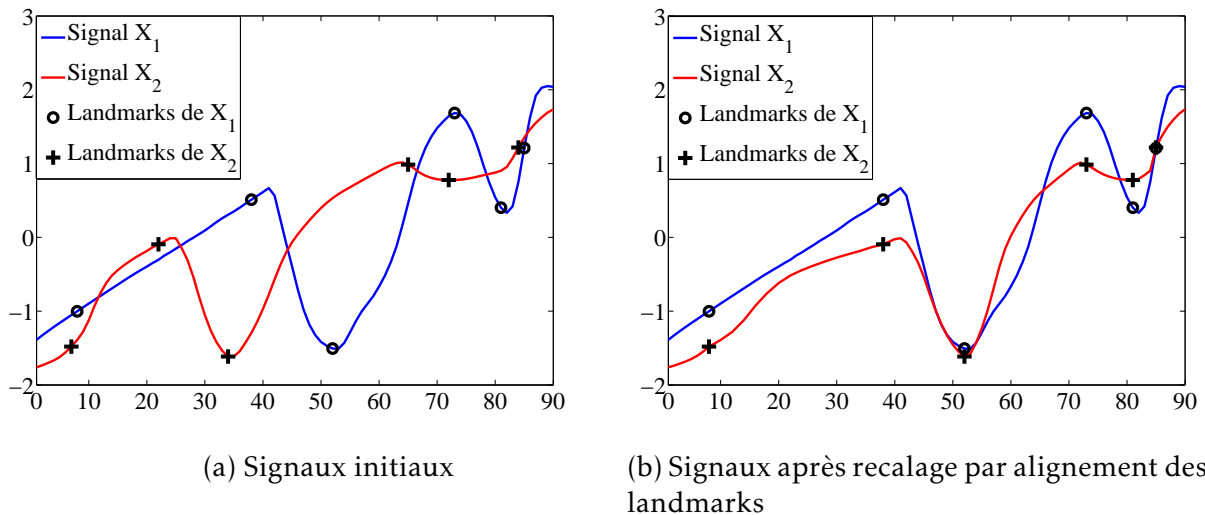


FIGURE 3.11 – Illustration de signaux X_1 et X_2 recalés après détection et recalage de landmarks

Notons que l'alignement global est communément utilisé lorsqu'aucun landmark n'est détectable ou bien lorsque ceux-ci n'apportent pas suffisamment d'information. Davantage de détails sur les fondements théoriques et les développements sur ce type d'approches peuvent être trouvés dans la littérature [143].

Une fois ce recalage effectué, la procédure d'obtention d'un signal moyen consiste à l'application de la formule 3.3. La figure 3.12 les résultats de cette méthode, sur les mêmes séquences utilisées comme exemple lors de la figure 3.10.

Il apparaît clairement la meilleure représentation de la dynamique des séquences obtenue par des alignements basés sur des landmarks, par rapport aux alignement linéaires présentés en figure 3.10. Cette méthode présente en effet l'intérêt de préserver les caractéristiques propres aux signaux et donc de pouvoir interpréter un signal moyen ou bien de servir de référence aux algorithmes de classification.

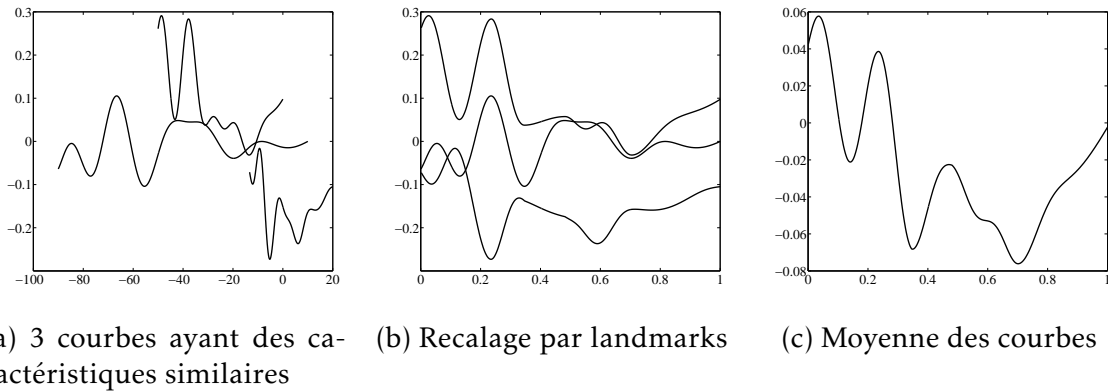


FIGURE 3.12 – Estimation de la forme moyenne d'un ensemble de courbes alignées par landmarks

3.3.1.3 Dynamic Time Warping

Afin de comparer les motifs extraits et de quantifier leurs similarités, nous considérons désormais une métrique adaptée au problème spécifique de la comparaison de motifs extraits des signaux de batterie. Ainsi, au-delà des mesures usuelles, une des méthodes référence pour quantifier les similarités de forme est actuellement la *Dynamic Time Warping (DTW)*.

Le concept de la DTW fut initialement introduit dans les années 70 pour l'obtention d'une fonction d'alignement entre deux signaux, associée à un coût de transformation (ou d'alignement). Le but premier est de comparer des séquences en respectant leurs formes et d'identifier l'alignement optimal en les étirant. Depuis son introduction, la DTW a été communément utilisée dans de nombreuses problématiques : le traitement de la parole, la reconnaissance de signaux, l'analyse de données ou l'imagerie.

Cette méthode est basée sur la distance de Levenshtein et trouve le chemin optimal entre deux séquences, considérant leurs distorsions temporelles. Ce chemin optimal produit une fonction d'alignement, ainsi qu'une mesure de similarité de formes. Formellement, nous avons deux séquences $X_1 := (x_1^1, \dots, x_N^1)$ de longueur $N \in \mathbb{N}$, et $X_2 := (x_1^2, \dots, x_M^2)$ de longueur $M \in \mathbb{N}$. Par la suite nous considérons l'espace des individus noté par \mathcal{F} .

Pour comparer deux individus $x^1, x^2 \in \mathcal{F}$, il est nécessaire de définir une mesure de coût local, par la fonction c :

$$c : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}_{\geq 0} \quad (3.5)$$

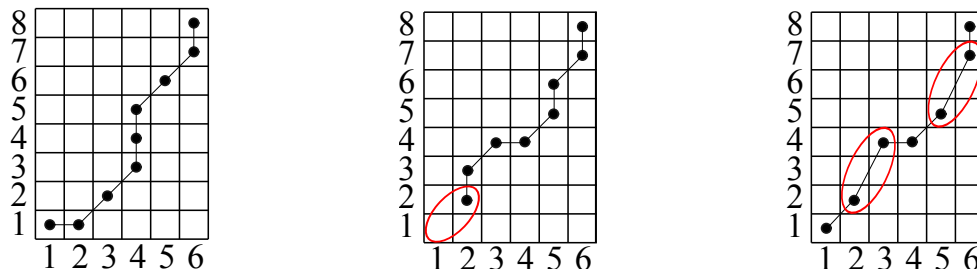
Typiquement, le coût $c(x^1, x^2)$ est bas si x^1 et x^2 sont similaires, et dans le cas contraire $c(x^1, x^2)$ est élevé. En évaluant cette mesure de coût local $c(x^1, x^2)$ pour chaque paire d'éléments des séquences X_1 et X_2 , on obtient une matrice de coûts $C \in \mathbb{R}^{N \times M}$, définie par $C(i, j) = c(x_i^1, x_j^2)$. Le but est trouver l'alignement entre X_1 et X_2 minimisant le coût global.

Formellement un chemin d'alignement (*warping path*) est une séquence $p = (p_1, \dots, p_L)$,

avec $p_l = (n_l, m_l) \in [1 : N] \times [1 : M]$, $\forall l \in [1 : L]$, satisfaisant les conditions :

$$\begin{cases} p_1 = (1, 1) \text{ et } p_L = (N, M) \\ n_1 \leq \dots \leq n_L \text{ et } m_1 \leq \dots \leq m_L \\ p_{l+1} - p_l \in \{(1, 0), (0, 1), (1, 1)\}, \forall l \in [1 : L - 1] \end{cases} \quad (3.6)$$

Un *warping path*¹ $p = (p_1, \dots, p_L)$ définit un alignement entre deux séquences X_1 et X_2 en assignant un élément $x_{n_l}^1$ de X_1 à un élément $x_{m_l}^2$ de X_2 . Les conditions d'alignement, décrites par l'équation 3.6, impliquent que les premiers éléments de X_1 et de X_2 , de même que leurs derniers éléments, soient associés. Ces conditions rendent donc non valables les chemins ne respectant pas les équations 3.6. Les figures 3.13 illustrent chacun des cas possibles, admissibles ou non.



(a) Chemin d'alignement admissible (b) Exemple d'un chemin non admissible, par non respect des bornes (c) Exemple de chemin non admissible, par non respect de la longueur des pas

FIGURE 3.13 – Illustration de chemins reliant deux séquences X_1 et X_2 , de longueurs respectives $N = 6$ et $M = 8$, les conditions non admissibles sont représentées en rouge

Les figures 3.14 présentent la matrice des coûts formée à partir de deux séquences (signal X_1 et signal X_2), ainsi que la représentation visuelle de l'alignement induit pour le chemin choisi.

Le coût global $c_p(X_1, X_2)$ d'un chemin d'alignement p entre X_1 et X_2 est défini à partir de la mesure locale c , selon l'équation 3.7.

$$c_p(X_1, X_2) = \sum_{l=1}^L c(x_{n_l}^1, x_{m_l}^2) \quad (3.7)$$

Le chemin d'alignement optimal entre X_1 et X_2 est alors le chemin d'alignement p^* , minimisant le coût global parmi tous les chemins admissibles. La distance $d_{DTW}(X_1, X_2)$, associée à la DTW, entre X_1 et X_2 est donc définie comme étant le coût global de ce chemin d'alignement optimal p^* .

$$d_{DTW}(X_1, X_2) = c_{p^*}(X_1, X_2) = \min\{c_p(X_1, X_2) \mid \forall p\} \quad (3.8)$$

Il est néanmoins important de noter le fait que cette distance d_{DTW} n'est pas à proprement parler une métrique, mais une pseudo-métrique, car elle ne satisfait pas l'inégalité triangulaire (équation 3.9).

$$\exists X_1, X_2, X_3, d_{DTW}(X_1, X_3) \not\leq d_{DTW}(X_1, X_2) + d_{DTW}(X_2, X_3) \quad (3.9)$$

La distance associée à la DTW permet donc la comparaison et la quantification des formes de signaux.

1. Chemin d'alignement.

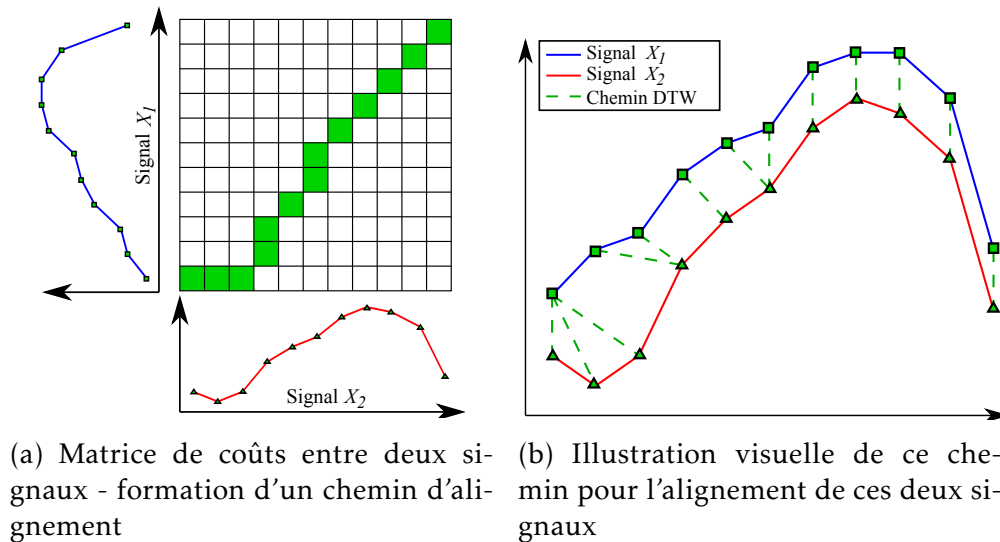


FIGURE 3.14 – Illustration d'un chemin d'alignement entre deux signaux X_1 et X_2 , (a) sur la matrice de coûts, (b) sur l'alignement induit sur les signaux

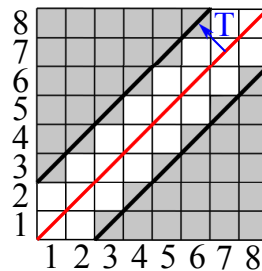


FIGURE 3.15 – Illustration de l'application d'une bande restrictive de Sakoe & Chiba avec $T = 1$, pour la construction d'un chemin d'alignement

3.3.1.4 Contraintes de Sakoe & Chiba

Une variante de la distance DTW, basée sur la modification de l'équation 3.7, est définie avec des contraintes de Sakoe & Chiba [144], qui permettent d'ignorer certains chemins avec un paramètre T limitant de la largeur de la bande des contraintes :

$$d_{DTW_{SC}}(X_1, X_2) = \min \left\{ \sum_{l=1}^L \delta_{n_l, m_l} c(x_{n_l}^1, x_{m_l}^2) \mid p \right\}, \quad (3.10)$$

avec

$$\delta_{n_l, m_l} \begin{cases} 1 & , \text{ si } |n_l - m_l| < T \\ \infty & , \text{ si } |n_l - m_l| \geq T \end{cases} \quad (3.11)$$

L'application de cette contrainte au sein de la distance $d_{DTW_{SC}}$ a pour principal effet de permettre la diminution de la complexité de l'algorithme DTW. Le figure 3.15 présente un exemple d'utilisation de ces bornes de Sakoe & Chiba.

3.3.1.5 Global Alignment

Le concept de la DTW est basé sur le chemin minimisant le coût total entre deux séquences (équation 3.8). Une alternative proposée par Cuturi [145] est nommée *Global*

Alignment (GA). Le GA consiste à prendre en compte désormais non plus le minimum mais le *soft-minimum* des coûts $c_p(X_1, X_2)$ entre deux séquences X_1 et X_2 . Ce procédé permet de considérer l'alignement optimal ainsi que tous les alignements proches de celui-ci. Le *soft-minimum* est défini par l'équation :

$$\text{soft-min}(c(X_1, X_2)) = -\log \sum_p \exp(-c_p(X_1, X_2)) \quad (3.12)$$

La distance GA, notée $d_{GA}(X_1, X_2)$, est donc obtenue par le calcul du soft-minimum des chemins entre X_1 et X_2 :

$$d_{GA}(X_1, X_2) = \text{soft-min}(c(X_1, X_2)) \quad (3.13)$$

De manière pratique, le GA permet de considérer l'ensemble des chemins entre X_1 et X_2 avec un poids associé relatif à la pertinence de ces chemins. En d'autres termes, tous les chemins reliant X_1 à X_2 sont pris en compte dans le calcul de la distance d_{GA} .

3.3.1.6 Application à l'étude des motifs

Dans le cadre de l'analyse de motifs extraits à partir des signaux issus d'une batterie en usage, il n'est pas envisageable de considérer la distance euclidienne comme une issue possible. En effet, les motifs extraits sont de longueurs diverses, ce qui n'est pas compatible avec l'emploi de cette métrique, sans prétraitement pouvant introduire un biais. La distance euclidienne ne permettant pas de prendre en compte les décalages temporels ou distorsions, qui sont fréquents dans ce type de signaux réels.

Les méthodes de recalage décrites permettent d'éviter ce problème de motifs de longueurs variées. Cependant, de tels recalages dans notre application ne permettent pas de conserver l'entière information contenue préalablement dans les motifs extraits. En effet, par définition, le recalage réduit les dynamiques propres à chaque signal, afin d'obtenir des signaux les plus similaires possibles. De ce fait, l'application de recalage directement sur les motifs ne peut être considérée pour la suite de l'étude.

La distance DTW est, elle, bâtie afin de permettre le calcul de dissimilarité entre motifs. Cette méthodologie mesure une distance entre motifs, deux-à-deux, quantifiant les dissimilarités de dynamique, quelque soit la longueur de ces motifs. Il est de même pour la distance GA. Dans la suite de l'étude ces distances DTW et GA seront donc privilégiées et leurs résultats comparés.

3.3.2 Méthodes de classification

À partir des mesures de similarité entre motifs, définies précédemment, il est désormais possible de quantifier les différences de dynamiques entre motifs. L'objectif est à présent d'utiliser ces informations pour estimer la classe d'état de santé d'une batterie. Dans un premier temps, le but est ici de valider l'hypothèse faite en section 3.2.3, d'une présence d'information relative à l'état de santé d'une batterie au sein de ces signaux mesurés lors d'accélérations spécifiques. Pour ce faire, nous nous plaçons dans le cadre d'une classification supervisée ayant pour objectif d'estimer une classe d'état de santé d'une batterie, uniquement à partir des motifs extraits selon la méthodologie décrite en section 3.2.2.

Autrement dit, cette étude a pour but de parvenir à estimer la classe d'état de santé d'une batterie uniquement en étudiant un motif extrait de ces usages, en le comparant à une base d'apprentissage préalablement établie.

Formellement, nous avons un ensemble d'apprentissage composé de n motifs x_i accompagnés de leurs classes d'état de santé respectives t_i : $B^{train} = \{(x_1, t_1), \dots, (x_n, t_n)\}$. Ainsi, la procédure de classification supervisée assigne une classe t à un nouveau motif x étudié. Pour ce faire, nous considérons des méthodes de classification usuelles dont le fonctionnement est décrit dans cette section.

3.3.2.1 k-Nearest Neighbors

L'approche des k plus proches voisins ou *k-Nearest Neighbors* (*kNN*), a largement été étudiée et a démontré être une technique non paramétrique puissante dans les problèmes de classification. Cet algorithme est simple à mettre en œuvre et donc communément utilisé comme référence.

Pour un nouveau motif, la méthode *kNN* assigne la classe votée majoritairement parmi des k plus proches voisins de cet objet, au sein de l'ensemble d'apprentissage B^{train} . Formellement, pour un ensemble d'apprentissage étiqueté B^{train} , et un nouvel objet x , la méthode *kNN* classe cet objet x en considérant k éléments au sein de l'ensemble d'apprentissage $\{x_1, \dots, x_k\} \in B^{train}$. Ces k éléments sont choisis comme étant les k plus proches de l'objet x en fonction d'une certaine métrique.

Ainsi, le motif x est classé à partir d'une statistique dépendant des classes de ces k plus proches voisins $\{t_1, \dots, t_k\}$. Par exemple, cette statistique peut correspondre à une détection de la classe présente majoritairement. Notons que le paramètre k est un entier positif, qui influe de manière importante sur la précision de l'algorithme. Ce paramètre k est défini dans cette étude à partir d'une procédure de validation croisée [146].

3.3.2.2 k-Means

L'algorithme *k-Means* est l'une des plus populaire et des plus employées parmi les méthodes de formation de *clusters*². La méthodologie de cet algorithme se base sur un ensemble initial de n observations puis utilise un processus itératif afin d'optimiser la cohérence des clusters formés.

L'objectif est donc de regrouper n observations en k clusters disjoints $\{C_1, \dots, C_k\}$, tels que la somme des variances intra-groupes soit minimisée. Les positions de chacun des centres de clusters $(c_1^{(0)}, \dots, c_k^{(0)})$ sont initialement choisies aléatoirement parmi les observations. Ainsi, à chaque étape les observations sont associées au cluster le plus proche. La distance entre un élément x_i et un cluster C_j étant définie comme étant la distance entre x_i et le centre de ce cluster c_k^p à l'étape p . À la fin d'une étape p , les barycentres de clusters $(c_1^{(p)}, \dots, c_k^{(p)})$ sont modifiés comme étant le centre des observations associées à chacun des clusters à cette étape. Cette procédure est ainsi itérée jusqu'à l'obtention d'un critère d'arrêt, usuellement défini par un critère de stabilité des centres de clusters.

2. Un ensemble d'objets sont regroupés dans un groupe appelé *cluster*.

Notons que la définition des centres des clusters, ici proposée par le biais de barycentres, peut également être faite par l'utilisation d'un point réel du cluster. Cette variante est appelée la méthode *k-Medoids*.

Cette procédure globale, à l'issue des itérations, permet d'associer chacune des données de l'ensemble d'apprentissage B^{train} avec un unique cluster parmi les k constitués. Considérant une classification supervisée, un nouvel objet x est associé au cluster le plus proche. De ce fait, le label attribué à cet objet x est le résultat d'une statistique des labels de ce cluster.

3.3.2.3 Classification Ascendante Hiérarchique

La Classification Ascendante Hiérarchique (CAH) est, comme la méthode k-Means (section 3.3.2.2), une méthodologie permettant de partitionner un ensemble d'observations en clusters disjoints, de sorte que les éléments appartenant au même cluster² soient proches d'après une métrique choisie. Ce notion de distance entre un élément et un cluster est ici définie comme étant la distance entre cet élément et le barycentre dudit cluster. Cette classification se fait alors suivant une procédure de regroupement itératif.

Initialement, l'ensemble des n observations sont répartis dans n clusters distincts, chacun contenant un unique élément, puis les deux plus proches sont fusionnés itérativement, jusqu'à la formation d'un unique cluster regroupant toutes les observations. Cette étape de regroupement ascendant forme un arbre hiérarchique qui est alors élagué afin de conserver le nombre défini k de clusters. Ainsi, les éléments étant sous chaque branche sont affectés à un cluster, ce qui forme le partitionnement des observations en k clusters $\{C_1, \dots, C_k\}$. La définition du critère d'élagage est délicate et de nombreuses approches ont été proposées. Dans cette étude nous utilisons la méthode du *gap statistics* [147] qui sera détaillée dans le prochain chapitre (section 4.2.1.2).

Le procédé décrit est nommé de type *bottom-up*³, notons qu'il est également possible de procéder à la réciproque appelée *top-down*. La figure 3.16 illustre ce procédé *bottom-up*⁴.

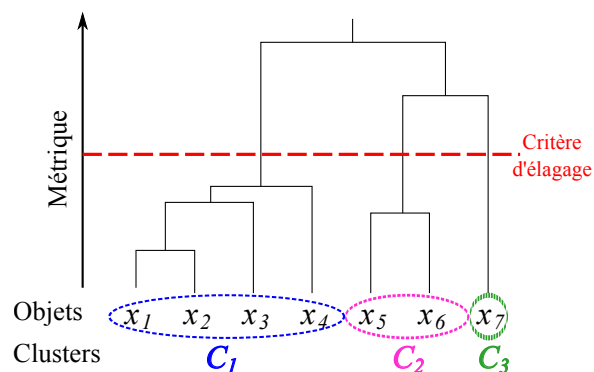


FIGURE 3.16 – Illustration d'une CAH entre 7 objets x_i , avec un critère d'élagage induisant la formation de 3 clusters disjoints C_j

3. Regroupe les les individus itérativement.

4. Scinde l'ensemble des données itérativement.

Utilisant ce regroupement, de la même manière que le k-Means, un nouvel objet est associé au label résultant d'une statistique des labels du cluster le plus proche.

3.3.3 Résultats des méthodes de classification

L'ensemble des méthodologies de classifications sont appliquées aux motifs extraits parmi les signaux mesurés. Pour ce faire, nous employons la distance d_{DTW} (section 3.3.1.3), afin de quantifier les dissimilarités entre motifs. La distance d_{GA} (section 3.3.1.5) n'est pas ici considérée, n'ayant pas démontré de résultats satisfaisant lors de l'application des méthodes de classification. En effet, les résultats obtenus avec cette distance induisent une perte de précision par rapport à la DTW, malgré les avantages théoriques décrits en section 3.3.1.5.

Nous utilisons différentes classes de motifs discontinues, dont un exemple de construction avec la batterie Bat_1 est illustrée en figure 3.7. Ces motifs correspondent à des signaux de courant, tension et puissance extraits lors d'accélération spécifiques. Ces motifs sont aléatoirement séparés en une partie d'apprentissage $B^{train} = \{(x_1, t_1), \dots, (x_n, t_n)\}$ étiquetés à une classe d'état de santé, puis une partie de test $B^{test} = \{(x_1, t_1), \dots, (x_m, t_m)\}$ pour laquelle nous cherchons à estimer les classes réelles. Les niveaux d'état de santé estimés $\hat{t} = \{\hat{t}_1, \dots, \hat{t}_m\}$ sont ensuite comparés aux valeurs réelles $t = \{t_1, \dots, t_m\}$. Pour chaque méthodologie, l'erreur d'estimation e est alors calculée selon l'équation 3.14.

$$e = \frac{\text{Card}(x_i \in B^{test} | \hat{t}_i = t_i)}{B^{test}} \quad (3.14)$$

Notons que les paramètres de chacune des méthodes sont préalablement déterminés par validation croisée opérée sur les données d'apprentissage B^{train} , qui correspondent à 75 % des données disponibles. Cette opération est effectuée par bootstrap 50 fois, ce qui permet d'obtenir des erreurs d'estimation e représentatives.

3.3.3.1 Approche sur la batterie Bat_1

La première expérimentation porte sur les motifs de la batterie Bat_1 , avec les motifs d'accélération de 10 à 60 km.h⁻¹ de la batterie Bat_1 , séparés en 3 classes d'après la figure 3.7 reprise ici en figure 3.17. La table 3.1 recense les erreurs d'estimation e obtenues avec ces motifs pour chacune des méthodes de classification présentées.

	kNN	k-Means	CAH
I	40.2	39.0	41.2
U	41.3	41.2	36.6
P	39.0	46.0	42.6

TABLE 3.1 – Erreurs d'estimation e (en %) obtenues à partir de motifs issus d'accélération de 10 à 60 km.h⁻¹ de la batterie Bat_1 , pour 3 classes avec la DTW

Cette présentation des erreurs d'estimation de motifs appartenant à 3 classes démontre la capacité d'estimation de la méthodologie de classification détaillée ici. En effet, le choix étant restreint à trois classes, une estimation aléatoire fournit une erreur d'estimation de $e = 66 \%$, ce qui est très largement supérieur à toutes les erreurs

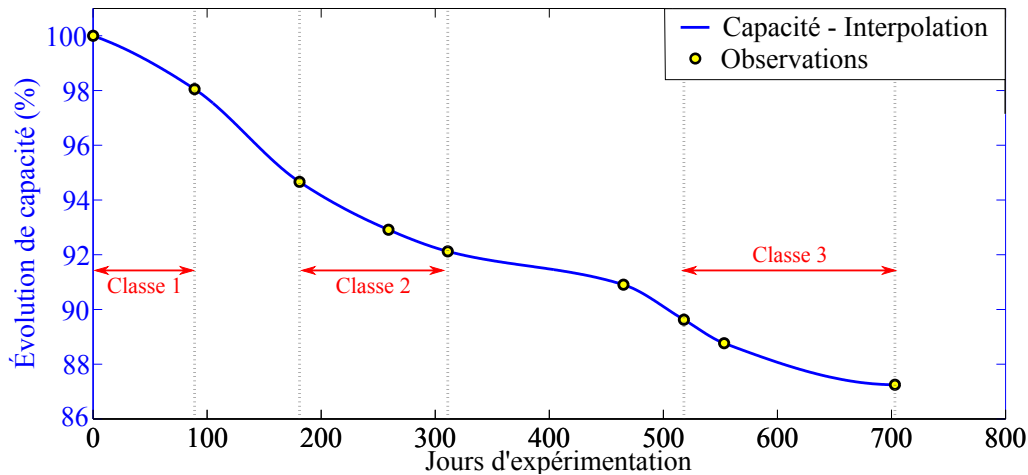


FIGURE 3.17 – Illustration de la création de 3 classes distinctes d'état de santé de la batterie Bat_1

obtenues par notre méthodologie. Ceci confirme et démontre donc la présence d'informations dans des signaux sur l'état de santé d'une batterie.

Le meilleur résultat obtenu avec cette expérimentation est donc un taux d'erreur minimal de $e = 36.6 \%$, ce qui représente un niveau de confiance très intéressant car la base d'apprentissage ne recouvre qu'une faible perte de SOH (figure 3.17). De ce fait, les modifications de comportement de la batterie sont faibles entre chacune de ces trois classes, ce qui explique une difficile identification de l'état de santé de la batterie. Ce niveau de précision permet néanmoins de valider l'hypothèse de présence d'information relative à l'état de santé de la batterie au sein des signaux mesurés (I, U, et P).

D'autre part, il apparaît une régularité des résultats obtenus par chacune des méthodes de classification, et pour tous les types de motifs. L'étude de la batterie Bat_1 ne permet donc pas de conclure sur la pertinence d'un algorithme en particulier ou bien d'un signal parmi I, U, P contenant plus d'informations que les autres.

3.3.3.2 Application aux accélérations de 20 à 40 km.h^{-1}

Les expérimentations menées sur cette batterie Bat_1 étant particulières, car faites sur circuit (section 2.2.1), nous observons les résultats de cette méthodologie portée sur la batterie Bat_3 pour des accélérations de 20 à 40 km.h^{-1} .

La table 3.2 présente les résultats des méthodes de classifications appliquées aux données de la batterie Bat_3 , séparées en 4 classes d'état de santé discontinues, l'erreur associée à une estimation aléatoire est donc ici de $e = 75 \%$.

Ces résultats démontrent la possibilité d'estimation d'état de santé uniquement à partir d'un signal (I, U ou P) mesuré directement sur la batterie lors d'accélérations de 20 à 40 km.h^{-1} pendant des usages réels. Les performances détaillées dans cette table 3.2 propose une erreur d'estimation minimale de $e = 60 \%$ obtenue par la méthode k-Means avec des signaux de courant. L'analyse de ces résultats, comme lors de la section 3.3.3.1 ne permet néanmoins pas de conclure à la précision particulière d'une méthode et/ou d'un signal par rapport aux autres.

		kNN	k-Means	CAH
20 – 40 km.h ⁻¹	I	61.1	60.0	65.1
	U	60.7	62.5	66.8
	P	60.8	61.0	60.1

TABLE 3.2 – Erreurs d'estimation e (en %) obtenues à partir de motifs issus des accélérations 20 à 40 km.h⁻¹ sur la batterie Bat_3 , pour 4 classes avec la DTW

3.3.3.3 Application aux accélérations de 10 à 60 km.h⁻¹

Intuitivement, il apparaît que plus l'accélération considérée est longue, plus la batterie atteint ses performances limites, altérées par les phénomènes de vieillissement. De ce fait, des signaux (I, U, P) issus d'une accélération maximale prolongée par rapport aux accélérations de 20 à 40 km.h⁻¹, devraient en théorie atteindre des erreurs d'estimations e plus faibles.

Analysons les résultats de classification résultant des accélérations de 10 à 60 km.h⁻¹ sur la batterie Bat_3 . Cette étude proposant de faire un comparatif des précisions en fonction des critères de vitesses considérés, à partir des résultats obtenus en section 3.3.3.2. Ce comparatif des erreurs d'estimations calculées par l'équation 3.14 est détaillé dans la table 3.3.

		kNN	k-Means	CAH
20 à 40 km.h ⁻¹	I	61.1	60.0	65.1
	U	60.7	62.5	66.8
	P	60.8	61.0	60.1
10 à 60 km.h ⁻¹	I	58.9	64.2	66.1
	U	50.7	53.6	61.1
	P	60.6	65.5	63.1

TABLE 3.3 – Erreurs d'estimation e (en %) obtenues à partir de motifs issus selon deux critères d'accélération sur la batterie Bat_3 , pour 4 classes avec la DTW

Les résultats détaillés dans cette table 3.3 démontrent une nette amélioration globale des estimations lorsque la plage d'accélération considérée est plus longue. Ceci confirme la présence, au sein des signaux I, U et P mesurés, de comportements caractéristiques des limitations de performances de la batterie.

Il est également important de noter la très faible modification des résultats de la méthode CAH selon les deux critères d'accélération étudiés. Au contraire, la méthode kNN semble ici la plus performante des méthodologies de classification choisies, permettant d'obtenir une erreur d'estimation minimale de $e = 50.7$ % avec les signaux de tension issus d'accélérations de 10 à 60 km.h⁻¹.

Les signaux de tension présentent les meilleures résultats, ce qui suggère un plus grand niveau d'information contenu dans ces motifs. Nous pouvons cependant remarquer que les faibles performances obtenues avec les motifs de puissance, ce qui peut être du au nivelage des modifications de dynamique propre au courant et à la tension.

Cette approche induit une préférence pour l'utilisation de la méthode kNN avec des signaux de tension issus d'une longue plage d'accélération.

3.3.3.4 Influence des critères d'extraction

Ce type de méthodologie de classification supervisée étant uniquement basé sur les données d'apprentissage, la précision obtenue est fortement reliée à la représentativité de la base d'apprentissage. Il est implicite qu'en construisant un modèle de classification basé sur plus de motifs d'apprentissage, celui-ci permettrait d'atteindre des erreurs d'estimations plus faible. Or, dans le contexte des usages réels le nombre de motifs extraits dépend uniquement des profils de vitesse appliqués par le conducteur et des critères d'extraction définis.

Dans notre cas nous considérons des usages de véhicules électriques réels, pour lesquels aucune consigne n'a été donnée au conducteur. Ainsi, en considérant des données fixes, les seules variables induisant une modification du nombre de motifs extraits sont les critères d'extractions définis en section 3.2.2 (figure 3.5).

De ce fait, pour agrandir la base d'apprentissage il est uniquement possible de modifier les critères d'extraction des motifs. Ces critères étant composés de la durée du motif $S_t = (t_{min}, t_{max})$ et des seuils de vitesse $S_v = (v_{init}, v_{term})$. Il est envisageable d'élargir ces seuils pour permettre une extraction d'un plus grand nombre de motifs. Ce procédé est cependant fait au détriment du niveau de précision de l'algorithme. Les seuils sont le résultat d'un compromis entre la fréquence des extractions (et donc des estimations) et de la précision requise. La figure 3.18 présente l'évolution de la précision de l'algorithme kNN avec des signaux de courant, et du nombre de motifs extraits, en fonction de différentes plages de $S_t = (t_{min}, t_{max})$, pour $v_{init} = 10$ et $v_{term} = 60$ pour la batterie Bat_2 .

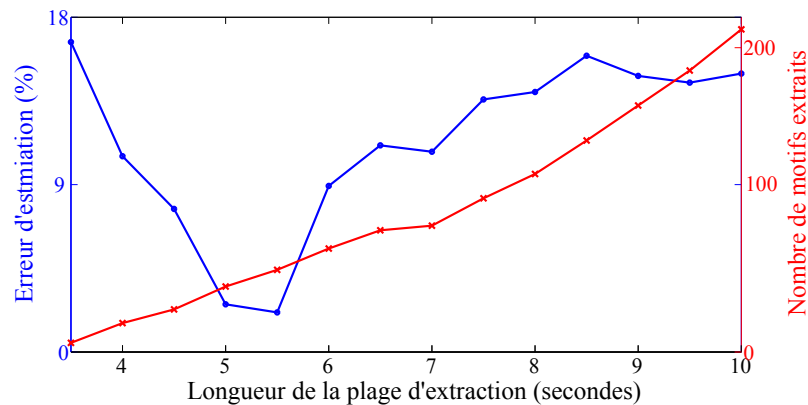


FIGURE 3.18 – Illustration du choix des critères d'extraction, démontrant le compromis entre le nombre de motifs et la précision de l'estimation

Afin d'illustrer l'évolution de l'erreur d'estimation e (équation 3.14) de la méthode kNN, en fonction de la fréquence des estimations, la figure 3.19 présente le rapport entre ces deux notions pour les mêmes conditions que celles détaillées pour la figure 3.18, démontrant le problème de généralisation de l'algorithme.

Nous pouvons donc voir que la définition des critères d'extraction est dépendant de la fréquence et de la précision que l'utilisateur souhaite obtenir. En effet, les esti-

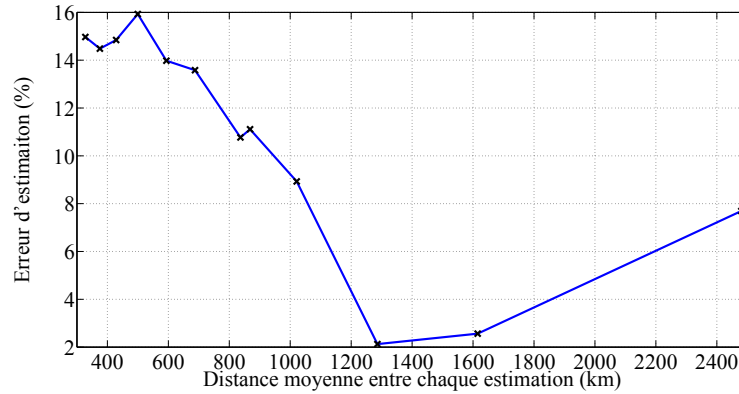


FIGURE 3.19 – Résultat d'une classification de motifs de la batterie Bat_3 , avec la méthode kNN

mations les plus précises dans ce cas sont obtenues environ tous les 1 300 km d'usage d'un véhicule électrique, selon l'application étudiée ici.

3.3.3.5 Extension à 8 classes

Dans le but d'obtenir les résultats des classification de manière plus détaillée, nous considérons désormais une décomposition des motifs de tension issus d'accélération de 10 à 60 km.h^{-1} de la batterie Bat_3 , en 8 classes discontinues. Les résultats de la classification obtenue par la méthode kNN sur ces données sont présentées en figure 3.20.

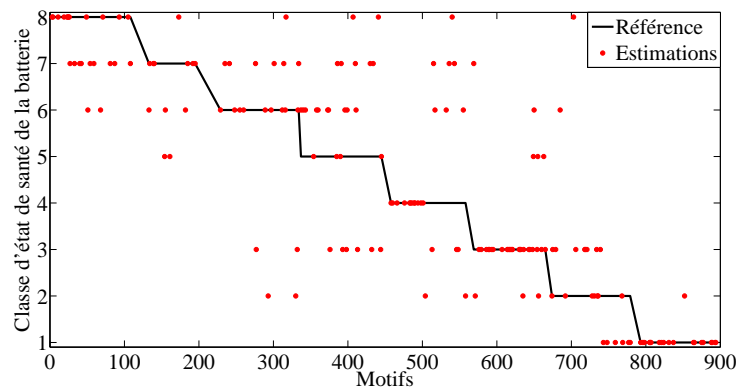


FIGURE 3.20 – Représentation de la matrice de confusion de la méthode kNN issue de la classification des signaux de tension provenant d'accélération de 10 à 60 km.h^{-1} avec la batterie Bat_3

La figure 3.20 présente la tendance des estimations qui suivent la dynamique de dégradation de l'état de santé de la batterie. En effet, même lorsqu'elle s'avère fautive la classe estimée est généralement proche de la classe réelle. Ces résultats confirment la bonne performance globale obtenue par la méthodologie basée sur les signaux provenant d'accélération. De tels résultats permettent une utilisation plus approfondie des

informations contenues dans ces signaux mesurés directement sur la batterie.

3.4 Vers une estimation continue de l'état de santé d'une batterie

Les résultats présentés jusqu'ici ont justifié la validité de l'approche proposée, considérant uniquement des signaux mesurés sur la batterie (I, U) lors d'usage réel d'un VE, le signal de puissance étant calculé à partir du courant et de la tension. L'investigation de motifs issus de la batterie a permis d'obtenir et de mettre en place des estimations de son état de santé.

Les méthodologies explicitées en section 3.3.2 consistant à une approche de classification ont démontré l'intérêt d'une telle approche, fournissant une indication macroscopique sur le niveau de vieillissement d'une batterie. De tels résultats obtenus sous forme de classe d'état de santé sont adaptés à la détection de défauts et au diagnostic global d'une batterie.

Cependant, dans le cadre d'un usage en VE réel, il est nécessaire d'obtenir une estimation d'état de santé continue. L'objectif de cette section est d'étendre la détection des motifs à la régression afin d'identifier un état de santé continu de la batterie, et ainsi de pouvoir mettre en place des estimations précises du niveau de vieillissement de cette batterie, toujours avec les contraintes d'une analyse en ligne des données.

3.4.1 Méthodologie

Pour obtenir un état de santé d'une batterie, à partir de motifs extraits lors de son usage réel sur un véhicule électrique, nous utilisons une méthodologie basée sur des méthodes à noyaux. Les fondements employés sont dans un premier temps explicités dans cette section, puis la mise en place de la méthodologie dans le cadre de l'estimation du niveau de vieillissement d'une batterie est détaillée par la suite.

3.4.1.1 Relevance Vector Machines

L'approche RVM (*Relevance Vector Machines*), initialement introduit par Tipping [93], est basé sur une formulation Bayésienne d'un modèle linéaire avec un *prior* adapté proposant des résultats dans une représentation parcimonieuse. Considérons l'ensemble d'apprentissage $B^{train} = \{(x_1, t_1), \dots, (x_N, t_N)\}$ constitué des motifs x_i et de leur niveau de vieillissement associé t_i , et considérons que $t_i = f(x_i) + \epsilon_i$, où les bruits ϵ_i sont supposés être indépendants et de distribution Gaussienne avec une moyenne nulle et une variance σ^2 , i.e. $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, $\forall i \in \{1, \dots, N\}$.

Le but est ici d'apprendre un modèle de dépendance entre les valeurs cibles t_i et les entrées x_i . Ce qui permettrait pour des nouveaux motifs non observés x d'obtenir des prédictions précises t . Typiquement, les prédictions sont basées sur une fonction $f(x)$ définie sur l'espace des observations, et l'apprentissage peut se définir comme un processus de déduction de ces paramètres. Cette fonction prend la forme :

$$f(x) = \sum_{i=1}^N w_i K(x, x_i) + w_0 \quad (3.15)$$

où $K(x, x_i)$ est une fonction noyau et $w = [w_1, \dots, w_N]^T$ sont les poids et w_0 le biais.

Plus de détails sur l'approche RVM peuvent être trouvés en annexe C.

3.4.1.2 Extension du RVM aux noyaux multiples

L'utilisation de différent noyaux au sein du même procédé permet la combinaison de différentes caractéristiques. Dans le cas de phénomènes complexes, une approche à noyaux multiples peut être pertinente puisque chacun des M noyaux utilisés est sujet à extraire une caractéristique différente, en fonction de leurs définitions et des paramètres employés [148]. Afin d'assigner un noyau spécifique à chaque phénomène, la fonction de prédiction est notée :

$$f(x) = \sum_{i=1}^N \sum_{j=1}^M w_{i,j} \cdot K_j(x, x_i) + w_0 \quad (3.16)$$

Une manière possible d'écrire cette fonction est de définir une base de noyaux. Cette définition décompose un noyau K en différent blocs, formant un noyau multiple. Pour M noyaux, ce noyau multiple est composé comme une base de noyaux de la forme :

$$K = [1 \ K_1 \ K_2 \ \dots \ K_M] \quad (3.17)$$

Si nous considérons que toutes les colonnes sont indépendantes, nous pouvons finalement écrire la fonction de prédiction comme :

$$f(x) = \sum_{i=1}^{N \cdot M} w_i \cdot K_i(x) + w_0 \quad (3.18)$$

Cette formulation 3.18 montre que l'on peut étendre le RVM à un cas d'un noyau multiple avec une approche fondée sur une base de noyaux.

3.4.1.3 Noyaux utilisés

L'application de telles méthodes à noyaux requiert la définition desdits noyaux. Définissons ici successivement plusieurs formes de noyaux K entre deux variables x_k et x_i , qui seront considérés dans la suite de l'étude.

Le noyau Gaussien K_{Gauss} est défini, avec un paramètre σ correspond à la largeur de bande, par :

$$K_{Gauss}(x_k, x_i) = \exp\left(-\frac{\|x_k - x_i\|^2}{2\sigma^2}\right) \quad (3.19)$$

Afin d'utiliser la métrique DTW dans l'algorithme RVM, il est nécessaire de définir une forme de noyaux basée sur cette distance. Plusieurs tentatives ont été faites pour employer la distance DTW sous la forme d'un noyau. Ainsi, le noyau associé à la Gaussian Dynamic Time Warping (GDTW) [149], avec un paramètre γ , est explicité par l'équation 3.20.

$$K_{GDTW}(x_k, x_i) = \exp(-\gamma d_{DTW}(x_k, x_i)) \quad (3.20)$$

Notons que ce noyau K_{GDTW} n'est cependant pas théoriquement défini positif.

Certaines méthodes à noyaux, telle que la méthode Support Vector Machine (SVM), nécessitent l'usage de noyaux définis positifs. De ce fait, des versions alternatives au

noyau K_{GDTW} ont été proposées cherchant à obtenir un noyau défini positif, notamment par l'emploi des contraintes de Sakoe & Chiba (section 3.3.1.4). Le noyau K_{SC} est donc défini à partir de cette distance $d_{DTW_{SC}}$:

$$K_{SC}(x_k, x_i) = \exp(-\gamma d_{DTW_{SC}}(x_k, x_i)), \quad (3.21)$$

où DTW_{SC} représente la distance d_{DTW} réduite par une bande définie par Sakoe & Chiba. Notons que ce noyau K_{SC} est pseudo défini positif.

En utilisant la distance d_{GA} explicitée en section 3.3.1.5, un noyau K_{GA} associé est formé par :

$$K_{GA}(x_k, x_i) = \sum_p \exp(d_{GA}(x_k, x_i)), \quad (3.22)$$

avec p les chemins d'alignements possibles.

Une variante de ce noyau également basée sur la distance d_{GA} est utilisée (section 3.3.1.5), en utilisant le même principe que les critères de Sakoe & Chiba (équation 3.11), autorisant une bande de chemins définie par un paramètre T . Ce noyau est noté K_{TGA} , basé sur le Triangular Global Alignment (TGA).

3.4.2 Résultats

Le cadre décrit est appliqué sur les données réelles provenant des batteries étudiées (section 2.2). La méthodologie considérée est testée avec des motifs extraits avec des critères d'accélération 20 à 40 km.h⁻¹ ou bien 10 à 60 km.h⁻¹. Comme il a été évoqué précédemment, un modèle d'accélération plus long contient plus d'informations. Toutefois, les profils les plus longs d'accélération exigent de plus grands ensembles de données afin d'obtenir suffisamment de motifs variés pour le procédé d'apprentissage.

Les données utilisées sont séparées, de la même manière que lors de la section 3.3.3, en une partie d'apprentissage B^{train} et une partie de test B^{test} . Les données d'apprentissage sont employées pour construire un modèle d'estimation qui est par la suite mis en œuvre dans des conditions d'usage embarqués.

La première étape consiste à créer un modèle RVM permettant d'estimer l'état de santé d'une batterie durant son usage, uniquement à partir de ces signaux. L'erreur e_2 entre les estimations $\{t_1, \dots, t_n\}$ du niveau de capacité d'une batterie, et les valeurs réelles $\{t_1^*, \dots, t_n^*\}$ est donc calculée par l'équation :

$$e_2 = \frac{1}{n} \sum_{i=1}^n \left| 1 - \frac{t_i}{t_i^*} \right| (\%) \quad (3.23)$$

Une approche préliminaire consiste à étudier les performances obtenues avec l'utilisation d'un unique noyau, construit selon les définitions présentées en 3.4.1.3. La méthodologie développée est ici appliquée à des noyaux simples construits à partir de motifs extraits de la batterie Bat_3 , avec des critères d'accélération de 10 à 60 km.h⁻¹. Ces résultats sont présentés en table 3.4.

Les performances présentées par ce tableau 3.4 démontrent la bonne précision de la méthodologie présentée, avec l'usage d'un noyau unique. Les erreurs d'estimations e_2 calculées étant autour de 2.50 % pour l'estimation du niveau de capacité d'une batterie. Ce niveau de performance devient la référence pour la suite des expérimentations

	I	U	P
K_{GDTW}	2.66	2.65	2.49
K_{SC}	2.71	2.66	2.51
K_{GA}	2.80	2.50	2.20
K_{TGA}	3.21	2.27	2.88

TABLE 3.4 – Comparatif des erreurs e_2 d'estimation du SOH (en %) obtenues pour différents noyaux pour chaque type de motifs (I, U, P), avec un modèle RVM dont les paramètres sont définis par validation croisée avec la batterie Bat_3 , en **gras** les meilleurs résultats

durant lesquelles l'objectif sera de réduire ces erreurs. D'autre part, ce tableau met en évidence l'amélioration des résultats par l'emploi de la distance d_{GA} , par rapport à la distance d_{DTW} . En revanche, l'extension des notions par l'emploi de critères de Sakoe & Chiba dans la définition de chemins n'améliore pas les résultats. En effet, les noyaux K_{SC} et K_{TGA} sont associés à des performances d'estimations moins bonnes que les noyaux K_{GDTW} et K_{GA} . Cette remarque nous amène donc à considérer dans la suite des expérimentations uniquement ces deux noyaux K_{GDTW} et K_{GA} .

Comme expliqué dans la section 3.4.1.2, l'utilisation de plusieurs noyaux peut engendrer une augmentation de l'information contenue. L'ajout de variables dans la construction du modèle permettant de prendre en compte de potentielles nouvelles informations. Ainsi, dans l'application de cette méthodologie nous comparons les résultats obtenus pour chacune des combinaisons de noyaux possibles.

Dans cette approche, il est donc permis de prendre en compte des noyaux construits à partir des valeurs de température (T) et de niveau de charge (SOC) correspondant aux instants d'extraction des motifs. Ces variables peuvent alors ajouter de l'information relative aux conditions d'extraction des motifs et donc permettre une meilleure précision d'estimation de l'état de santé d'une batterie. Notons que ces variables T et SOC sont toujours prises en compte dans les modèles par le biais de noyaux Gaussien K_{Gauss} . Au contraire, les variables I, U, P sont, elles, employées soit avec un noyau K_{GDTW} , soit avec un noyau K_{GA} .

Le comparatif effectué sur les motifs extraits de la batterie Bat_3 , lors d'accélération de 10 à 60 km.h⁻¹, est présenté en table 3.5. Ce tableau fournit un détail des différentes erreurs d'estimations e_2 obtenues pour différentes combinaisons de noyaux. Ces différentes combinaisons consistent en l'utilisation d'au moins d'une des trois variables I, U, P puis en l'usage optionnel des variables T et SOC . Les paramètres respectifs de ces modèles RVM sont ici définis préalablement au calcul de l'erreur d'estimation par un procédé de validation croisée.

Notons les excellents niveaux de précisions obtenus par la méthodologie proposée. Ceux-ci proposent en effet une erreur d'estimation moyenne inférieure à $e_2 = 2\%$ pour la plupart des modèles, avec une meilleure performance de $e_2 = 0.45\%$ avec la combinaison des noyaux $K_{GA}(I)$, $K_{GA}(P)$, $K_{Gauss}(T)$ et $K_{Gauss}(SOC)$.

Dans le cadre d'une utilisation de tels modèles en embarqué dans un véhicule électrique, le niveau de précision global ($e_2 < 2\%$) présenté par cette méthodologie permettrait donc d'obtenir sans protocole spécifique des estimations régulières et fiables de l'état de santé de la batterie. Des erreurs d'estimation inférieures à $e_2 = 1\%$ per-

	2 variables		3 variables		4 variables
	I, U	I, SOC	I, P, T	I, U, T	I, U, P, T
K_{GDTW}	2.01	2.50	0.94	0.77	0.92
K_{GA}	1.70	2.34	0.57	0.78	0.71
	U, T	U, SOC	I, U, SOC	I, P, SOC	I, U, T, SOC
	K_{GDTW}	0.99	2.39	1.77	0.94
K_{GA}	0.82	0.84	0.85	0.93	0.58
	I, P	P, T	U, P, T	U, P, SOC	I, U, P, SOC
	K_{GDTW}	1.86	2.41	1.19	1.54
K_{GA}	2.05	2.11	0.81	0.94	0.72
	I, T	P, SOC	I, T, SOC	U, T, SOC	I, P, T, SOC
	K_{GDTW}	2.37	2.34	2.21	0.96
K_{GA}	2.29	2.15	0.59	0.75	0.45
	U, P		I, U, P	P, T, SOC	U, P, T, SOC
	K_{GDTW}	1.70		1.55	1.49
K_{GA}	1.76		1.68	0.95	0.83

TABLE 3.5 – Comparatif des erreurs e_2 d'estimation du SOH (en %) de la méthode RVM, en fonction de la combinaison des noyaux pour des accélérations de 10 à 60 km.h⁻¹ avec la batterie *Bat*₃, en **gras** les meilleurs résultats pour chaque nombre de variables

mettent d'obtenir des informations sur l'état de santé d'une batterie avec un important degré de confiance. En effet, pour une autonomie maximale réelle d'un véhicule de 100 km, une erreur d'estimation de $e_2 = 1$ % ne présente pas d'inconvénient majeur pour l'utilisateur.

Les résultats démontrent également la plus grande précision des noyaux K_{GA} par rapport aux noyaux K_{GDTW} , basés sur la DTW. Cette remarque confirme l'un des désavantages lié à la DTW ayant conduit à la construction du GA, par l'utilisation du *soft-minimum* (section 3.3.1.5). De plus, le noyau K_{GA} fournit de meilleures propriétés théoriques que les noyaux basés sur la DTW [145]. La figure 3.21 présente la répartition des résultats de cette table 3.5, pour les combinaisons entre noyaux K_{GA} et K_{Gauss} , en fonction du nombre de variables considérées par les modèles.

Cette figure 3.21 illustre l'amélioration globale des estimations par l'ajout de nouvelles variables dans la construction du modèle RVM, justifié par la prise en compte d'un plus grand nombre d'informations identifiant les effets du vieillissement de la batterie. L'ajout des variables de température permet ainsi d'améliorer considérablement les résultats, ce qui traduit l'impact important des phénomènes climatiques dans les dynamiques des motifs. Il est néanmoins important de noter que les valeurs de températures ne permettent pas à elles seules d'obtenir quelque conclusion sur le niveau de capacité d'une batterie. En effet, la construction d'un modèle RVM uniquement basé sur la combinaison des noyaux $K_{Gauss}(T)$ et $K_{Gauss}(SOC)$, sur les mêmes données que celles utilisées dans la table 3.5, délivre une erreur d'estimation très élevée de $e_2 > 70$ %.

La figure 3.22 présente les estimations obtenues dans l'un des meilleurs cas détaillé dans la table 3.5. Il apparait la très bonne précision d'estimation de l'état de santé de la méthodologie explicitée. Dans ce cas, les estimations présentent une erreur moyenne

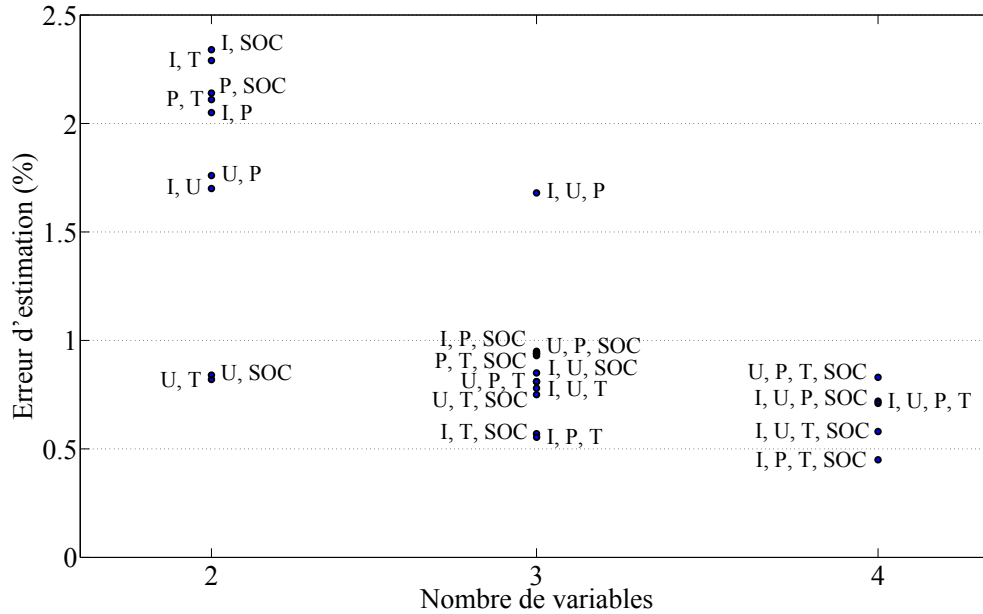


FIGURE 3.21 – Répartition des erreurs e_2 d'estimation de SOH (en %), en fonction des variables considérées par l'algorithme RVM, avec des noyaux K_{GA} et K_{Gauss} , pour des accélérations de 10 à 60 km.h⁻¹ avec la batterie Bat_3

de $e_2 = 0.45$ % du niveau de capacité (table 3.5).

Un tel niveau de précision démontre la complète habileté de l'algorithme développé dans cette étude à pouvoir estimer régulièrement le niveau de santé d'une batterie, uniquement à partir de données simplement mesurables, lors des usages réels et sans aucune préconisation particulière. Il est néanmoins important de noter le fait que la base d'apprentissage utilisée dans cette section couvre l'ensemble de la période étudiée. Dans l'optique d'une extension de ces travaux il est nécessaire d'étendre cette base d'apprentissage afin de calculer l'erreur d'estimation du SOH sur un période différente de celle ayant été utilisée pour construire le modèle.

3.5 Déduction du temps de vie restant d'une batterie

3.5.1 Méthodologie

La méthodologie d'estimation de l'état de santé d'une batterie est appliquée en ligne sur des batteries en usage. Ce modèle permet d'obtenir, à différents instants, une estimation de l'état de santé de la batterie à l'étude. Se basant sur ses estimations du SOH $\{t_{i*} | i = 1, \dots, N\}$, le but est désormais de parvenir à prédire le temps restant de vie de la batterie, appelé Remaining Useful Life (RUL). De la même façon que dans le chapitre précédent (section 2.5.1), le RUL est défini comme la durée restante avant que la batterie n'atteigne le critère de fin de vie, noté End Of Life (EOL), choisi comme étant un niveau de SOH de 80 %.

Le RUL est donc obtenu à partir de l'interpolation des estimations de SOH à un

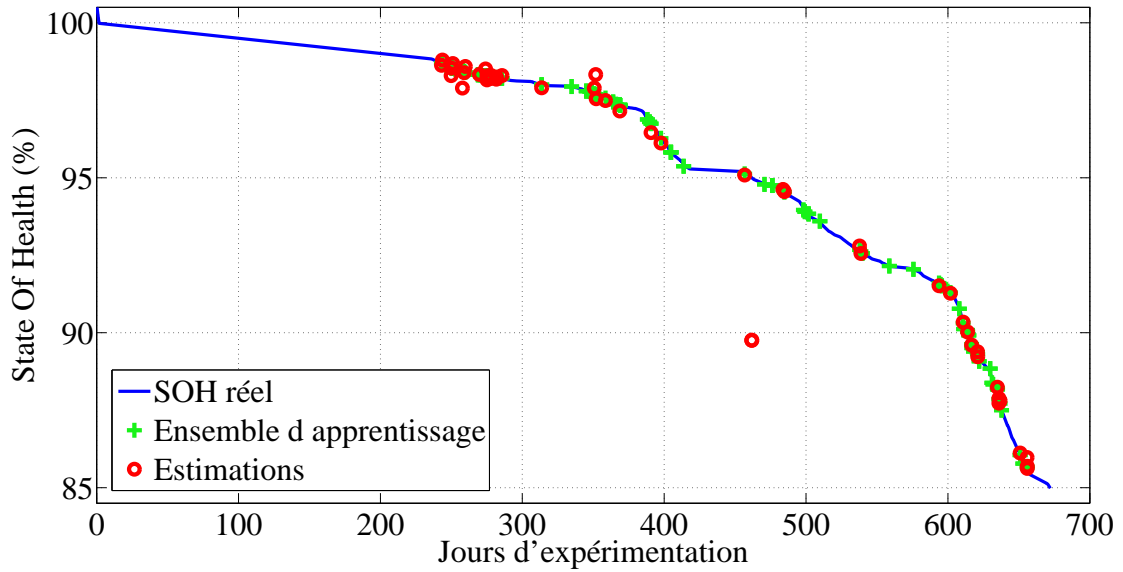


FIGURE 3.22 – Exemple du meilleur résultat d'estimation de SOH de la table 3.5, par RVM avec des noyaux $K_{GA}(I)$, $K_{GA}(P)$, $K_{Gauss}(T)$ et $K_{Gauss}(SOC)$, pour des accélérations de 10 à 60 km.h⁻¹ de la batterie *Bat*₃, avec $e_2 = 0.45\%$

instant donné. L'instant de fin de vie est déterminé comme étant l'instant correspondant au franchissement du critère prédéterminé EOL par la courbe d'interpolation des SOH. Contrairement à l'estimation continue du SOH, ce processus n'est pas nécessairement permanent et peut ainsi être effectué à des intervalles de temps réguliers. Ceci correspond à la dynamique du SOH relativement lente, et qui ne peut pas par conséquent modifier significativement le RUL d'un usage à l'autre.

Formellement, à un instant V , la méthodologie proposée considère une régression polynomiale pour représenter les estimations de SOH, obtenues jusqu'à ce temps V , $\{t_{i^*}(v_i) | i = 1, \dots, N\}$, avec $v_i \leq V$. La régression polynomiale estime les coefficients d'un polynôme $P(v)$ de degré d qui s'adapte aux estimations des états de santé t_{i^*} à un instant v . Ainsi, le polynôme P de degré d est défini tel que :

$$P(v) = p_0 + p_1 v + \dots + p_d v^d = \sum_{j=0}^d p_j v^j. \quad (3.24)$$

L'objectif est de minimiser la différence entre l'estimation du polynôme P aux points t_i , et les valeurs expérimentales t_{i^*} . Pour ceci, on utilise la méthode des moindres carrés, cela revient à minimiser la somme des carrés des écarts des erreurs entre valeurs expérimentales et valeurs calculées par le polynôme. Les seules variables à déterminer sont alors les coefficients p_j du polynôme. Nous minimisons donc la quantité :

$$E = \sum_{i=1}^N (t_{i^*} - P(v_i))^2 = \sum_{i=1}^N \left(t_{i^*} - \sum_{j=0}^d p_j v_i^j \right)^2. \quad (3.25)$$

Cette valeur est alors considérée comme une fonction des variables $\{a_j | j = 0, \dots, d\}$. Pour minimiser E , il suffit alors de trouver, pour chaque variable a_q , la valeur annulant

la dérivée partielle de E par rapport à cette variable a_q . Pour tout q tel que $0 \leq q \leq d$, cela s'écrit alors :

$$\frac{\partial E}{\partial a_q} = 0 \Rightarrow 2 \sum_{i=1}^N \left(-v_i^q \left(t_{i*} - \sum_{j=0}^d p_j v_i^j \right) \right) = 0 \quad (3.26)$$

$$\Rightarrow \sum_{i=1}^N t_{i*} v_i^q = \sum_{i=1}^N \left(\sum_{j=0}^d p_j v_i^{j+q} \right) \quad (3.27)$$

$$\frac{\partial E}{\partial a_q} = 0 \Rightarrow \sum_{i=1}^N t_{i*} v_i^q = \sum_{j=0}^d \left(p_j \sum_{i=1}^N v_i^{j+q} \right), \forall q | 0 \leq q \leq d \quad (3.28)$$

Cela revient à écrire sous forme matricielle $B = M \cdot A$, avec :

$$\begin{cases} A = (p_0 \quad \dots \quad p_d), \\ B = \begin{pmatrix} \sum_{j=1}^N t_{i*} v_i^0 \\ \vdots \\ \sum_{j=1}^N t_{i*} v_i^d \end{pmatrix}, \\ M = \begin{pmatrix} \sum_{i=1}^N v_i^0 & \dots & \sum_{i=1}^N v_i^d \\ \vdots & & \vdots \\ \sum_{i=1}^N v_i^d & \dots & \sum_{i=1}^N v_i^{2d} \end{pmatrix} \end{cases} \quad (3.29)$$

Avec ces notations, la solution devient un système linéaire $A = M^{-1}B$, permettant alors d'identifier les variables p_j formant le polynôme P (équation 3.24).

Considérant, cette construction du polynôme P , le but est de construire une densité de probabilité du RUL. Pour cela, nous utilisons une technique de bootstrap [150]. Cette méthode est basée sur la création de nouveaux échantillons de données obtenus par tirage avec remise à partir de l'échantillon initial, ce procédé est appelé *rééchantillonnage*. L'application du bootstrap à l'échantillon composé des estimations de SOH fournit un échantillonnage statistique de ces estimations, générant chacun une prédiction du RUL.

Au temps V , nous échantillonnons les estimations passées du SOH $\{t_{i*}(v_i) | i = 1, \dots, N\}$, avec $v_i \leq V$. Ce qui permet alors, par bootstrap, l'obtention de B échantillons de N estimations $\{t_{i*}^B | i = 1, \dots, N\}$. À partir de chacun de ces B ensembles formés, nous calculons le polynôme P correspondant aux estimations $\{t_{i*}^B | i = 1, \dots, N\}$. Chacun de ces polynômes $\{P_k | k = 1, \dots, B\}$ permet alors la détection du temps restant d'utilisation associé \widehat{RUL}_k .

Nous obtenons donc par ce processus une famille de B prédictions du RUL $\{\widehat{RUL}_k | k = 1, \dots, B\}$. La distribution de ces \widehat{RUL}_k fournit alors une prédiction de la fonction de densité du temps de vie restant au temps V .

3.5.2 Application et résultats

Utilisons à présent la méthodologie décrite dans la section 3.5.1. Ainsi, nous employons l'historique des estimations disponibles à un instant V . La création d'un polynôme adapté à ces estimations permet alors l'obtention d'une estimation du temps

de vie (RUL) restant. Dans le cadre de l'estimation du SOH, nous privilégions le choix de polynômes de degré $d = 2$, du fait de la forme observée par les références de SOH (figures 2.9 et 2.10). La figure 3.23 montre un polynôme, formé à partir d'estimations de SOH obtenues par la méthodologie présentée en section 3.4.1, proposant une prédiction du RUL.

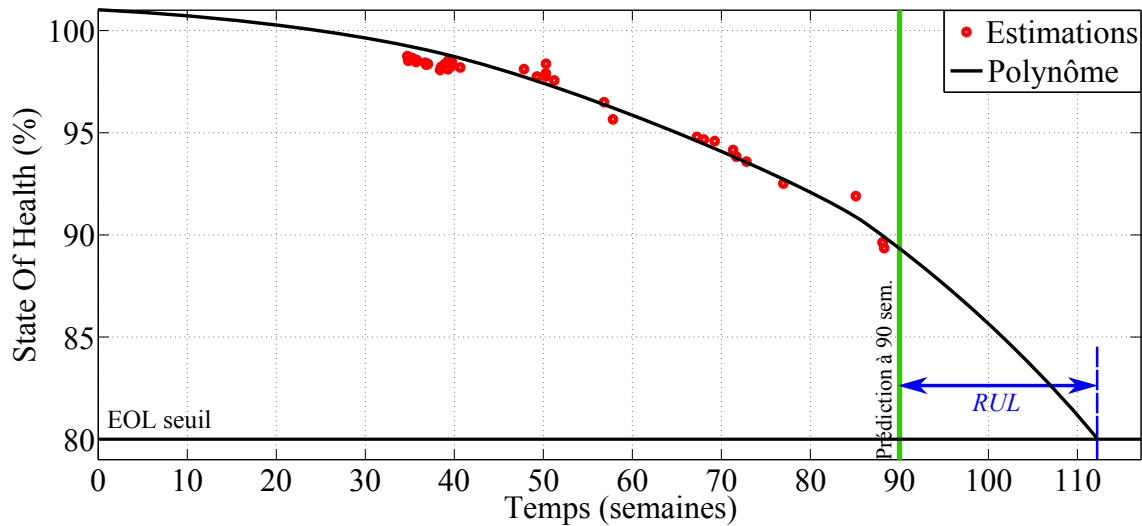


FIGURE 3.23 – Illustration d'un polynôme basé sur des estimations de SOH, générant une prédiction de RUL

La prédiction du RUL peut être faite à chaque instant T , induisant selon la méthodologie explicitée, la formation d'une densité de probabilité (*pdf*) de ce RUL prédit. Ce procédé est illustré pour trois instants, par la figure 3.24.

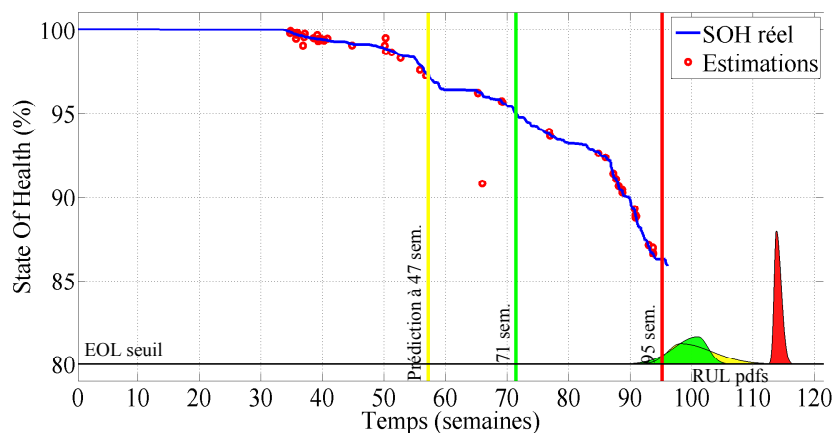


FIGURE 3.24 – Prédiction du RUL de la batterie Bat_2 , à partir des estimations du SOH obtenues par l'analyse des motifs d'accélération, à trois différents instants

Cette figure 3.24 illustre l'évolution de la confiance en la prédiction en fonction de l'instant considéré. En effet, plus les connaissances du comportement du SOH sont accumulées, plus la prédiction \widehat{RUL} est précise. Le niveau d'incertitude du RUL se réduit

alors au fil des estimations de SOH considérées. La première prédiction du RUL a un horizon de 310 jours, c'est à dire 44 semaines (pour une densité de 95 %) alors que la dernière a un horizon de moins de 120 jours (17 semaines). Cette remarque ne présente pas d'inconvénient majeur à l'utilisation de cette méthodologie, en effet lorsque la durée de vie restante est importante il est plus difficile d'obtenir une prédiction précise du fait de la complexité des phénomènes de vieillissement de la batterie.

L'analyse des prédictions et des incertitudes du SOH obtenues sont résumées dans la table 3.6. La diminution de la proportion de l'incertitude PI calculée à partir de l'intervalle de confiance IC_{RUL} du SOH et l'instant de la prédiction T , selon l'équation :

$$PI = \frac{L(IC_{RUL})}{\min(IC_{RUL}) - T}, \quad (3.30)$$

avec $L(IC_{RUL})$ l'incertitude représentant la longueur de l'intervalle de confiance IC_{RUL} .

Prédiction (semaines)	EOL (semaines)	Intervalle de confiance à 95 % (semaines)	Incertainde (semaines)	Proportion d'incertainde (%)
47	98	[91,110]	19	0.43
71	100	[91,105]	14	0.70
95	114	[112,117]	5	0.29

TABLE 3.6 – Comparatif de l'incertitude de prédiction de RUL à différent instants

La table 3.6 met en évidence la diminution de l'incertitude au fil des expérimentations, ce qui permet de diviser la plage de l'intervalle de confiance par quatre entre la première et la troisième prédiction.

D'autre part, nous pouvons observer un décalage des prédictions faites à chacun des trois instants pris en compte. Ainsi, la prédiction faite après 47 semaines de mise en service de la batterie associe une fin de vie aux alentours de 100 semaines d'expérimentation, alors que la prédiction au 95^{ème} semaine la fin de vie est prédite autour de la 114^{ème} semaine d'Usage. Ce décalage est alors expliqué par la différence de comportement observé dans les estimations obtenues.

Ces prédictions du RUL proposent alors une alternative à la méthodologie proposée au chapitre précédent (section 2.5.1). La figure 3.25 illustre le comparatif des prédictions obtenues pour chacune des deux méthodes.

Cette figure 3.25 représente la grande précision obtenue par la méthodologie développée à partir de l'extraction simple de motifs. En effet, les estimations de SOH ont prouvé leur fiabilité et nous avons ici une prédiction du RUL *pdf* d'une très faible incertitude (5 semaines), alors que la méthode proposée au chapitre précédent (chapitre 2) admet une incertitude de 13 semaines. Cette remarque permet de conclure favorablement à une utilisation possible de la méthodologie d'extraction de motif en embarqué pour la prédiction du RUL d'une batterie en utilisation réelle, considérant de plus l'adéquation avec les contraintes temps réel de la mise en œuvre de ce procédé.

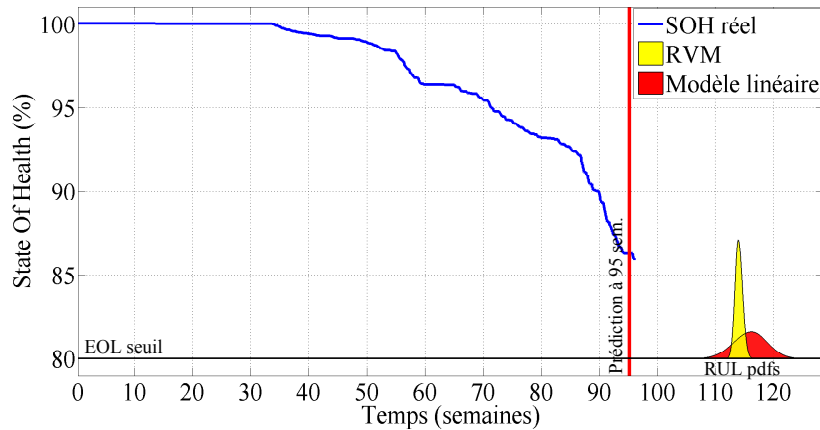


FIGURE 3.25 – Comparatif des densités de probabilités des RUL, selon deux méthodes d’estimation du SOH proposées (RVM et modèle linéaire du chapitre 2), pour les données de la batterie Bat_2

3.6 Conclusion - Discussion

Ce chapitre a, dans un premier temps, mis en évidence la modification comportementale de signaux issus de la batterie. Une méthodologie d’extraction de motifs spécifique a été détaillée, permettant d’obtenir un ensemble de motifs correspondant à des conditions d’usages prédéterminés. Dans la suite de l’étude ces conditions d’usages ont alors été définies comme étant des phases d’accélération ou de freinages spécifiques.

L’ensemble de ces signaux extraits fournissent alors une base de données permettant l’identification de différences de dynamiques de la batterie. Ainsi, l’approche d’analyse de motifs de courant, tension et de puissance proposée dans ce chapitre met en évidence les potentielles informations qu’ils contiennent. Une identification des différences de dynamiques entre motifs pour des états de santé de la batterie différents, justifie en partie la validité d’une telle approche uniquement basée sur les signaux mesurés. En effet, les motifs moyens représentés pour différents états de santé présentent des comportements différents, et difficilement interprétables.

À partir de ces observations, il est envisageable de procéder à une méthodologie de classification des motifs permettant d’obtenir un diagnostic d’état de santé d’une batterie simplement avec l’utilisation des signaux correspondant à un profil vitesse prédéterminé. Cette approche détaillée dans la section 3.3.2 fournit de bon résultats de classification. Ce qui permet d’envisager l’application ce type d’approche dans le cadre d’un suivi mené de manière non détaillé, par exemple lors d’une analyse rapide des niveaux de santé des batteries d’une flotte de véhicule. Une approche de classification du niveau de santé d’une batterie peut également être employée pour la détection de dysfonctionnement de la batterie ou d’une détection de ruptures, dans le cas où les motifs extraits mettraient en évidence un état de santé considérablement dégradé.

Une approche basée sur des méthodes à noyaux est ensuite développée. L’utilisation de méthode à noyaux autorise la prise en compte de plusieurs variables, ce qui autorise alors la capacité de combiner plusieurs informations disponibles. L’approche développée dans ce chapitre est basée sur les RVM, et démontre la très bonne précision

des estimations continues faites en embarqué, uniquement à partir des motifs extraits de la batterie.

L'ensemble des méthodes évoquées dans ce chapitre sont donc uniquement basées sur les signaux directement mesurés, et ont démontré leur capacité à obtenir des résultats satisfaisant, sans prise en compte de modèle physique ou électrochimique, et ce pour tout type de batterie étudiée. Une telle méthodologie présente alors une approche novatrice permettant de traiter la problématique de l'estimation de l'état de santé d'une batterie lors de ses utilisations, sans protocole spécifique.

Cependant, l'application des méthodologies développées nécessite l'utilisation d'une base de données étiquetées obtenues à partir d'expérimentations faites sur le même type de batterie que celles analysées par la suite. Cette base de données forme une base d'apprentissage permettant l'application des algorithmes explicités dans ce chapitre, que ce soit la classification de l'état de santé ou bien d'estimation de SOH. La base d'apprentissage est utilisée pour la construction d'un ou plusieurs modèles qui seront embarqués lors des usages d'un véhicule électrique.

L'amélioration des performances obtenues par les méthodologies développées dans ce chapitre nécessite l'usage d'une base de données exhaustive, contenant des motifs provenant des différentes conditions d'expérimentation possibles. L'obtention d'une telle base d'apprentissage étant longue et onéreuse, il doit être envisagé de considérer la relation entre la précision des estimations de SOH et les données d'apprentissage disponibles, comme introduit par la figure 3.18.

La mise en place en embarqué de ces approches présente donc l'avantage d'être simplement exécutable lors de roulages réels. Le seul procédé consiste à construire au préalable les modèles basés sur les données d'apprentissage. Ces modèles peuvent ensuite être appliqués en embarqué sur un VE en usages réels, fournissant une estimation du SOH de la batterie employée (ou une classe de SOH selon la méthode choisie), lors de chaque accélération spécifique détectée. Le procédé effectué en embarqué est uniquement l'extraction des motifs de signaux mesurés, d'après un critère de vitesse prédéfini, puis d'appliquer le (ou les) modèle(s) construit(s) en amont à ce motif, permettant l'obtention en temps réel d'une (ou plusieurs) estimation(s) du niveau de santé de la batterie. La mise à jour de ces algorithmes peut être faite hors ligne, néanmoins une telle mise à jour requiert l'augmentation de la base d'apprentissage utilisée afin de recalculer de nouveaux modèles d'estimation.

Chapitre 4

Vers le suivi d'une flotte de véhicules électriques

4.1 Objectifs et problématique	120
4.2 Classification des usages	121
4.2.1 Méthodologie	121
4.2.1.1 Mesure de similarité entre usages	123
4.2.1.2 Clustering des usages	125
4.2.2 Application de la classification	127
4.2.2.1 Exemple 1	127
4.2.2.2 Exemple 2	129
4.3 Aide aux préconisations d'usages	130
4.3.1 Méthodologie	130
4.3.2 Application de cette méthodologie	133
4.3.2.1 Exemple 1	133
4.3.2.2 Exemple 2	136
4.4 Mise en œuvre de modèles d'estimation sur une flotte	138
4.4.1 Modèles	139
4.4.2 Résultats sur une flotte	140
4.4.2.1 Estimation du SOH	140
4.4.2.2 Influence des préconisations d'usages	141
4.4.2.3 Amélioration des estimations	142
4.5 Gestion de la santé des batteries d'une flotte	145
4.5.1 Méthodologie de prise en compte de check-up	146
4.5.1.1 Estimations confirmées	146
4.5.1.2 Estimations non confirmées	147
4.5.2 Stratégies d'optimisation	148
4.5.2.1 Choix de l'instant d'un check-up	148
4.5.2.2 Véhicule choisi pour un check-up	150
4.5.2.3 Combinaison des stratégies	151
4.5.3 Exemple d'application	151
4.5.4 Perspectives	154
4.6 Conclusion - Discussion	155

4.1 Objectifs et problématique

Le suivi d'une flotte de véhicules est devenu un besoin majeur et indispensable des entreprises utilisant plusieurs véhicules. De nombreux logiciels existent désormais, proclamant un suivi adapté des véhicules permettant, principalement une réduction des coûts de fonctionnement [151, 152, 153, 154, 155]. Ces logiciels de *fleet management* sont en effet conçus pour optimiser l'usage des différents véhicules en fonction des contraintes de conducteurs disponibles et de trajets. Leur concept est basé sur la minimisation des temps de repos d'un véhicule, sur l'optimisation de la productivité humaine, et des trajets effectués pour réduire les consommations et surtout les coûts associés à ces usages. De tels logiciels sont basés sur un objectif de production et ne considèrent pas les contraintes de maintenance.

Cette stratégie est donc valable et largement répandue dans les méthodes de gestion de flotte actuelles. Cependant aucune des hypothèses faites n'est actuellement compatible avec les contraintes imposées par les véhicules électriques (VE). L'application de ces logiciels de suivi de flotte avec des VE n'est donc pas cohérente.

D'autre part, les solutions de suivi de performances de ces outils ne sont en réalité que des simples suivis de consommation des véhicules et ne permettent pas d'effectuer un diagnostic sur l'état de santé d'un véhicule et de ses composantes. Les analyses de performances, basées sur les données, vendues par ces logiciels ne sont, par exemple, que des possibilités de visualisation ou d'affichage des historiques de vitesse ou de données GPS. Celles-ci n'offrent pas de quantification réelle de modes d'Usage et du vieillissement des batteries composant la flotte.

Il est important de remarquer la présence des tels outils, de suivi des performances et des maintenances, au sein des arguments de vente des logiciels de suivi de flotte. Ceci confirme l'intérêt des gestionnaires de flottes pour l'obtention d'informations détaillées sur les performances de leurs véhicules, pour les détections automatique de possible dysfonctionnements. Cependant, les logiciels actuels ne semblent pas proposer réellement de telles fonctionnalités et se contentent donc d'afficher des historiques des données directement mesurées.

Étudions désormais plus spécifiquement le contexte des VE. Le développement d'un outil de suivi d'une flotte composée uniquement de VE comporte des problématiques inhérentes aux caractéristiques des VE. L'optimisation des utilisations des véhicules et des conducteurs est totalement distincte de l'application aux véhicules thermiques. En effet, les recharges des VE nécessitent leur immobilisation pour une longue durée. Comme nous l'avons argumenté lors du chapitre 1, la solution d'utilisation des recharges rapides ne constitue qu'un remplacement d'un problème par un autre. Ce type de recharge induisant des phénomènes de dégradation considérablement plus importants (section 1.1.2). Cette problématique nécessite donc un procédé d'optimisation complexe, entièrement dédié à cette application.

Comme détaillé lors du chapitre 1, la connaissance précise de l'autonomie d'un véhicule est primordiale pour un utilisateur individuel (*range anxiety*). Cette information est encore plus importante pour un gestionnaire de flotte qui désire optimiser ces trajets et donc utiliser les capacités limites de ces véhicules. Le risque de prendre un VE pour un trajet, sans être certain qu'il sera capable d'atteindre la destination souhaitée, peut provoquer des conséquences coûteuses pour le gestionnaire. Un outil de gestion de flotte conçu spécifiquement pour les VE doit fournir l'autonomie disponible pour

chaque véhicule. Cet objectif nécessite la mise en place d'un indicateur d'état de santé (SOH) préalablement à l'estimation de l'autonomie maximale disponible. Un tel indicateur du SOH propose l'optimisation de la gestion des flottes avec la prise en compte de l'état de santé des batteries utilisées.

Il est également utile d'identifier les facteurs impactant l'état de santé de chaque batterie. Cela permettrait de mettre en évidence des conditions d'Usage contraignantes, proposant une aide à la gestion des conditions de vie des batteries. Concrètement, un VE composé d'une batterie identifiée comme ayant un vieillissement accéléré, provoqué par de nombreux trajets de longue distance, peut se voir utilisé pour des distances plus courtes, au contraire d'un autre VE ayant été détecté comme subissant de longs stockages. Une telle stratégie aurait alors pour conséquence d'optimiser la durée de vie globale des batteries d'une flotte et par conséquent de pouvoir optimiser au mieux la longévité d'une flotte de VE.

Dans ce chapitre nous développons des outils adaptés à la gestion de flotte dans le cas de VE. Le premier objectif est de mettre en place un suivi des Usages des véhicules de la flotte, en terme de contraintes liées au vieillissement des batteries, parmi les variables identifiées comme impactant le vieillissement lors du chapitre 2. Autrement dit, les conditions de vie de chacun des VE sont étudiées et comparées, ce qui rend possible le regroupement des VE par similarités d'Usage. Ce procédé fournit une mise en place de préconisation d'usages, avec pour objectif d'éviter d'éventuelles dégradations prématurées des batteries utilisées. Le second outil développé dans ce chapitre a pour but de permettre un suivi de diagnostic des batteries à l'étude à partir des méthodologies évoquées dans les chapitres précédents, et donc d'estimer le SOH de chacune de ces batteries à partir des données collectées. Ces estimations sont basées sur la méthode d'estimation du SOH hors ligne notée Estimation Hors Ligne (EHL), et en ligne notée Estimation En Ligne (EEL), respectivement décrites dans le chapitre 2, et le chapitre 3. L'amélioration et la mise à jour de ces estimations, par la combinaison des informations apportées par ces deux méthodes, est explorée. Enfin, l'impact de possibles caractérisations complètes de batterie est détaillé afin de réduire l'incertitude des estimations par le biais de différentes stratégies évoquées et comparées. La figure 4.1 présente l'approche globale utilisée dans ce chapitre.

4.2 Classification des usages

La première approche consiste en une analyse des Usages de chacun des VE de la flotte étudiée. Le terme d'Usage de batterie, employé dans ce chapitre, correspond aux conditions globales de vie de cette batterie (stockages, cyclages et environnementales). L'objectif est de parvenir à identifier de potentielles similarités d'usage entre les VE composant la flotte considérée, fournissant par la suite des préconisations adaptées à chaque groupe d'Usage. La méthodologie décrite dans cette section correspond à la mise en place d'une classification automatique des VE en fonction de leurs modes d'Usages, sur une même période.

4.2.1 Méthodologie

Afin de décrire le plus explicitement possible les conditions de vie d'une batterie, nous sélectionnons, d'après les études précédentes, les variables énumérées en table

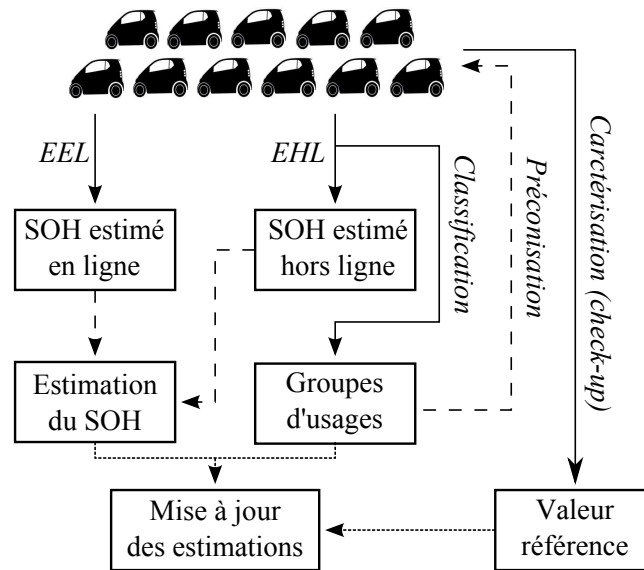


FIGURE 4.1 – Méthodologie globale développée dans ce chapitre pour la gestion de l'état de santé des batteries d'une flotte de véhicules

4.1. Ces variables proviennent ou sont calculées à partir de mesures directement issues des batteries. Les 7 variables évoquées sont choisies par leurs simplicités de mesures et par leurs représentativités des conditions vécues par la batterie étudiée pour le suivi des usages des VE.

Période	Variables
Charge ou décharge	T_{bat} , $E_{bat,sorti}$, ΔSOC , Nb_{altI}
Stockage	t_{cal} , SOC_{cal} , T_{cal}

TABLE 4.1 – Liste des variables utilisées dans la méthodologie de suivi des usages

Pour un VE donné, chacune de ces variables contient de nombreuses valeurs enregistrées. Lors d'un stockage, les trois variables citées en table 4.1 sont calculées, à raison d'une valeur par stockage. Les charges et décharges permettent de mesurer l'ensemble des autres variables évoquées en table 4.1, selon une fréquence régulière de 10 Hz. Pour réduire les volumes de données, nous synthétisons chacune de ces variables, lors de charges ou de décharges, par une unique valeur. Autrement dit, une décharge permet d'évaluer la température moyenne d'un roulage, le nombre de changements de signe cumulés du signal de courant, le ΔSOC global de cette décharge et la quantité totale d'énergie délivrée.

Un exemple de données obtenues à partir d'une charge (Ch) et de deux périodes de

stockages (St_1, St_2) , sur un véhicule j , est détaillé dans l'équation 4.1.

$$V_j = \begin{pmatrix} \text{mean}(T_{bat}(Ch)) \\ \sum (E_{bat,sorti}(Ch)) \\ \Delta SOC(Ch) \\ \sum Nb_{altI}(Ch) \\ t_{cal}(St_1), t_{cal}(St_2) \\ SOC_{cal}(St_1), SOC_{cal}(St_2) \\ T_{cal}(St_1), T_{cal}(St_2) \end{pmatrix} \quad (4.1)$$

Par souci de simplification des notations, dans la suite de l'étude le terme $V_j(v)$ désigne les valeurs de la $v^{\text{ème}}$ variable mesurée sur le véhicule j .

4.2.1.1 Mesure de similarité entre usages

La classification des usages entre VE requiert la définition d'une métrique permettant la comparaison et la quantification des différents Usages. Une telle métrique est ici basée sur les variables présentées en table 4.1 représentatives des conditions d'utilisation des batteries associées.

Considérant une période de comparaison P_c entre deux VE V_1 et V_2 , il est nécessaire de définir un critère de quantification entre les usages mesurés. Les variables mesurées sur chacun de ces véhicules sont comparées deux-à-deux. Par exemple, les valeurs de température batterie enregistrées lors de la période P_c , $V_1(i)$ et $V_2(i)$, sont comparées. L'objectif est désormais de définir une métrique entre deux vecteurs de longueurs différentes d'une même variable, par exemple entre $V_1(1) = \{T_{bat}^{V_1}(1), \dots, T_{bat}^{V_1}(n)\}$ et $V_2(1) = \{T_{bat}^{V_2}(1), \dots, T_{bat}^{V_2}(m)\}$. Ce problème intervient régulièrement dans le domaine du traitement d'images, sous la forme du *bag-of-features* [156].

Afin de prendre en compte les répartitions entre ces deux vecteurs, $V_1(v)$ et $V_2(v)$, nous basons le système de comparaison sur les fonctions de densité déduites de ces vecteurs. Pour considérer les répartitions sur le même intervalle, les deux histogrammes sont mesurés sur une plage de comparaison commune I pour les deux vecteurs étudiés, $V_1(v)$ et $V_2(v)$. Cette plage I est définie comme contenant l'ensemble des données des vecteurs $V_1(v)$ et $V_2(v)$. Ceci permet la création de deux fonctions de densité empiriques $f(V_1^v)$ et $f(V_2^v)$, évaluées sur le même intervalle I (figure 4.2). La fonction de densité empirique $f(V_j^v)$ de la variable v mesurée sur le véhicule j sera désormais notée $f_j^v = f(V_j^v)$.

Néanmoins, les valeurs acquises sont possiblement peu nombreuses du fait d'une période de comparaison P_c réduite, ou bien d'un faible nombre de charges/décharges effectuées. D'autre part, les valeurs extraites peuvent toutes être différentes d'un véhicule à l'autre, tout en étant très proches ce qui représente tout de même des Usages similaires. Lors de telles éventualités, la comparaison directe des densité empiriques, issues des données mesurées, devient difficile.

Pour pallier à cette problématique de comparaison, nous procédons, préalablement à la comparaison, à un lissage de chacune des fonctions de densité empirique sur un intervalle I_v^F , spécifique à la variable v étudiée, contenant la plage I . En effet, certaines variables, telles que le ΔSOC et la température T_{bat} , sont toujours contenues dans des intervalles fixes quelque soit le véhicule, pouvant donc être préalablement définis.

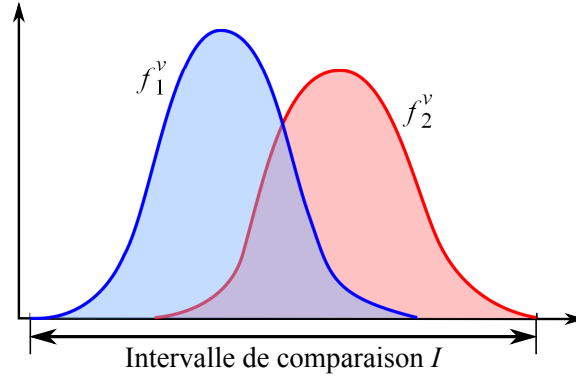


FIGURE 4.2 – Illustration de deux fonctions de densité f_1^v et f_2^v d'une même variable v , et de l'intervalle de comparaison I

Cette méthodologie permet de considérer les voisinages des valeurs et de réduire les erreurs générées. La fonction de lissage ici privilégié est un lissage Gaussien, appliqué à une fonction de densité empirique $f(V_j^v)$ de longueur n . Ce procédé de lissage est effectué selon l'équation 4.2.

$$\widehat{L}_j^v(x) = \frac{\sum_{i=1}^n K_\lambda(x, x_i) \cdot f_j^v(x_i)}{\sum_{i=1}^n K_\lambda(x, x_i)}, \text{ avec } K_\lambda(x, x_i) = \frac{1}{\lambda} \exp\left(-\frac{\|x_i - x\|^2}{2\lambda}\right) \quad (4.2)$$

Ce lissage Gaussien d'une fonction de densité f_j^v nécessite la définition d'un paramètre λ , contrôlant la largeur du voisinage considéré par le noyau K_λ . Ce paramètre λ est primordial pour obtenir une bonne régularisation. Dans ce travail nous définissons la valeur de ce paramètre par le *Silver's Rule Of Thumb (SROT)*, défini par l'équation 4.3 [157, 158], avec n la longueur de f_j^v et σ la variance de f_j^v .

$$\lambda_{SROT} = \left(\frac{4}{3}\right)^{1/5} \min\left(\sqrt{\sigma}, \frac{IQR(f_j^v)}{1.349}\right) \cdot n^{-1/5}, \quad (4.3)$$

avec IQR l'espace interquartile.

Cette approche fournit la fonction de densité empirique lissée \widehat{L}_j^v d'une variable v mesurée sur un véhicule j , sur un intervalle prédéfini. La comparaison des Usages de deux véhicules correspond alors à la comparaison deux-à-deux des fonctions de densité lissées de l'ensemble des 7 variables étudiées.

Nous définissons une mesure h quantifiant les similarités entre deux fonctions de densités empirique lissées d'une variable v entre deux véhicules V_1 et V_2 (équation 4.4).

$$h(V_1^v, V_2^v) = \bigcap \left(\widehat{L}_1^v(x_i), \widehat{L}_2^v(x_i) \right) \quad (4.4)$$

Notons que $0 \leq h(V_1^v, V_2^v) \leq 1$, car $\sum \widehat{L}_1^v = \sum \widehat{L}_2^v = 1$. La figure 4.3 présente une illustration visuelle de cette quantification des similarités entre densités, comme étant la proportion d'aire commune aux fonctions de densité.

À partir de cette définition de la similarité entre deux densités d'une même variable, nous avons une combinaison linéaire $D(V_1, V_2)$ entre deux véhicules V_1 et V_2

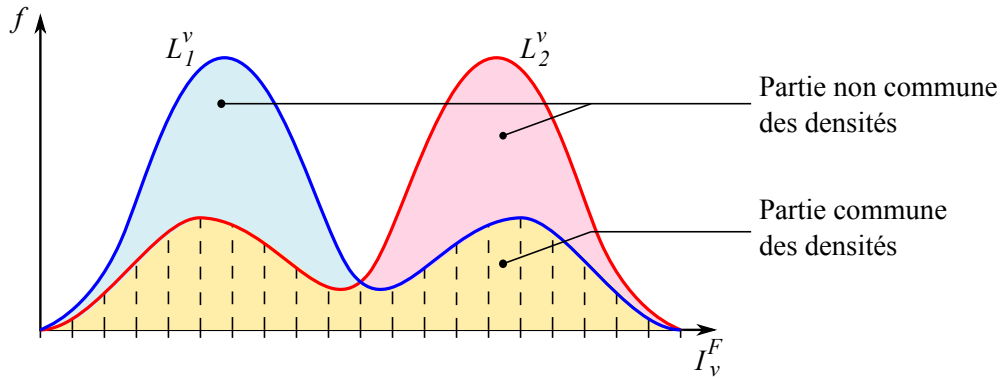


FIGURE 4.3 – Illustration du calcul de similarité entre deux fonctions de densité L_1^v et L_2^v

portant sur les 7 variables évoquées en table 4.1, calculée selon l'équation 4.5.

$$D(V_1, V_2) = \sum_{v=1}^7 \frac{1}{\alpha_v} \cdot h(V_1^v, V_2^v), \quad (4.5)$$

avec α_v un poids associé à la variable v . Notons que ces poids doivent être normés ($\sum \alpha_v = 1$) pour conserver la notion de métrique. L'imputation de ces poids permet d'ajouter à l'étude de potentielles importances auprès de chacune des variables considérées.

Les définitions introduites jusqu'ici permettent de quantifier les similarités d'Usage entre deux véhicules. Ce procédé autorise l'application de cette mesure D à l'ensemble des véhicules deux-à-deux.

4.2.1.2 Clustering des usages

La quantification des similarités d'Usage entre véhicules permet la mise en place d'une procédure de clustering des conditions d'utilisations des VE et de leur batteries. Nous considérons donc les similarités évaluées entre chacun des s véhicules composant la flotte lors d'une période P_c .

La méthodologie choisie utilise la Classification Ascendante Hiérarchique (CAH) présentée en section 3.3.2.3. Les individus étudiés sont ici des VE, et la métrique de comparaison est la distance D définie en équation 4.5. Le processus de regroupement des individus, effectué par la CAH, forme un arbre hiérarchique complet reliant tous les VE considérés. Cet arbre est ensuite élagué avec l'aide d'un critère d'élagage, à définir, impliquant la formation de différents groupes de VE (figure 4.4).

Dans la pratique, le critère d'élagage reste à déterminer afin d'obtenir les classes de VE, générées en fonction de leurs conditions d'Usage lors de la période P_c choisie. Ce critère d'élagage est ici déterminé par le biais du *gap statistics* [147].

Le concept du *gap statistics* repose sur les sorties des algorithmes de clustering, en comparant les changements de dispersion intra-clusters en fonction du nombre de clusters choisi. Considérons des véhicules V_j regroupés en k clusters $\{C_1, C_2, \dots, C_k\}$, avec C_r les indices des individus du cluster r , et $n_r = |C_r|$. Notons \mathbf{D}_r la somme des

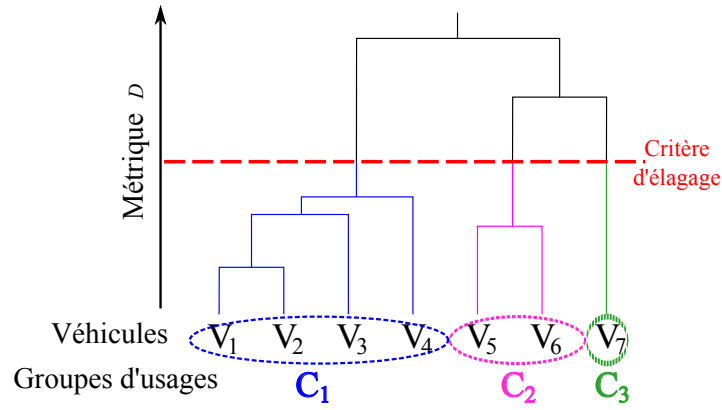


FIGURE 4.4 – Illustration de la Classification Ascendante Hiérarchique (CAH) mise en place selon la distance D entre les usages de 7 VE, avec un critère d'élagage formant trois groupes différents C_1 , C_2 et C_3

distances entre les individus d'un même cluster r :

$$\mathbf{D}_r = \sum_{j, j' \in C_r} D(V_j, V_{j'}) \quad (4.6)$$

Nous assimilons la somme intra-cluster normalisée des carrés autour du centre d'un cluster par W_k (équation 4.7).

$$W_k = \sum_{r=1}^k \frac{1}{2n_r} \mathbf{D}_r \quad (4.7)$$

L'idée fondamentale de la procédure du *gap statistics* est de standardiser les graphes de $\log(W_k)$ en les comparant avec une distribution de référence nulle des données. Le nombre optimal de clusters \hat{k} est alors la valeur k pour laquelle $\log(W_k)$ tombe le plus loin en dessous de cette courbe de référence. Ainsi, nous avons :

$$\text{Gap}_n(k) = \mathbf{E}_n^*(\log(W_k)) - \log(W_k), \quad (4.8)$$

avec \mathbf{E}_n^* l'espérance d'un échantillon de taille n de la distribution référence. Celle-ci est généralement obtenue à partir d'un échantillon uniforme des données originelles.

L'estimation \hat{k} est donc la valeur maximisant $\text{Gap}_n(k)$ après avoir pris la distribution d'échantillonnage en compte. Notons que cette méthodologie est également applicable à tous les algorithmes de clustering, et à toutes les mesures D .

Nous estimons $\mathbf{E}_n^*(\log(W_k))$ par la moyenne de B copies $\log(W_k^*)$, générées avec un échantillon de Monte Carlo $\{X_1^*, \dots, X_n^*\}$ bâti à partir de la distribution référence. Il est donc nécessaire d'évaluer la distribution d'échantillonnage du *gap statistics*. Notons $\text{std}(k)$ l'écart type des B répliquas de $\log(W_k^*)$, qui représente l'erreur de simulation, et est transformé en la quantité :

$$s_k = \sqrt{1 + \frac{1}{B}} \cdot \text{std}(k) \quad (4.9)$$

Finalement, le nombre optimal de cluster \hat{k} est le plus petit k tel que $\text{Gap}(k) \geq \text{Gap}(k+1) - s_{k+1}$. La figure 4.5 illustre les valeurs et la méthodologie évoquée. La figure

en haut à gauche induit un choix de $k = 3$ par la *elbow method*. Le *gap statistics* permet de formaliser cette observation visuelle. La figure en haut à droite présente la comparaison des $\log W_k$ et la moyenne des distributions moyennes. Enfin, les figures situées en bas illustrent l'évolution du gap (gauche), comportant un pic évident en $k = 3$ et le critère de choix (droite). Le \hat{k} sélectionné est le plus petit k pour lequel la quantité tracée en bleu devient positive. Cela donne donc un nombre optimal de clusters de $\hat{k} = 3$, qui s'avère ici être le nombre initial de groupes sélectionnés.

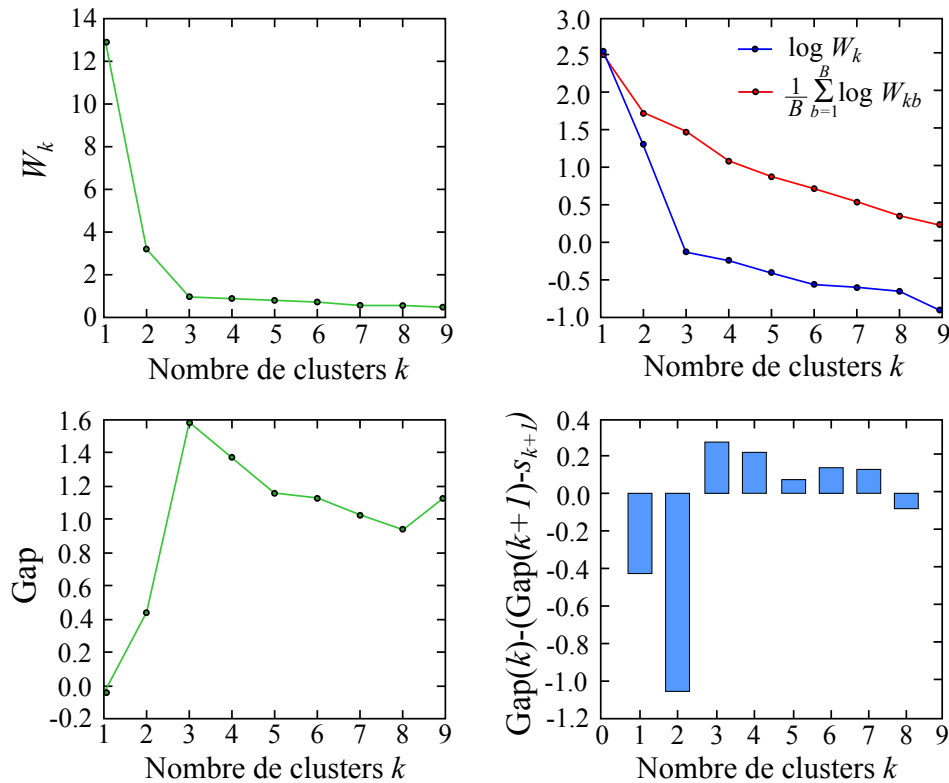


FIGURE 4.5 – Exemple de résultats obtenus par le *gap statistic*, pour des données réparties en 3 groupes

Cette procédure de classification forme la répartition des VE, composant une flotte, au sein de différents groupes formés à partir de leurs conditions d'Usage.

4.2.2 Application de la classification

4.2.2.1 Exemple 1

La méthodologie globale de classification des Usages mise en place dans la section 4.2.1 est ici appliquée aux données issues de 2 VE en usage réel. Afin de simuler la présence d'une flotte de véhicules parmi les données obtenues, nous procédons artificiellement à la séparation de ces données comme provenant de 20 VE différents $\{V_1, \dots, V_{20}\}$, grâce à la longue durée des expérimentations. Cette construction artificielle d'une flotte de VE, à partir des données de *Bat*₂ et *Bat*₃, est détaillée en figure 4.6. Par souci de simplification, nous prenons par la suite l'hypothèse que chacun de ces 20 VE a été utilisé lors d'une période commune de 4 semaines. Considérant cette

hypothèse de construction, on s'attachera à vérifier l'influence des températures ambiantes sur les constructions des groupes.

	Jan	Fév	Mar	Avr	Mai	Juin	Juil	Aoû	Sep	Oct	Nov	Déc	
Bat_3	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	V_{11}	V_{12}	V_1	V_2	2011
													2012
Bat_2	V_{19}	V_{20}					V_{13}	V_{14}	V_{15}	V_{16}	V_{17}	V_{18}	2011
													2012

FIGURE 4.6 – Décomposition des données sous la forme d'une flotte de 20 VE, avec 1 VE par mois d'usage

Étudions désormais une flotte de 20 VE ayant été utilisés sur une même période P_c . Le but est d'appliquer la méthodologie décrite en section 4.2.1.1, afin de séparer ces VE en différents groupes formés en fonction de leurs Usages respectifs lors de la période P_c . Pour cela, nous considérons les variables décrites en table 4.1 acquises lors d'une période P_c .

Dans cet exemple, nous définissons la période de comparaison P_c comme étant égale à quatre semaines complètes consécutives. Les 20 VE sont comparés à partir de leurs données enregistrées durant ces quatre semaines. La procédure de comparaison des Usages avec la métrique D est appliquée (section 4.5), construisant une matrice de distance entre chacun des VE étudiés. Le nombre optimal de clusters \hat{k} est ensuite déterminé par le *gap statistics*, comme valant $\hat{k} = 4$ dans cet exemple (figure 4.7).

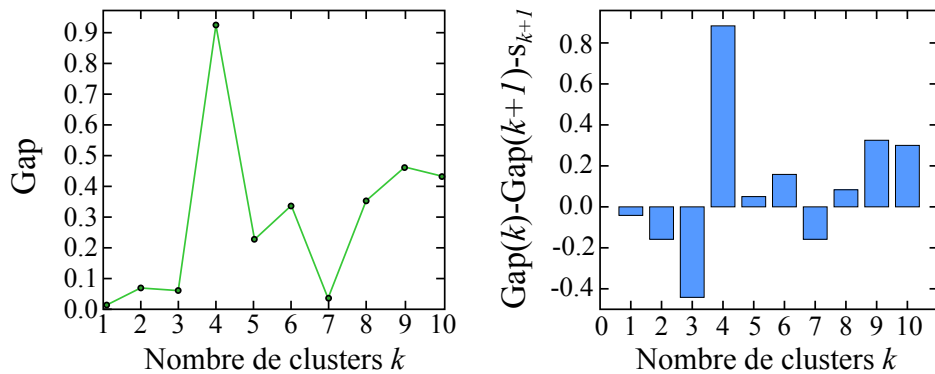


FIGURE 4.7 – Résultats du *gap statistic* effectué sur les données des 20 VE utilisés pendant 4 semaines, estimant un nombre optimal de groupe de $\hat{k} = 4$

L'arbre hiérarchique associé ainsi que l'application de l'élagage optimal associé à la création de 4 groupes de VE distincts sont illustrés en figure 4.8.

Dans l'exemple proposé ici, les 20 VE sont partitionnés en 4 groupes distincts. Ce qui implique la présence de quatre types d'Usages différents parmi l'ensemble des véhicules composant la flotte étudiée, lors du mois considéré. Nous pouvons observer la composition de ces 4 groupes semblant non affectée par les conditions de températures extérieures, les véhicules V_2 , V_6 et V_9 se trouvant dans la même classe alors qu'ils correspondent respectivement à un mois de Décembre, d'Avril et de Juillet.

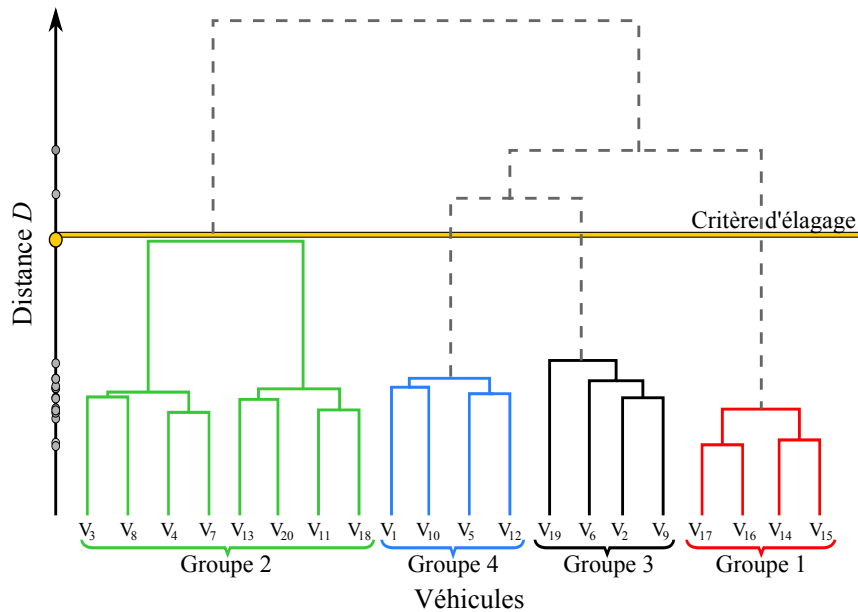


FIGURE 4.8 – Classification Ascendante Hierarchique de 20 VE utilisés lors de 4 semaines, avec la formation de 4 groupes de véhicules

4.2.2.2 Exemple 2

La méthodologie explicitée en section 4.2.1 peut également être effectuée pour des périodes P_c d'expérimentation successives, de manière dissociée ou bien en considérant une fenêtre glissante f_g . Par exemple, pour $f_g = 1$ semaine, la période considérée P_c est décalée d'une semaine entre chaque classification.

L'évolution de ces groupes de véhicules constitue un outil de suivi des Usages faits pour chacun des VE. Prenons ici les mêmes données que dans la section précédente 4.2.2.1. Ces données sont ici réparties en 10 VE $\{V_1, \dots, V_{10}\}$, utilisés sur 10 semaines chacun. Le processus de classification est alors appliqué sur l'ensemble de la flotte avec une période de $P_c = 4$ semaines, puis en considérant une fenêtre glissante de $f_g = 1$ semaine. Ceci revient à l'application de 7 classifications (1 initiale après P_c , puis 6 successives en considérant la fenêtre glissante f_g). Les résultats de ces classifications et l'évolution des différents groupes formés sont illustrés en figure 4.9.

La figure 4.9 permet l'analyse des modifications de comportements des VE. Ainsi, nous pouvons remarquer la continuité des conditions d'Usage pour la majorité des VE tout au long de l'étude. Il existe néanmoins quelques changements de groupes au sein de la flotte. En effet, seuls les VE $\{V_1, V_2, V_3, V_4, V_9, V_{10}\}$ restent dans le même groupe tout au long de l'étude. Cela nécessite donc une investigation plus approfondie par la suite, afin de justifier les raisons ayant provoqué ces changements.

Cette approche de classification met en évidence les similarités ou les différences d'Usage entre les véhicules. L'application de la méthodologie offre également le suivi des Usages d'une flotte. L'extension naturelle de ce processus consiste en l'identification et l'interprétation des conditions d'Usage des véhicules ayant provoqué ces séparations ou ces similarités et suivre leurs évolutions dans le temps.

Il est envisageable ici de considérer l'évolution des groupes eux-même afin d'identifier la modification de comportement d'un groupe entier. Cette notion peut être mise

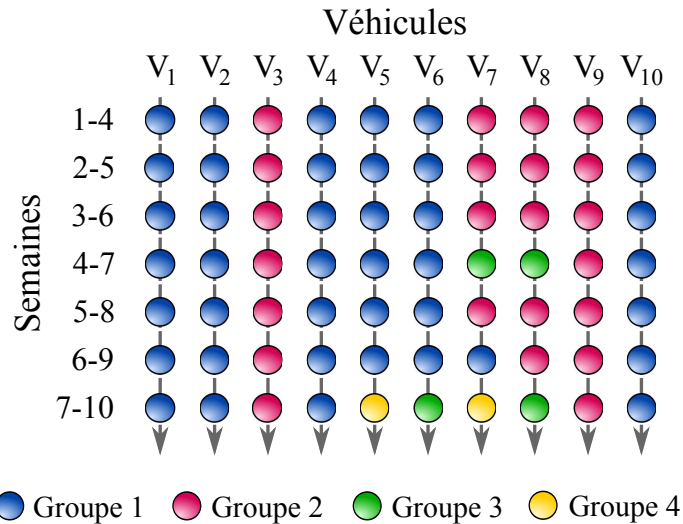


FIGURE 4.9 – Résultats de classification des usages d’une flotte de 10 VE lors d’une période de 10 semaines

en place par le biais d’un suivi d’un véhicule prototype représentatif du centre de ce groupe.

4.3 Aide aux préconisations d’usages

La méthodologie de classification des Usages mise en œuvre permet d’identifier des groupes d’Usage différents. L’objectif de cette section réside en l’utilisation des résultats de la méthodologie décrite en section 4.2.1.2, afin de permettre l’interprétation des groupes de VE au travers des descriptions des Usages n’étant pas initialement considérées pour le calcul des groupes.

L’identification des conditions discriminantes d’usage peut être utilisée, par un gestionnaire de flotte, dans le cadre de préconisations faites sur les usages futurs des VE.

4.3.1 Méthodologie

La production de préconisations pour les utilisations futures d’un VE et de sa batterie nécessite la mise en place d’une procédure spécifique. Celle-ci est basée sur l’application préalable de la méthodologie de classification des usages des VE composant la flotte étudiée. Considérons donc les k clusters $\{C_1, \dots, C_k\}$ dans lesquels sont répartis les s VE de la flotte, issus du processus décrit en section 4.2.1.

L’objectif est désormais d’interpréter les facteurs responsables de la formation de ces k clusters. Nous utilisons des variables différentes de celles ayant été employées lors de la construction de cette classification (énumérées en table 4.1). Les variables considérées pour l’interprétation des différents groupes d’usages sont listées en table 4.2. Ces variables correspondent à des indicateurs de la nature des roulages et du style de conduite employés, au contraire des variables de la table 4.1 qui représentent l’état de la batterie.

L’ensemble des variables de cette table 4.2 sont ensuite fusionnées par groupe.

Variables	
Mesurées	<i>Distance, Vitesse, T_{amb}</i>
Calculées	<i>Nb_{cycles}, PKE, I_{max}</i>

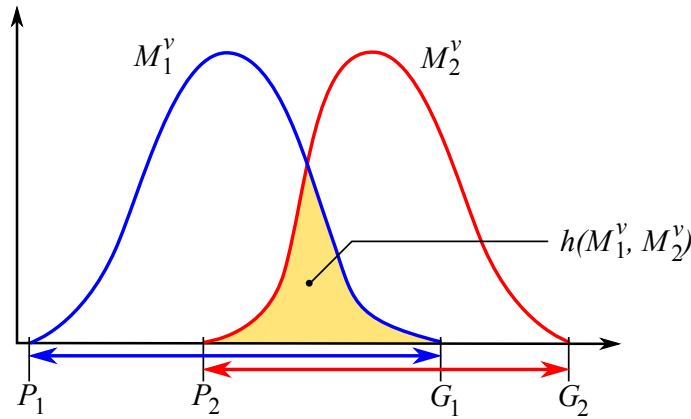
TABLE 4.2 – Liste des variables utilisées dans l'interprétation des classes d'Usage

Toutes les valeurs d'une même variable correspondant aux VE d'un même cluster sont regroupées par concaténation. Cela revient à considérer seulement k individus chacun représentatif d'un cluster C_r . La suite de la méthodologie détaillée cherche donc à quantifier et expliquer les différences de distribution des variables de la table 4.2 entre chacun de ces clusters.

Pour ce faire, nous estimons des fonctions de densité selon la méthode décrite en section 4.2.1.1. Toutes les variables de la table 4.2 sont utilisées pour la création de fonctions de densité empirique par groupe d'usage. La mesure M_r^v d'une variable v des VE représentative d'un cluster r est définie par :

$$M_r^v = \bigcup_{j \in C_r} V_j^v \quad (4.10)$$

L'impact respectif de chacune de ces variables dans la création des différents clusters est estimé entre un groupe et l'ensemble des autres groupes. Ce calcul est effectué à partir d'un taux de recouvrement entre les fonctions de densités empirique. Notons $P_r = \min(M_r^v)$ et $G_r = \max(M_r^v)$ les valeurs extrêmes, respectivement minimales et maximales, d'une fonction de densité M_r^v d'une variable v , et $h(M_1^v, M_2^v)$ la proportion de densité commune définie par l'équation 4.4 (figure 4.10).

FIGURE 4.10 – Illustration de deux fonctions de densité M_1^v et M_2^v d'une même variable v , et de la proportion de densité commune $h(M_1^v, M_2^v)$

Le taux de recouvrement $\tau(M_1^v, M_2^v)$, entre deux fonctions de densité M_1^v et M_2^v d'une même variable v , est défini par l'équation 4.11. Notons que par définition $0 \leq h \leq 1$ et $0 \leq \tau \leq 1$.

$$\tau(M_1^v, M_2^v) = \frac{\min(G_1, G_2) - \max(P_1, P_2)}{\max(G_1, G_2) - \min(P_1, P_2)} \cdot h(M_1^v, M_2^v) \quad (4.11)$$

Cette définition 4.11 quantifie les répartitions de densités, basé sur leurs intervalles d'observations pondéré par la proportion commune des densités h . Ce concept et les différents cas de recouvrement possibles sont illustrés en figure 4.11. Un indicateur μ est également introduit par l'équation 4.12 pour simplifier l'interprétation des répartitions. Ce taux de recouvrement τ est calculé entre toutes les variables de la table 4.2 et entre chacun des groupes d'Usages C_r .

$$\mu(M_1^v, M_2^v) = \begin{cases} +1 & , \text{ si } G_1 > G_2 \\ -1 & , \text{ sinon} \end{cases} \quad (4.12)$$

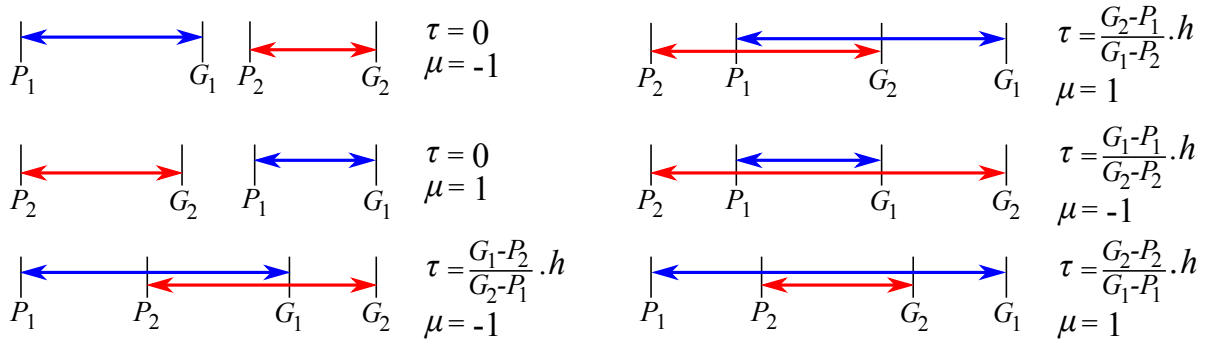


FIGURE 4.11 – Illustration des six différents cas du calcul du taux de recouvrement τ entre deux fonctions de densité M_1^v et M_2^v d'une même variable v , avec $h = h(M_1^v, M_2^v)$ la proportion de densité commune

Le taux de recouvrement τ présenté en équation 4.11 est défini pour deux groupes de VE. L'extension de ce procédé au cas de plusieurs groupes est simplement faite en considérant une approche *one-vs-all*, pour laquelle nous comparons une fonction de densité d'un groupe par rapport à la fonction de densité de l'ensemble des autres groupes.

Nous avons un taux de recouvrement τ pour chacune des variables mesurées sur les différents groupes de VE. Afin de considérer seulement les variables ayant un impact significatif dans la composition des différents groupes d'usages, il est nécessaire de fixer un seuil $0 \leq \tau_c \leq 1$. Les variables d'un cluster, telles que $1 - \tau > \tau_c$ sont identifiées comme contraignantes dans l'Usage des VE de ce groupe. La sensibilité liée à la définition de ce seuil τ_c implique un niveau de recommandation plus ou moins élevé. Celui-ci doit être fixé en fonction des besoins et des attentes du gestionnaire de la flotte.

Cette notion de recouvrement permet également d'obtenir des informations interprétables sur ces différences d'Usage entre VE. Autrement dit, un faible taux de recouvrement entre deux densités se traduit par un décalage de ces fonctions. La présence d'un tel décalage transmet une information sur les Usages effectués par les VE de la flotte étudiée. Par exemple, si une variable *Distance* d'un groupe 1 présente des valeurs inférieures à celles des autres groupes (deux premiers cas présentés dans la figure 4.11), on peut conclure à des parcours du premier groupe moins contraignants que ceux des autres groupes.

Ces interprétations sur les conditions des variable v , entre les différents groupes de VE, lors d'une période considérée, engendrent la proposition de possibles préconisa-

tions sur les conditions d'utilisations des VE. L'application de ce seuil τ_c est illustrée par la figure 4.12.

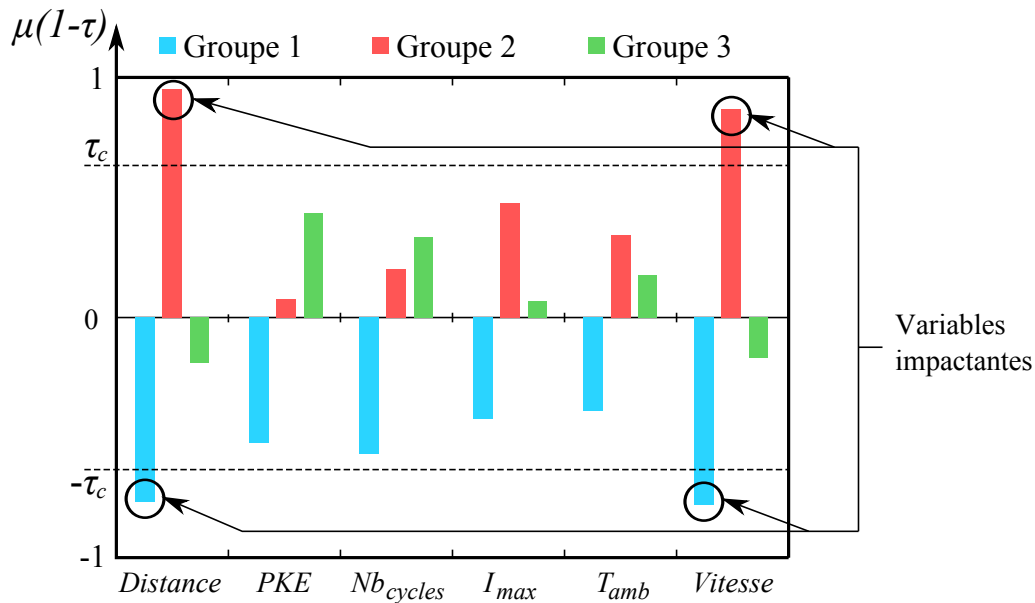


FIGURE 4.12 – Exemple d'impact des variables dans la construction de 3 groupes, avec l'application d'un seuil τ_c

4.3.2 Application de cette méthodologie

Dans cette section nous étudions l'application de la méthodologie décrite en section 4.3.1, en particulier pour l'interprétation des compositions des différents groupes d'usages.

Il existe deux cas d'applications à cette méthodologie :

- Déterminer à un instant donné les conditions d'usages responsables de la formation des groupes de EV, amenant à des préconisations évitant les Usages contraignants pour la batterie
- Faire un suivi sur le long terme de la classification des VE afin de détecter des changements d'usages significatifs.

Ces deux cas sont ici traités pour trouver deux exemples distincts, reprenant les hypothèses faites dans les exemples respectivement présentés dans les sections 4.2.2.1 et 4.2.2.2.

4.3.2.1 Exemple 1

Nous avons ici une flotte constituée de 20 VE en service pendant 4 semaines consécutives (cf. section 4.2.2.1). L'application du procédé de classification des usages de ces véhicules sur la période totale de $P_c = 4$ semaines génère la composition de $k = 4$ groupes distincts de VE (figure 4.8). L'objectif est d'obtenir une justification sur les différences d'Usage entre les 20 EV, ayant généré ce partitionnement. Pour cela, nous employons la méthodologie mise en place en section 4.3.1.

Les résultats obtenus sur l'exemple 1 présentés en section 4.2.2.1 sont ici utilisés. Les distributions des variables utilisées pour la description des Usages des groupes (listées en table 4.2) sont difficilement interprétables sans traitement préalable. La figure 4.13 illustre la répartition des valeurs obtenues pour chacune des variables au sein des différents groupes de VE, démontrant la complexité d'interprétation directe.

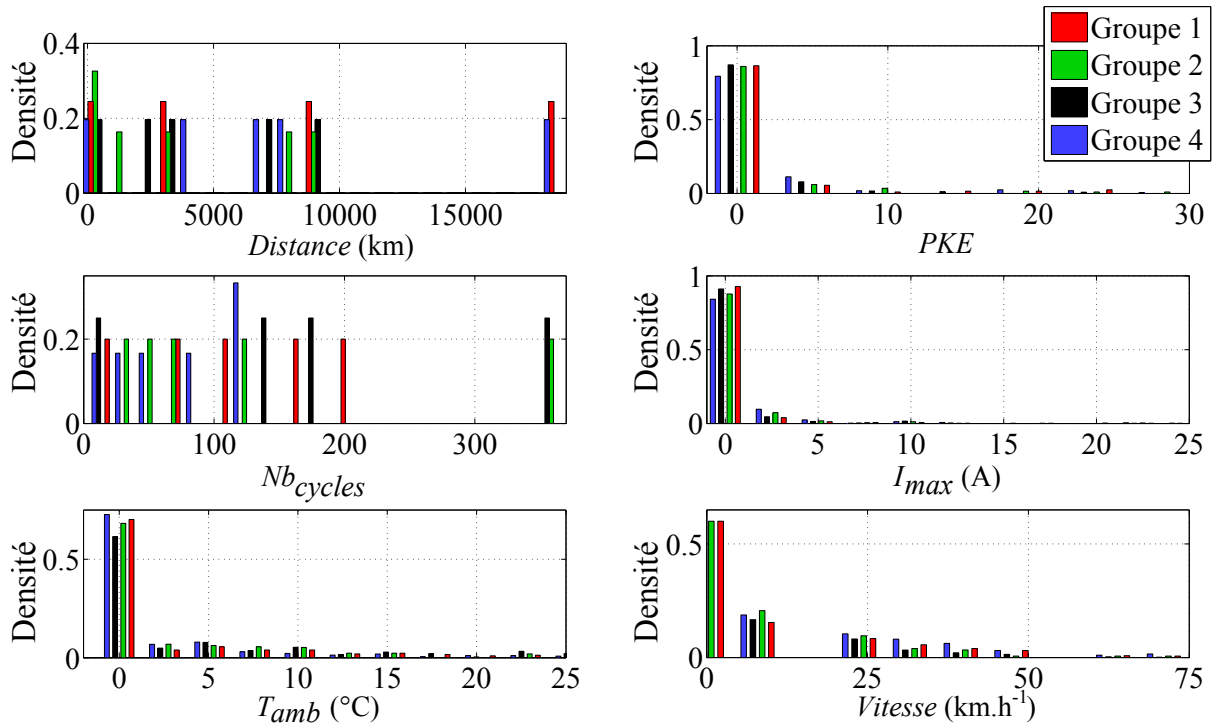


FIGURE 4.13 – Répartition des valeurs mesurées pour chacune des variables, en fonction du groupe d'Usage

Nous utilisons la méthodologie développée en section 4.3.1. Les impacts τ des variables sont calculés pour chacun des 4 groupes de VE. Il est nécessaire de définir un critère de seuil $\tau_c \in [0, 1]$, permettant la dissociation des variables impactantes.

En effet, le seuil τ_c est responsable de la sensibilité d'identification des variables comme étant influentes dans la séparation des usages de VE en différents groupes. Celui-ci doit être déterminé afin d'extraire automatiquement les conditions d'Usage caractéristiques de chacun des groupes de VE. Nous définissons ici ce seuil τ_c comme un paramètre dépendant du nombre de groupes de VE formés (équation 4.13). Ce choix est effectué empiriquement pour représenter une sensibilité cohérente sur le nombre de variables détectées comme impactant la formation des groupes.

$$\tau_c = \max(0.10, 0.60 - 0.10 \cdot \hat{k}) \quad (4.13)$$

Cette définition permet de fixer le seuil en fonction du nombre de groupes identifiés \hat{k} . Cela se traduit par un seuil bas dans le cas d'un grand nombre de groupe de VE, qui correspond aux conditions où les taux τ seront faibles, car résultant de la comparaison des répartitions entre tous les groupes.

Dans notre cas, le calcul du seuil explicité par la formule 4.13 donne $\tau_c = 0.20$.

L'application de ce seuil aux calculs précédents permet l'extraction des variables impactantes dans l'interprétation des différents groupes (figure 4.14).

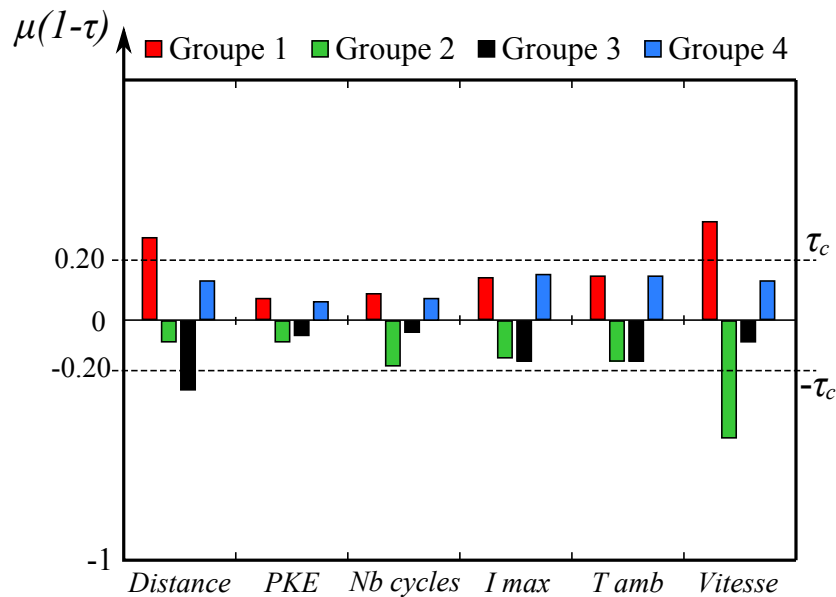


FIGURE 4.14 – Calcul de l'influence de chacune des variables dans la composition des différents groupes de VE, avec un seuil $\tau_c = 0.20$

Ces résultats fournissent une interprétation des Usages représentatifs des 4 groupes de VE composant la flotte. Nous pouvons déduire que les VE présents dans le groupe 1 sont utilisés avec des vitesses plus élevées comparativement aux autres groupes, pour de longues distances, ce qui peut se traduire par des Usages effectués en dehors des centres-villes, contrairement aux autres groupes. Les VE du groupe 2 s'avèrent, eux, être utilisés de manière peu rapide, ce qui semble typique des usages citadins. Le groupe 3 est lui identifié comme représentant des VE utilisés sur de faibles distances, ce qui peut par exemple être expliqué par une proximité domicile-travail. Les résultats également illustrent un comportement moyen des VE du groupe 4, semblant être représentatif d'un Usage intermédiaire.

Les variables identifiées comme significatives des Usages des différents groupes de VE sont énumérées dans la table 4.3.

	Variables significatives	
	Basse	Élevée
Groupe 1	-	<i>Distance, Vitesse</i>
Groupe 2	<i>Vitesse</i>	-
Groupe 3	<i>Distance</i>	-
Groupe 4	-	-

TABLE 4.3 – Liste des variables significatives des usages par groupe de EV

Tous ces résultats et ces interprétations sont confirmés visuellement par la figure 4.14. Par exemple, les VE composant le groupe 1 $\{V_{14}, V_{15}, V_{16}, V_{17}\}$ correspondent à

de longs trajets des véhicules. Ceux-ci présentent des variables de distance et de vitesse supérieures à celles mesurées sur les autres VE. Nous pouvons remarquer que toutes les classes sont associées à des PKE similaires, ce qui est caractéristique des véhicules en partage, tels que ceux étudiés ici. Dans cet exemple, les types de conduites ne sont pas détectés significativement différents, l'interprétation du comportement des conducteurs de cette flotte requiert la comparaison de ces valeurs de PKE par rapport à une référence¹.

Les VE du groupe 3 ont eux été utilisés lors de la période considérée lors de trajets de courte distance. Enfin, les éléments du groupe 4 sont identifiés comme utilisés de manière totalement moyenne. En effet, aucune variable n'est spécifiquement caractéristique des Usages des VE de ce groupe 4.

La méthodologie de détection des variables représentatives des Usages d'un groupe de VE démontre ici son potentiel dans les interprétations qu'elle induit. Nous avons donc retrouvé, à partir des résultats extraits, les Usages représentatifs de chacun des groupes. Ces conclusions sont de plus en accord avec les caractéristiques d'Usage de chacun des VE de la flotte.

En se basant sur ces résultats, des préconisations peuvent être donc établies afin de limiter les dégradations des batteries, suggérant une diminution des Utilisations des VE du groupe 1 et au contraire, pour remplir les demandes de la flotte, une Utilisation plus importante des VE du groupes 3. En revanche, le comportement des conducteurs concernés par les utilisations de cette flotte ne semblent pas engendrer de vieillissement prématuré. L'interprétation complète de ces résultats nécessitent la comparaison des conditions d'Usage avec des références externes d'usages moyens obtenus sur des véhicules cibles.

4.3.2.2 Exemple 2

Dans cette section, nous appliquons la méthodologie d'identification des variables significatives des Usages des VE composant une flotte. Nous considérons ici la flotte composée lors de la section 4.2.2.2. Nous avons une flotte comportant 10 VE utilisés lors de 10 semaines successives.

L'étude continue de cette flotte est faite par l'application successive d'une fenêtre glissante f_g d'une semaine. L'application de la méthodologie de classification des Usages des VE sur une période de 4 semaines présente des résultats évolutifs sur les 10 semaines étudiées (figure 4.9).

Nous pouvons donc analyser les modifications d'usages des véhicules, et en déduire les causes des changements de classification de certains VE d'une semaine sur l'autre. Afin de montrer le fonctionnement de cette méthodologie sur un exemple, nous considérons uniquement les modifications de comportement des VE de la flotte entre les semaines 6-9 et les semaines 7-10. Pour rappel, les classes attribuées aux VE de la flotte en ces instants sont présentées en figure 4.15.

L'objectif est donc d'interpréter les facteurs ayant généré ces changements de classification. Le calcul des influences respectives des variables dans la composition des groupes d'usage est effectuée. Les résultats de cette analyse pour les deux périodes étudiées, sont illustrés en figure 4.16. Notons que lors des semaines 7-10, les groupes 2

1. Un PKE significativement élevé pour une ou plusieurs classes impliquerait une conduite *agressive*.

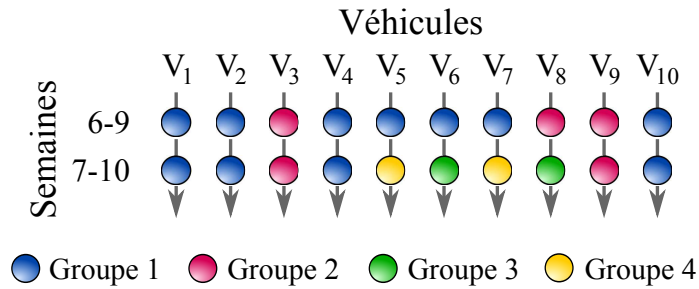
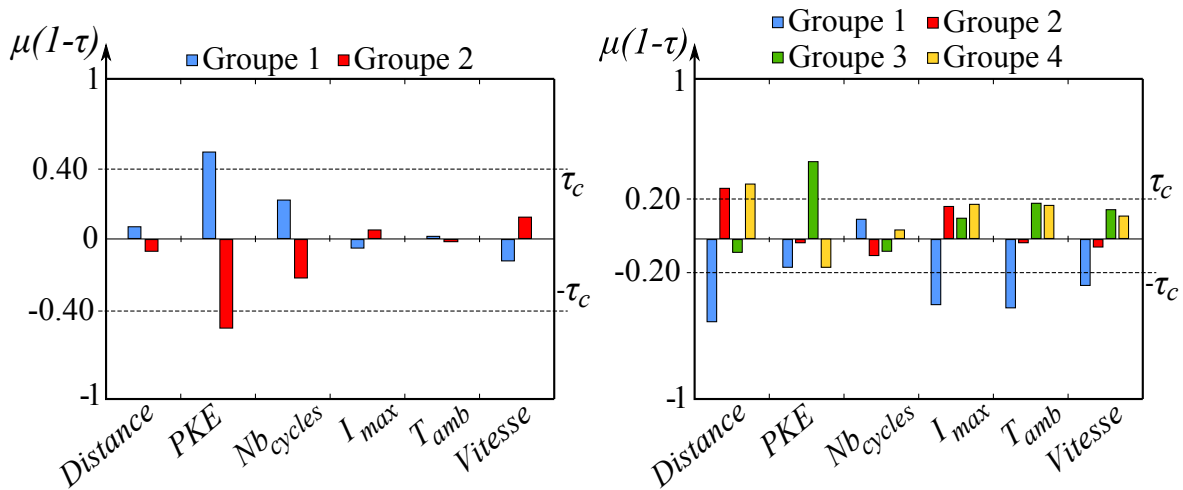


FIGURE 4.15 – Résultats de classification des usages d’une flotte de 10 EV, sur une partie d’expérimentation globale

et 4 présentent des caractéristiques similaires, seulement différenciées par le nombre de cycles et la température ambiante mesurée.



(a) Période des semaines 6-9, avec 2 groupes (b) Période des semaines 7-10, avec 4 groupes

FIGURE 4.16 – Calcul de l’influence de chacune des variables pour une flotte de 10 EV

La liste des variables impactantes pour chaque VE, lors des deux périodes considérées, est présentée en table 4.4. Cela permet de détecter de possibles changements d’usage d’un VE, et de comprendre les nombreux changements de classification observés.

L’interprétation de ces résultats illustre la légère modification du PKE calculé sur le véhicule V_8 entre ces deux périodes. Le comportement des utilisateurs de ce véhicule lors de la semaine 10 semble responsable du changement de classe de V_8 . D’autre part, le véhicule V_6 présente également des valeurs de PKE élevées, même si réduites lors des semaine 7-10, ces deux véhicules font donc l’objet de préconisations sur les comportements des conduites futures. Les répartitions de ces valeurs de PKE sont présentées en figure 4.17, illustrant notamment la modification d’Utilisation du véhicule V_6 .

Nous avons également remarqué sur la figure 4.15 le changement de classe des véhicules V_5 et V_7 . La table 4.4 permet de conclure à une réduction des PKE mesurés sur ces deux VE, accompagné par une augmentation significative de la longueur des

	Variables significatives			
	Semaines 6-9		Semaines 7-10	
	Basse	Élevée	Basse	Élevée
V_1	-	<i>PKE</i>	<i>Distance, I_{max}, T_{amb}, Vitesse</i>	-
V_2	-	<i>PKE</i>	<i>Distance, I_{max}, T_{amb}, Vitesse</i>	-
V_3	<i>PKE</i>	-	-	<i>Distance</i>
V_4	-	<i>PKE</i>	<i>Distance, I_{max}, T_{amb}, Vitesse</i>	-
V_5	-	<i>PKE</i>	-	<i>Distance</i>
V_6	-	<i>PKE</i>	-	<i>PKE</i>
V_7	-	<i>PKE</i>	-	<i>Distance</i>
V_8	<i>PKE</i>	-	-	<i>PKE</i>
V_9	<i>PKE</i>	-	-	<i>Distance</i>
V_{10}	-	<i>PKE</i>	<i>Distance, I_{max}, T_{amb}, Vitesse</i>	-

TABLE 4.4 – Liste des variables significatives des usages pour chaque VE, sur deux périodes successives

trajets effectués.

L'analyse des Usages des VE composant une flotte a permis d'identifier les comportements spécifiques de chaque véhicule, permettant également de procéder à des préconisations portant sur une ou plusieurs variables déterminées. Ces préconisations ayant pour but d'éviter des phénomènes de dégradations prématurées sur les batteries utilisées. Ce procédé a également démontré son intérêt dans le cadre du suivi d'une flotte sur plusieurs périodes. Un tel suivi d'une flotte s'avère utile pour l'analyse des évolutions des Usages effectués et ainsi de mieux contrôler l'ensemble des VE.

4.4 Mise en œuvre de modèles d'estimation sur une flotte

Les méthodes décrites précédemment permettent le suivi des Usages d'une flotte de VE, formant des groupes de VE basé sur leurs Usages. L'objectif de cette section est de mettre en place un procédé liant ces groupes d'Usage avec les estimations d'état de santé de batteries lors d'usages sur des VE. Pour cela, il est nécessaire de mettre en place une méthodologie d'estimation de l'état de santé des éléments d'une flotte étudiée.

Dans les chapitres précédents nous avons évoqué deux stratégies d'estimation de l'état de santé d'une batterie : une approche d'estimation hors ligne (notée ici EHL, chapitre 2), et une estimation en ligne (notée EEL, chapitre 3). Ces deux méthodes peuvent être utilisées et adaptées au cadre de suivi d'une flotte de VE. La figure 4.18 représente les approches proposées dans cette section.

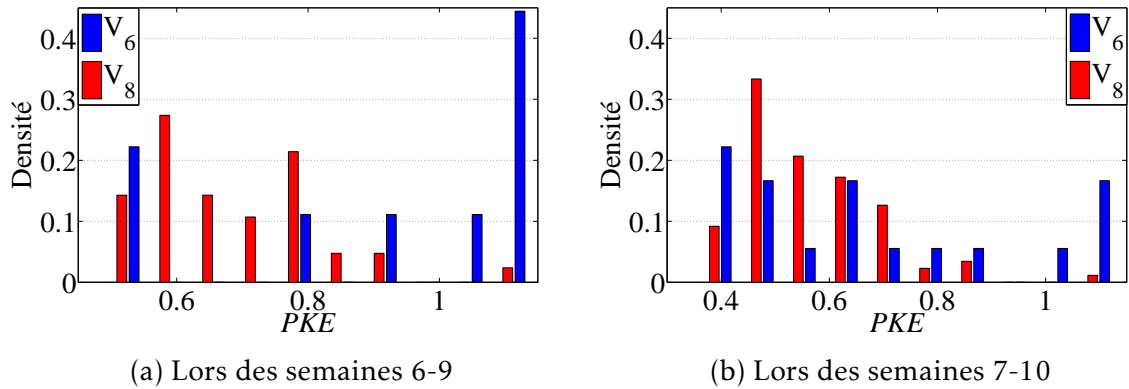


FIGURE 4.17 – Comparatif des densités des mesures de PKE des véhicules V_6 et V_8 sur deux périodes différentes

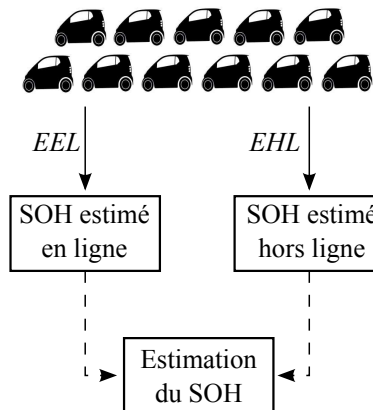


FIGURE 4.18 – Méthodologies d'estimation du niveau de performance des VE composant une flotte

4.4.1 Modèles

Nous considérons dans cette section les estimations d'état de santé d'une batterie obtenues à partir des méthodes développées précédemment.

La première stratégie d'estimation consiste à la méthode hors ligne basée sur des historiques de données enregistrées, notée EHL (section 2.4). Celle-ci peut être adaptée à une flotte de VE. Le modèle d'estimation est ici supposé construit au préalable selon la méthode décrite en section 3.3. Ce procédé détaillé précédemment fournit une estimation de SOH pour chacune des batteries de la flotte étudiée, au terme d'une période fixe.

L'alternative proposée consiste à prendre en compte les différents groupes d'usages de VE obtenus par la méthode présentée en section 4.2.1.2. Dans ce cas, une estimation d'évolution de SOH est calculée pour chacun des EV, après une période fixe. Une estimation de l'évolution de capacité est donc obtenue pour chaque batterie de la flotte, comme explicité dans l'exemple en section 2.4, chacune associée à un intervalle de confiance (figure 2.24).

L'application d'un tel modèle présente donc des estimations individuelles du SOH,

avec l'apport d'un intervalle de confiance, pour chacune des batteries de la flotte.

La deuxième stratégie possible considère la méthodologie d'estimation de SOH en ligne développée en section 3.3, notée EEL. Ce procédé estime l'état de santé d'une batterie en temps réel lors de ces Usages sur un VE. Chacun des VE de la flotte étudiée se voit attribuer ici des estimations de SOH provenant de profils de vitesse spécifiques, selon les critères définis en section 3.2.2.

Notons que ces deux méthodologies d'estimation du SOH de batterie sont compatibles et peuvent être mises en œuvre de manière simultanée. L'association de ces deux procédés permettant de conforter ou de préciser les estimations obtenues.

L'objectif est ici d'évaluer le comportement des estimations faites sur les VE d'une flotte, afin de proposer par la suite une stratégie d'optimisation de ces estimations du SOH des batteries.

4.4.2 Résultats sur une flotte

4.4.2.1 Estimation du SOH

Les méthodologies d'estimation du SOH sont appliquées à la flotte de 10 VE détaillée en section 4.2.2.2. Cette flotte comporte 10 VE, chacun utilisé sur une période globale de 10 semaines.

Nous faisons dans un premier temps le choix d'appliquer un modèle d'estimation du SOH hors ligne sur les 10 semaines d'Usage, avec une fenêtre glissante f_g de 1 semaine. Ce procédé permet l'obtention d'estimations régulières à la fin de chaque semaine, pour chaque batterie utilisée. Les résultats de cette application, sur la flotte étudiée, sont illustrés sur la figure 4.19.

Cette figure 4.19 illustre les différences de comportement des estimations du niveau de capacité des batteries lors d'usages réels, par le biais d'un intervalle de confiance $IC = [IC_{inf}, IC_{moy}, IC_{sup}]$ pour chaque batterie.

Les résultats des classifications des Usages des VE, calculés dans la section 4.2.2.2, sont également représentés par les couleurs remplissant les différents intervalles de confiance. Notons que les VE ne sont pas associés à des classes avant la semaine 4, car l'hypothèse choisie considère les 4 dernières semaines d'Usages pour regrouper les VE en différents groupes (section 4.2.1.2).

Il est désormais nécessaire d'introduire un critère de comparaison des résultats obtenus. Pour cela nous définissons l'erreur d'estimation globale $E_{Fl}(s)$ faite sur la flotte Fl , à la semaine s , explicitée par l'équation 4.14. Cette erreur est calculée sur le centre des intervalles de confiance IC_{moy} , avec $t_j^*(s)$ le SOH référence du véhicule j lors de la semaine s .

$$E_{Fl}(s) = \frac{1}{|Fl|} \sum_{j=1}^{|Fl|} \left| 1 - \frac{IC_{moy}^j(s)}{t_j^*(s)} \right| \quad (4.14)$$

Cette erreur globale calculée sur la flotte, accompagnée de sa variance parmi l'ensemble des VE, est présentée en figure 4.20. Nous pouvons observer l'augmentation de cette erreur au fil de la durée d'expérimentation. Néanmoins, l'erreur obtenue après 10 semaines d'Usages d'une flotte de 10 VE correspond à une erreur moyenne des SOH prédits d'environ 3 %. Après 2 mois et demi d'Usage ce résultat implique un bon niveau de confiance vis à vis des estimations de SOH obtenues.

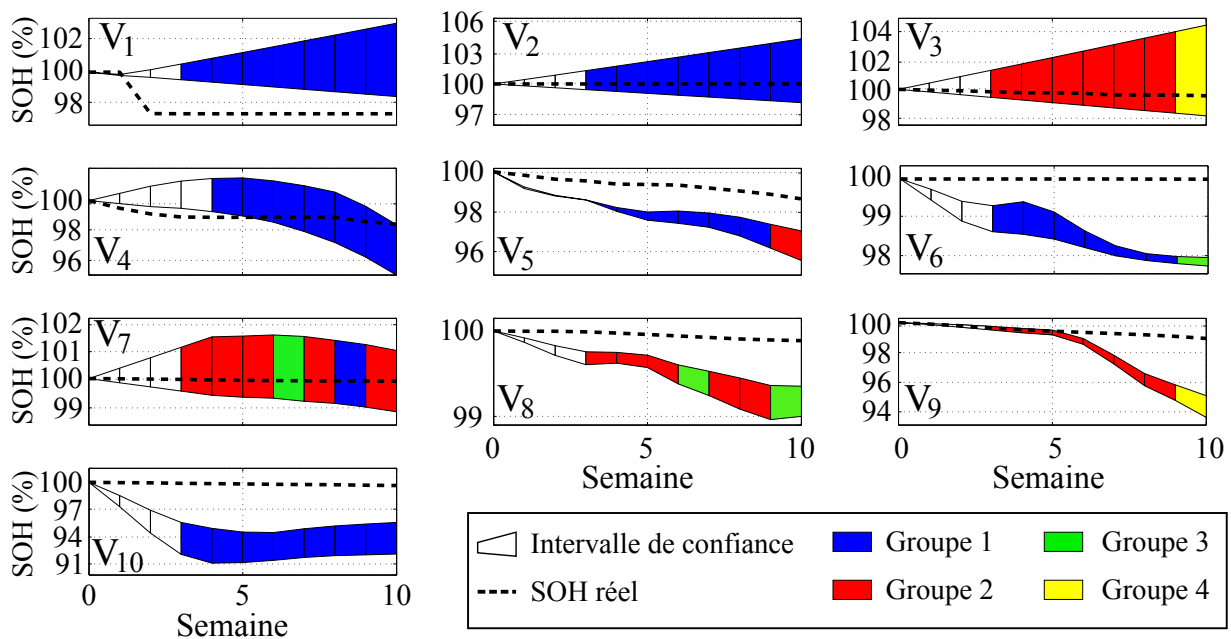


FIGURE 4.19 – Illustration des estimations de capacité de batterie d'une flotte composée de 10 VE, avec la méthodologie basée sur les historiques de données des dernières 4 semaines, pour une fenêtre glissante de 1 semaine - Les groupes d'usages, obtenus par le processus de classification, sont présentés par un code couleur

4.4.2.2 Influence des préconisations d'usages

Les impacts d'éventuelles préconisations d'usages sont quantifiables grâce à l'application du modèle décrit en section précédente. Ce modèle sert de simulateur capable de mesurer l'impact de diverses stratégies sur le vieillissement estimé d'une batterie, suivant la figure 4.21.

Nous pouvons ainsi comparer, par simulation, l'évolution des estimations de vieillissement pour des modes d'Utilisation différents. Par exemple, comparons des Usages références et les mêmes conditions avec des préconisations : une diminution des roulages hebdomadaires de 150 km, et une réduction de la vitesse moyenne des roulages de 10 km.h^{-1} .

La comparaison de ces deux stratégies présente l'incidence des préconisations directement visibles sur les dégradations d'une batterie, sans être particulièrement contraignant pour l'utilisateur. Cet exemple, comparant deux stratégies d'Usages, est illustré sur le véhicule V_6 , avec la représentation en figure 4.22 de la capacité estimée sans modification des Usages comparativement aux estimations considérant l'application des préconisations.

Des préconisations d'Utilisations simples à mettre en place permettent une optimisation de la durée de vie des batteries. Dans l'exemple proposé le SOH de la batterie du véhicule V_6 connaît un ralentissement de son vieillissement de 0.22 % de SOH avec l'application de ces deux préconisations simples sur 10 semaines. Ce résultat permet d'envisager une prolongation de la durée de vie de cette batterie d'environ 4 semaines.

Cet exemple démontre donc l'amélioration de la durée de vie d'une batterie par le

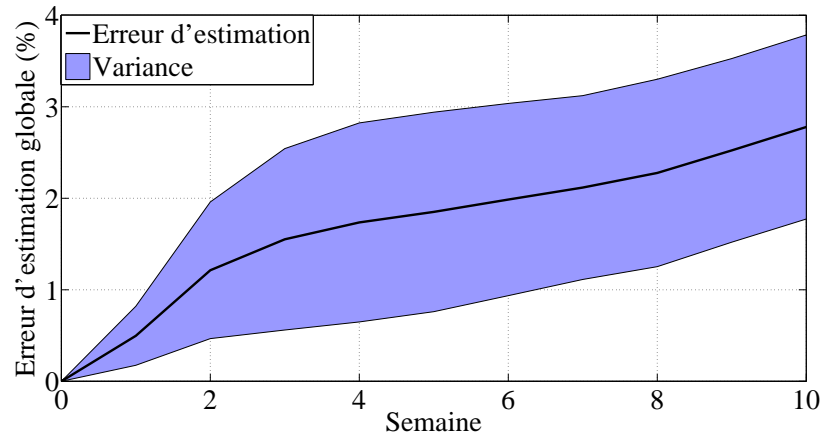


FIGURE 4.20 – Erreur d'estimation moyennes du SOH d'une flotte de VE sur 10 semaines, calculée à partir des estimations faites hors ligne

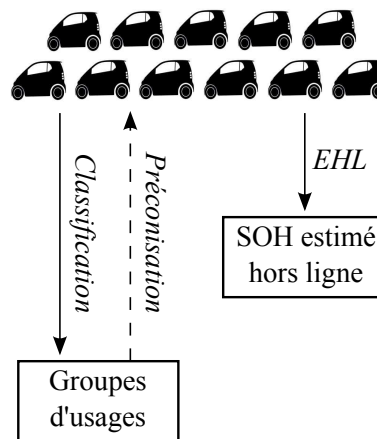


FIGURE 4.21 – Illustration du fonctionnement du simulateur permettant de tester diverses stratégies d'usages

biais de préconisations simples et adaptées à l'Usage du véhicule.

4.4.2.3 Amélioration des estimations

L'objectif est désormais d'analyser les résultats d'estimation de SOH provenant de la méthode de la section 3.3, afin d'améliorer la précision d'estimation sur une flotte de VE (section 4.4.2.1). Le principe étant de combiner les deux types d'estimation, hors ligne (EHL) et en ligne (EEL), pour réduire le taux d'erreur.

Nous considérons ici la méthodologie développée en section 3.3. Les hypothèses portent sur des critères d'extraction correspondant à des accélérations $10 - 60 \text{ km.h}^{-1}$, durant 5 et 6 secondes. Notons que ce procédé nécessite l'utilisation d'un modèle spécifique à chaque modèle de batterie. Nous construisons un modèle à partir des données de la batterie Bat_2 , et un autre basé sur celles de la batterie Bat_3 . Ces deux modèles sont respectivement employés sur les VE de la flotte utilisant ces batteries.

Les résultats de ce procédé sur la flotte étudiée sont détaillés en figure 4.23. Notons

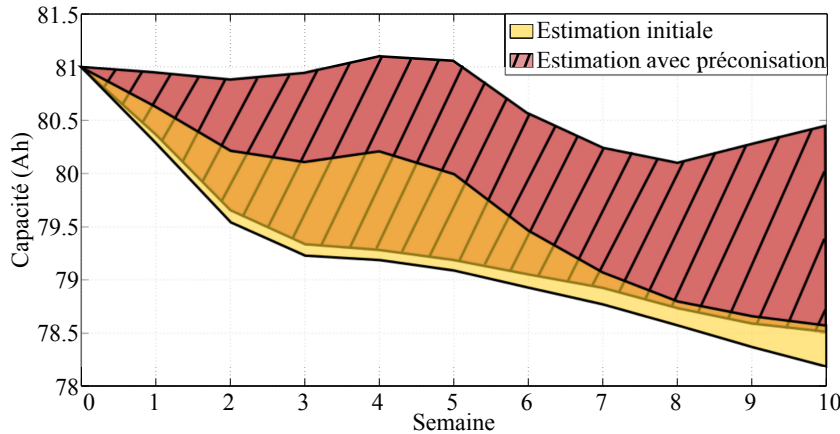


FIGURE 4.22 – Estimations de la capacité de la batterie du véhicule V_6 , sans préconisation et avec une réduction de la vitesse moyenne de 10 km.h^{-1} et une diminution des distances parcourues hebdomadairement de 150 km

le comparatif de ces résultats avec ceux fournis par la méthode hors ligne (figure 4.19).

Cette figure 4.23 met en évidence la complémentarité des deux méthodes d'estimations proposées. Les véhicules V_3 , V_4 et V_5 présentent en particulier des estimations par motifs (EEL) stables sur l'ensemble de la période étudiée, ce qui démontre un bon niveau de confiance. Au contraire, les estimations par motifs (EEL) obtenues avec les véhicules V_6 et V_8 ne proposent pas une grande stabilité d'estimation.

Il est important de remarquer, pour certains véhicules (V_1 , V_2 , V_{10}), l'absence d'estimations obtenues à partir des motifs. Dans ces cas, les véhicules n'ont pas subi d'accélération maximale de 10 à 60 km.h^{-1} lors de la période étudiée.

Une stratégie de combinaison de ces deux approches est alors développée ayant pour but de fiabiliser ces estimations.

Considérons un véhicule V_j de la flotte étudiée. Les estimations de SOH $\{t_1^j, \dots, t_n^j\}$ provenant de l'algorithme d'extraction de motifs d'accélération sont donc obtenues en temps réel lors des usages de ce véhicule. Ces dernières sont prises en compte au terme de chaque période f_g (ici une semaine dans notre exemple). Ces estimations sont utilisées si et seulement si lors de la dernière période considérée f_g , le nombre d'estimations $|t_{f_g}^j|$ est supérieur à un critère Ψ_1 et que leur variance démontre une cohérence des estimations obtenues. Ces deux critères sont explicités dans l'équation 4.15.

$$\begin{cases} |t_{f_g}^j| \geq \Psi_1 \\ \frac{\text{Var}(t_{f_g}^j)}{\text{Mean}(t_{f_g}^j)} \leq \Psi_2 \end{cases}, \quad (4.15)$$

avec Ψ_1 et Ψ_2 des critères définis selon la durée de la période f_g . Ainsi, plus la période f_g est courte, plus le nombre d'estimations est faible. Cependant le SOH est relativement stable sur cette période, ce qui implique une faible variance désirée (i.e. des paramètres Ψ_1 et Ψ_2 peu élevés).

Dans le cas où ces critères 4.15 sont satisfaits à l'instant I , l'intervalle de confiance du SOH estimé du véhicule V_j est uniquement basé sur l'estimation par l'extraction

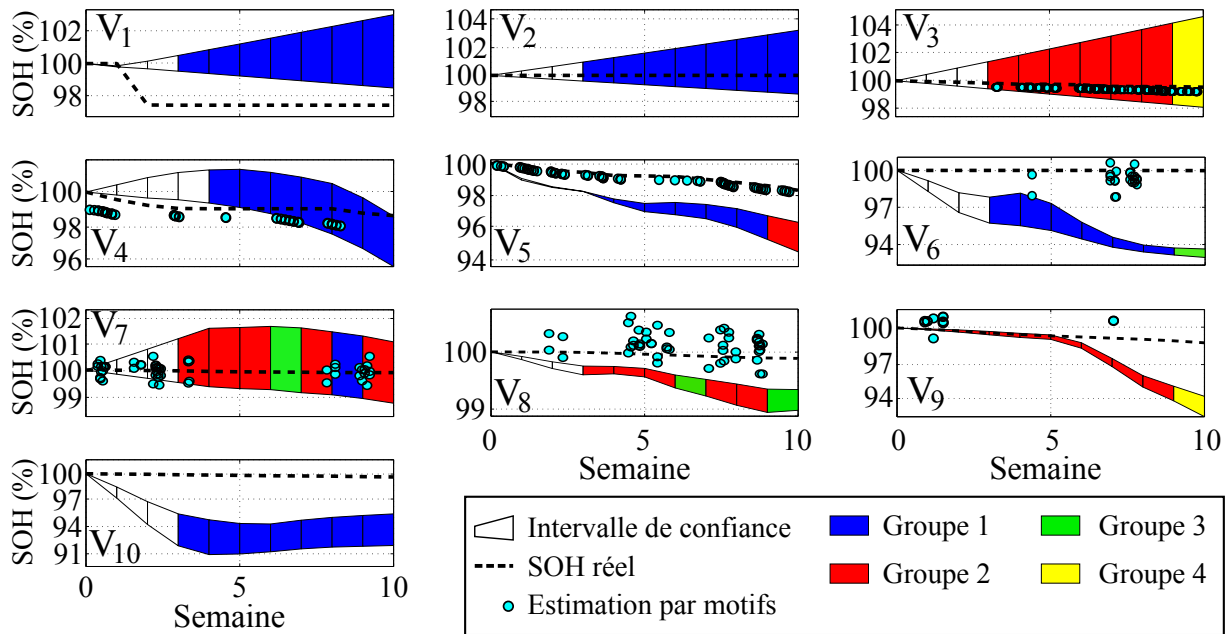


FIGURE 4.23 – Illustration des estimations de capacité de batterie d'une flotte composée de 10 VE, avec la méthodologie présentée en figure 4.19, accompagnée des estimations effectuées en ligne par la méthodologie d'extraction de motifs et des estimations hors ligne basée sur les historiques de données

des motifs (EEL). Formellement, l'estimation du SOH est écrite selon l'équation 4.16, avec les termes -2 et 2 choisis empiriquement pour proposer un niveau de confiance, symétrique, autour des estimations.

$$IC^j(I) = \text{Mean}\left(t_{fg}^j\right) + [-2, 0, 2] \cdot \text{Var}\left(t_{fg}^j\right) \quad (4.16)$$

Dans l'exemple de la flotte Fl étudiée jusqu'ici, $P_c = 1$ semaine. Les paramètres sont définis comme étant : $\Psi_1 = 4$ et $\Psi_2 = 2$ %. La figure 4.24 présente le résultat de l'application de ce procédé. Notons que dans cet exemple l'erreur globale d'estimation 4.14 est de $E_{Fl} = 1.62$ %, contre $E_{Fl} = 3$ % sans la prise en compte des estimations faites à partir des motifs extraits (section 4.4.2.1).

Cette figure 4.24 illustre, sur les véhicules V_3 , V_4 , V_5 , V_6 , V_7 et V_8 , l'impact des estimations de SOH résultant des estimations suite à l'extraction de motifs (EEL). Les trois premiers véhicules ici cités voient donc leurs estimations de SOH considérablement améliorées grâce à la cohérence des estimations faites par l'extraction de motifs. Les estimations de SOH obtenues à partir des motifs extraits sur les véhicules V_6 , V_7 et V_8 n'offrant pas un niveau de confiance très important, les estimations de SOH faites hors ligne (EHL) sont modifiées, ce qui permet la diminution du taux d'erreur.

Nous avons développé une méthodologie de prise en compte de deux types d'estimations différentes. Ce procédé fournit des estimations de SOH précises sur l'ensemble des VE d'une flotte, tout au long des Usages. En effet, parmi les VE disposant d'estimations faites en ligne (EEL), l'erreur maximale d'estimation du SOH observée sur un véhicule, à un instant donné, est de 1.8 %, ce niveau d'erreur n'engendre pas de

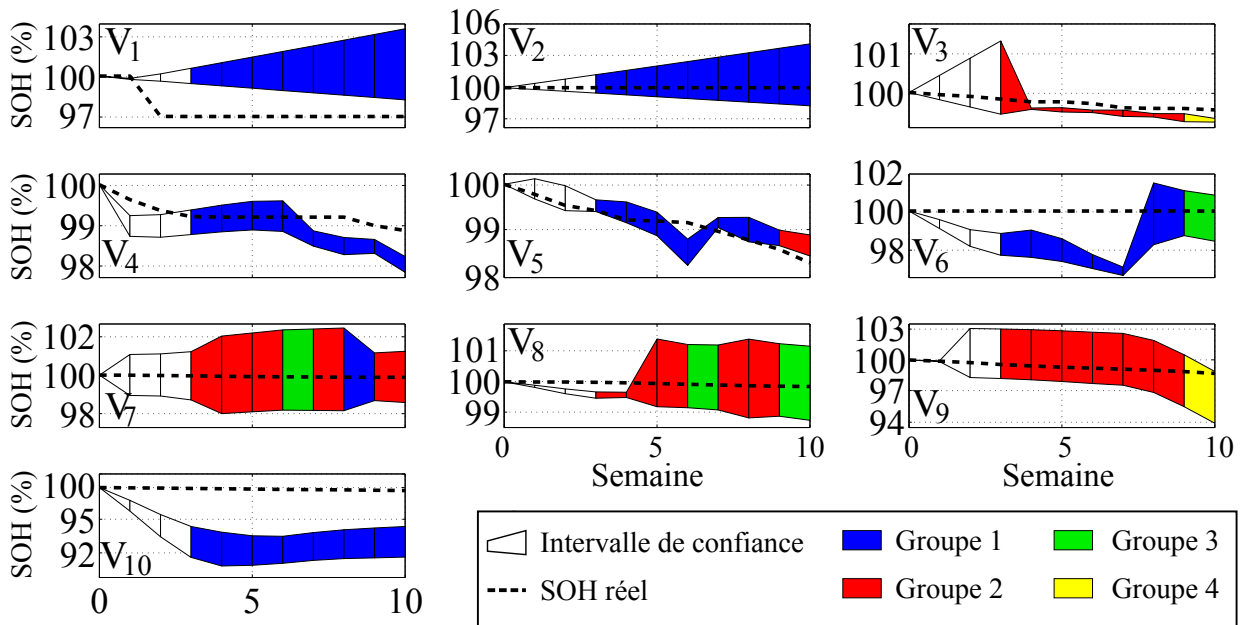


FIGURE 4.24 – Illustration des estimations de SOH pour chacun des 10 VE de la flotte étudiée, en considérant la méthodologie de combinaisons des deux types d'estimations hors ligne et en ligne

contraintes majeures pour un utilisateur de VE.

4.5 Gestion de la santé des batteries d'une flotte

Nous abordons dans cette section les problématiques liées à l'optimisation des précisions d'estimation de SOH d'une flotte de VE en usage réel, en considérant une contrainte liée au nombre de check-ups possibles².

Le contexte de cette section correspond au cas d'un gestionnaire d'une flotte de VE ayant pour objectif de suivre les SOH des batteries utilisées, le plus précisément possible, en imposant une contrainte d'un nombre limité CU_{max} de Check-Up (CU). Ce nombre CU_{max} étant bien sur (très) faible par rapport au nombre de VE composant la flotte, ainsi que par rapport à la durée d'expérimentation totale. Une telle contrainte est par ailleurs guidée par le budget disponible, les check-ups étant très généralement coûteux³.

Cette gestion de l'état de santé des batteries d'une flotte de VE requiert donc deux étapes :

1. L'identification des VE et des instants optimaux, permettant d'obtenir un maximum d'information à partir d'un check-up

2. Ces check-ups, ou caractérisations, servent de référence de SOH pour l'amélioration des estimations.

3. Le prix d'un check-up n'est pas clairement identifié, nous pouvons a-priori l'estimer entre 500 et 2 000 €.

2. La détermination de la prise en compte d'un check-up pour l'ensemble de la flotte

4.5.1 Méthodologie de prise en compte de check-up

Nous détaillons dans cette section le deuxième point évoqué ci-dessus. Nous considérons le résultat d'un check-up effectué à un instant t_{CU} , sur un véhicule V_{CU} . Ce résultat de check-up constitue une information importante devant être utilisée pour corriger les estimations de SOH. La batterie ayant servi au check-up à l'instant t_{CU} voit donc son estimation de SOH attribuée à la valeur exacte référence obtenue. Il est donc important d'établir une méthodologie permettant de répercuter (ou non) un résultat de check-up sur les autres VE de cette flotte.

La méthodologie choisie se base pour cela sur les classifications des Usages de la flotte. Les VE connaissant les mêmes conditions d'Usage sont sujets aux mêmes dégradations. Les informations apportées par la classification des Usages permet d'identifier les VE ayant les mêmes conditions d'exploitation. La connaissance de l'état de santé d'une batterie, apporté par un check-up, offre donc des informations sur les autres véhicules utilisés de la même manière.

Le processus développé considère que les seuls VE pouvant être impactés par un check-up appartiennent au même groupe d'Usage que le véhicule V_{CU} . Cette notion d'appartenance à un même groupe d'Usage nécessite d'être définie de manière précise. L'ensemble des VE impactés par un check-up CU est noté $V_{Imp(CU)}$. L'équation 4.17 formalise ce critère d'appartenance, avec $C(V_j^t)$ la classe associée au véhicule V_j , lors de semaine s , avec δ la fonction de Dirac ⁴.

$$V_{Imp(CU)} = \left\{ V_j, \text{ tel que } \sum_{s=1}^{t_{CU}} \delta_{(C(V_j^s), C(V_{CU}^s))} > 0.7 \cdot t_{CU} \right\} \quad (4.17)$$

L'équation 4.17 revient à utiliser un check-up d'un véhicule V_{CU} , pour modifier uniquement les estimations faites sur les VE possédant une concordance d'historiques de classes d'Usage d'au moins 70 % (ce paramètre étant défini empiriquement). Ce critère est illustré en figure 4.25.

Deux cas possibles se présentent : si le résultat d'un check-up $t_j^*(t_{CU})$ est compris dans l'intervalle de confiance du SOH estimé, ou non. Ceci est simplement déterminé par la condition $\mathcal{B} : \mathbb{1}_{IC(t_{CU})}(t_j^*(t_{CU}))$.

4.5.1.1 Estimations confirmées

Si les estimations de SOH sont confirmées par un check-up (la condition \mathcal{B} est satisfaite), les estimations faites sur les véhicules impactés par ce check-up $V_{Im(CU)}$ sont modifiées. L'information d'un check-up est utilisée pour réduire l'incertitude des niveaux de confiances $IC(t_{CU}) = [IC_{inf}, IC_{moy}, IC_{sup}]$ résultant des modèles d'estimation EHL et EEL.

4. $\delta_{ij} = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{sinon} \end{cases}$.

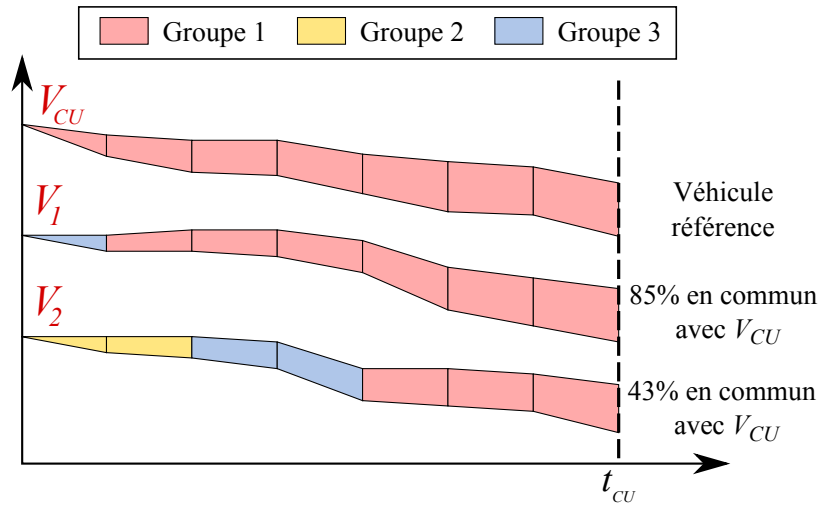


FIGURE 4.25 – Illustration de la sélection des véhicules impactés par un check-up effectué sur le véhicule V_{CU} . Dans cet exemple seul le véhicule V_1 est impacté par le check-up, car ayant un historique de classification commune d'au moins 70 % avec V_{CU} , au contraire de V_2 .

Les estimations $IC^j(t_{CU})$, faites sur un véhicule $V_j \in V_{Im(CU)}$ impacté par le check-up effectué à l'instant t_{CU} , deviennent $IC_{corr}^j(t_{CU})$, suivant l'équation 4.18 :

$$IC_{corr}^j(t_{CU}) = IC^j(t_{CU}) + t_j^*(t_{CU}) - \frac{IC_{sup}^{CU} + IC_{inf}^{CU}}{2} + [R, 0, -R] \cdot (IC_{sup}^j - IC_{inf}^j), \quad (4.18)$$

avec $R = \frac{\epsilon}{IC_{sup}^{CU} - IC_{inf}^{CU}}$, et ϵ un paramètre définissant le taux de restriction de l'incertitude de l'intervalle de confiance.

Une illustration de ce processus est présentée en figure 4.26 dans le cas où l'estimation est vérifiée conforme.

4.5.1.2 Estimations non confirmées

Si la condition \mathcal{B} n'est pas satisfaite, le check-up démontre les erreurs d'estimations portées sur le véhicule V_{CU} étudié. C'est à dire que les incertitudes des estimations de SOH faites sur les batteries des véhicules $V_{Im(CU)}$ ne peuvent être réduites car le niveau de confiance du modèle est insuffisant.

Les estimations $IC^j(t_{CU})$ faites sur un véhicule $V_j \in V_{Im(CU)}$ en un instant t_{CU} , sont modifiées pour devenir $IC_{corr}^j(t_{CU})$, d'après l'équation 4.19.

$$IC_{corr}^j(t_{CU}) = 1 + \left(\frac{2 \cdot t_j^*(t_{CU})}{IC_{sup}^{CU} + IC_{inf}^{CU}} - 1 \right) \cdot \frac{IC^j(t_{CU})}{\rho}, \quad (4.19)$$

avec un paramètre ρ permettant de déterminer le niveau de décalage de l'intervalle de confiance, en fonction du check-up.

La figure 4.27 illustre l'utilisation de cette méthodologie.

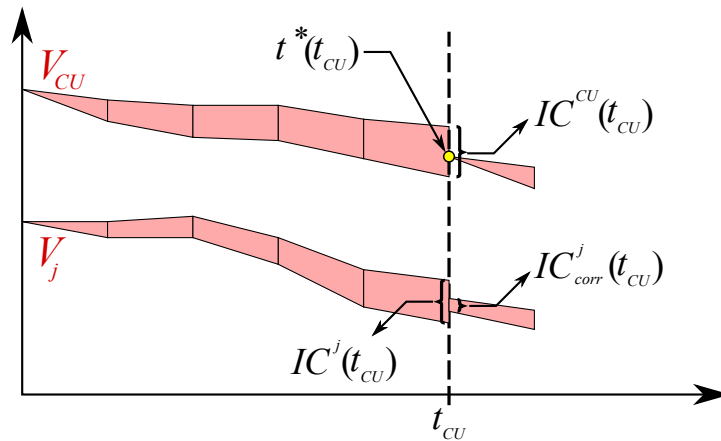


FIGURE 4.26 – Stratégie de mise à jour des estimations $IC^j(t_{CU})$ faites sur le véhicule $V_j \in V_{Imp(CU)}$, à partir du résultat d'un check-up $t^*(t_{CU})$ effectué sur le véhicule V_{CU} - Dans le cas d'un check-up confirmant les estimations

4.5.2 Stratégies d'optimisation

Rappelons que nous considérons ici un nombre maximal de check-ups CU_{max} disponibles sur l'ensemble de la période d'expérimentation. Ce critère est fixé par des contraintes fixées par un gestionnaire d'une flotte, en fonction du budget et du temps d'immobilisation de batterie accordés. Chacun de ces check-ups impacte les estimations de SOH d'autres VE de la flotte Fl , selon les méthodes décrites en section 4.5.1.

L'objectif est à présent d'établir des stratégies permettant d'identifier les instants optimaux pour procéder à un check-up, ainsi que les VE de la flotte sujets de ces check-ups. Les différentes stratégies possibles se divisent donc en deux étapes :

- Choix de l'instant choisi pour procéder à un check-up
- Choix du véhicule dont la batterie est utilisée lors du check-up

Ces choix sont d'une importance primordiale dans l'impact de la réduction de l'erreur d'estimation E_{Fl} du SOH d'une flotte complète, cette erreur étant définie par l'équation 4.14. Il apparaît intuitivement que la décision de procéder à un check-up sur un VE engendre des résultats différents sur l'erreur globale E_{Fl} selon les choix. Nous présentons, dans les sections 4.5.2.1 et 4.5.2.2, différentes stratégies de décisions possibles pour chacun des deux choix évoqués.

4.5.2.1 Choix de l'instant d'un check-up

Notons ici $\mathcal{T}_{CU} = \{t_{CU_1}, \dots, t_{CU_{max}}\}$ les instants correspondant aux check-ups, et \mathcal{S} la durée totale en semaines de l'expérimentation. Détaillons ici 6 stratégies possibles pour le choix des instants \mathcal{T}_{CU} , les plus adaptés pour procéder aux CU_{max} check-ups.

Stratégie $S_1 = 1$: Ce choix correspond à une décision effectuée au début de l'expérimentation par le choix aléatoire de CU_{max} instants parmi les instants possibles. Ce procédé requiert donc la connaissance a-priori de la durée totale de l'expérimentation, ce qui n'est pas le cas dans la majorité des applications considérées.

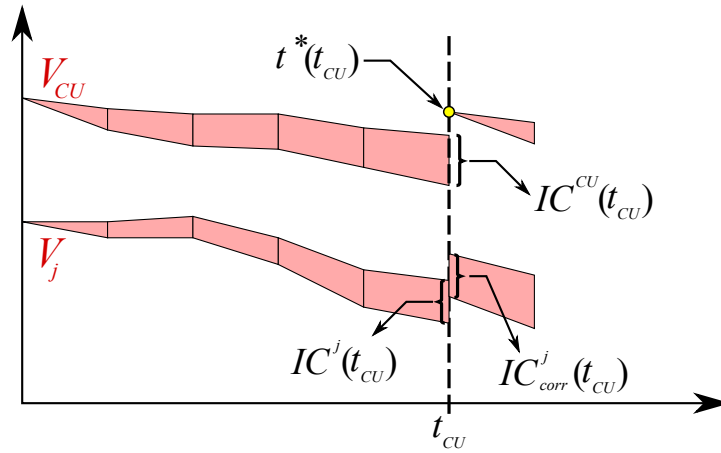


FIGURE 4.27 – Stratégie de mise à jour des estimations $IC^j(t_{CU})$ faites sur le véhicule $V_j \in V_{Imp(CU)}$, à partir du résultat d'un check-up $t^*(t_{CU})$ effectué sur le véhicule V_{CU} - Dans le cas d'un check-up infirmant les estimations

Stratégie $S_1 = 2$: Nous supposons ici également la connaissance a-priori de la valeur S . Le choix consiste à répartir par intervalles réguliers dans le temps l'ensemble des check-ups disponibles. Ce procédé est illustré en figure 4.28.

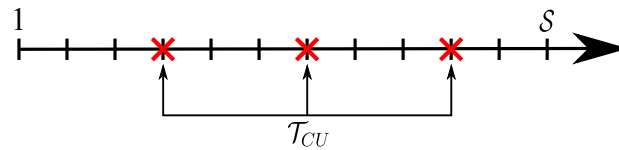


FIGURE 4.28 – Illustration de la stratégie $S_1 = 2$ pour le choix des instants de check-ups

Stratégie $S_1 = 3$: Sans la connaissance de la durée d'expérimentation S , nous procédons ici à un tirage aléatoire d'une valeur θ telle que $0 \leq \theta \leq 1$, à chaque nouvelle semaine. Ceci permet d'appliquer le procédé du choix aléatoire de la stratégie $S_1 = 1$ en l'absence de la valeur S .

Un check-up est donc effectué lors d'une semaine, si la condition $\mathbb{1}_{[0, \lambda_{S3}]}(\theta)$ est satisfaite, avec un paramètre λ_{S3} à fixer, dépendant du nombre de check-ups disponibles.

Stratégie $S_1 = 4$: Cette stratégie se base sur les incertitudes d'estimations cumulées par classe d'usage C_k . Lorsque la somme de celles-ci dépasse un seuil λ_{S4} , pour au moins une classe, un check-up est nécessaire. Ce critère est explicité par l'équation 4.20.

$$\mathcal{T}_{CU} = \left\{ t \mid \max_{C_k} \left(\sum_{j \in C_k} (IC_{sup}^j(t) - IC_{inf}^j(t)) \right) > \lambda_{S4} \right\} \quad (4.20)$$

Stratégie $S_1 = 5$: Avec cette stratégie, nous procédons à un check-up lorsqu'au moins un VE de la flotte possède une incertitude d'estimation supérieure au seuil λ_{S5} . Ceci revient à effectuer l'équation 4.21.

$$\mathcal{T}_{CU} = \left\{ t \mid \max_{j \in FI} \left(IC_{sup}^j(t) - IC_{inf}^j(t) \right) > \lambda_{S5} \right\} \quad (4.21)$$

Stratégie $S_1 = 6$: Selon le même principe que la stratégie $S_1 = 4$, nous calculons la somme des incertitudes de tous les VE de la flotte. Lorsque cette valeur franchit un seuil λ_{S6} un check-up est effectué. L'équation 4.22 formalise ce principe.

$$\mathcal{T}_{CU} = \left\{ t \mid \sum_{j=1}^{|FI|} \left(IC_{sup}^j(t) - IC_{inf}^j(t) \right) > \lambda_{S6} \right\} \quad (4.22)$$

4.5.2.2 Véhicule choisi pour un check-up

Considérant les instants \mathcal{T}_{CU} sélectionnés pour l'application d'un check-up, il reste à choisir un VE, noté \mathcal{V}_{CU} pour chacun de ces instants. Pour ce faire il existe de nombreuses stratégies possibles. La batterie choisie à cet instant sera sujette à un check-up. Nous énumérons ci-dessous quelques unes des stratégies permettant de définir un choix de VE parmi la flotte étudiée. Pour développer ces stratégies, nous notons $t_{CU} \in \mathcal{T}_{CU}$ l'instant lors duquel un VE doit être choisi pour effectuer un check-up.

Stratégie $S_2 = 1$: Cette stratégie consiste simplement à prendre un VE aléatoirement parmi la flotte dans sa globalité, à l'instant t_{CU} .

Stratégie $S_2 = 2$: À l'instant t_{CU} , le choix du véhicule se porte sur le véhicule ayant la plus grande incertitude (équation 4.23), au sein de la classe C_p proposant la plus grande incertitude cumulée (équation 4.24).

$$\mathcal{V}_{CU} = \operatorname{argmax}_{VE \in C_p} \sum_{j=1}^{|C_p|} \left(IC_{sup}^j(t_{CU}) - IC_{inf}^j(t_{CU}) \right), \quad (4.23)$$

$$\text{avec } Cl_p = \operatorname{argmax}_{C_r \in C} \sum_{j=1}^{|C_r|} \left(IC_{sup}^j(t_{CU}) - IC_{inf}^j(t_{CU}) \right) \quad (4.24)$$

Stratégie $S_2 = 3$: Cette stratégie correspond à un choix de véhicule parmi les VE de la classe C_r issue de l'équation 4.24.

Stratégie $S_2 = 4$: Sélectionnant la classe C_r avec l'incertitude cumulée la plus importante, par l'équation 4.24, le véhicule choisi est celui impactant le plus grand nombre

d'autres VE de la flotte (noté $Imp(CU)$), défini par la condition \mathcal{B} en section 4.5.1). Ceci est formalisé par l'équation 4.25.

$$\mathcal{V}_{CU} = \operatorname{argmax}_{VE \in Cl_r} |Imp(CU)| \quad (4.25)$$

Stratégie $S_2 = 5$: Ce procédé sélectionne le véhicule étant associé à la plus grande incertitude d'estimation parmi toute la flotte Fl (équation 4.26).

$$\mathcal{V}_{CU} = \operatorname{argmax}_{VE \in Fl} \sum_{j=1}^{|Fl|} \left(IC_{sup}^j(t_{CU}) - IC_{inf}^j(t_{CU}) \right) \quad (4.26)$$

Stratégie $S_2 = 6$: Cette stratégie permet de choisir un VE au sein de la flotte entière, tel qu'un plus grand nombre possible d'autres VE soient impactés par ce check-up (équation 4.27).

$$\mathcal{V}_{CU} = \operatorname{argmax}_{VE \in Fl} |Imp(CU)| \quad (4.27)$$

Stratégie $S_2 = 7$: Nous définissons cette stratégie comme sélectionnant, parmi la flotte, le VE générant une modification de la plus grande incertitude cumulée des véhicules impactés $Imp(CU)$ par ce check-up (équation 4.28).

$$\mathcal{V}_{CU} = \operatorname{argmax}_{VE \in Fl} \left(\sum_{j=1}^{|Imp(CU)|} \left(IC_{sup}^j(t_{CU}) - IC_{inf}^j(t_{CU}) \right) \right) \quad (4.28)$$

4.5.2.3 Combinaison des stratégies

Nous avons donc ici défini un grand nombre de stratégies possibles. La table 4.5 présente un résumé avec des notations simplifiées des propositions évoquées.

Toutes les associations de stratégies S_1 et S_2 sont compatibles les unes avec les autres, même si certains choix semblent plus cohérents ensembles. Par exemple, la stratégie $S_2 = 5$ semble plus adaptée avec $S_1 = 5$. Il est possible de générer 42 stratégies globale $S = (S_1, S_2)$ différentes, à partir de ces propositions.

Notons que les modifications survenant lors d'un check-up ne modifient aucunement les méthodologies de classification des usages ou bien d'estimation du SOH en temps réel. L'objectif de cette approche est d'identifier la stratégie optimale permettant de réduire les erreurs d'estimation au maximum lors de l'étude d'une flotte de VE. L'ensemble des stratégies décrites sont comparées par la suite.

4.5.3 Exemple d'application

Afin d'illustrer le fonctionnement et les performances obtenues par ce procédé de prise en compte de check-ups, nous prenons de nouveau l'exemple d'une flotte de 10 VE présentée en section 4.2.2.2.

Il est en premier lieu nécessaire de définir les nombreux paramètres évoqués en sections 4.5.1 et 4.5.2, et intervenant dans l'application de cette méthodologie. Les choix

Différentes stratégies			
S_1	Instant Descriptif	S_2	VE Descriptif
1	Aléatoire	1	Aléatoire
2	Intervalle régulier	2	$\begin{cases} \operatorname{argmax}_C(IC) \\ \operatorname{argmax}_{VE}(IC) \end{cases}$
3	Aléatoire continu	3	$\begin{cases} \operatorname{argmax}_C(IC) \\ VE \text{ aléatoire} \end{cases}$
4	$\sum_C(IC) > \lambda_{S4}$	4	$\begin{cases} \operatorname{argmax}_C(IC) \\ \operatorname{argmax}_{VE}(IC(Imp(CU))) \end{cases}$
5	$\max_{VE}(IC) > \lambda_{S5}$	5	$\operatorname{argmax}_{VE}(IC)$
6	$\sum_{VE}(IC) > \lambda_{S6}$	6	$\operatorname{argmax}_{VE}(VE(Imp(CU)))$
		7	$\operatorname{argmax}_{VE}(IC(Imp(CU)))$

TABLE 4.5 – Synthèse des différentes stratégies proposées pour la gestion de répartition des check-ups, avec des notations simplifiées

effectués dans cet exemple sont énumérés au sein des équations 4.29. Les paramètres sont ici choisis empiriquement.

$$\begin{cases} \epsilon &= \min(|IC_{sup}^{CU} - t_j^*|, |t_j^* - IC_{inf}^{CU}|) \\ \rho &= 20 \\ \lambda_{S3} &= \min(0.8, CU_{max}/16) \\ \lambda_{S4} &= 6 \cdot CU_{max}^{-1} \\ \lambda_{S5} &= 14 \cdot CU_{max}^{-1} \\ \lambda_{S6} &= 6 \cdot CU_{max}^{-1} \end{cases} \quad (4.29)$$

L'application de la méthodologie globale de prise en compte de check-up effectués selon des stratégies prédéterminées améliore les performances d'estimations du SOH des batteries utilisées au sein de la flotte. La figure 4.29 présente l'évolution des erreurs calculées selon l'équation 4.14, en fonction de la stratégie employée et du nombre de check-ups considérés. Seules quelques stratégies sont ici illustrées, les autres ne présentant pas de résultats significativement différents de ceux-ci.

La figure 4.29 démontre la pertinence de la méthodologie développée à partir d'un check-up. En effet, rares sont les stratégies ne permettant pas d'améliorer les estimations en procédant à des check-ups, en particulier lorsque le nombre de check-ups augmente.

La stratégie $S = (2, 5)$ fournit par exemple des erreurs similaires quelque soit le nombre de check-ups effectués. Cette stratégie $S = (2, 5)$ consistant à une sélection de l'instant d'un check-up avec un intervalle régulier dans le temps, et la sélection du

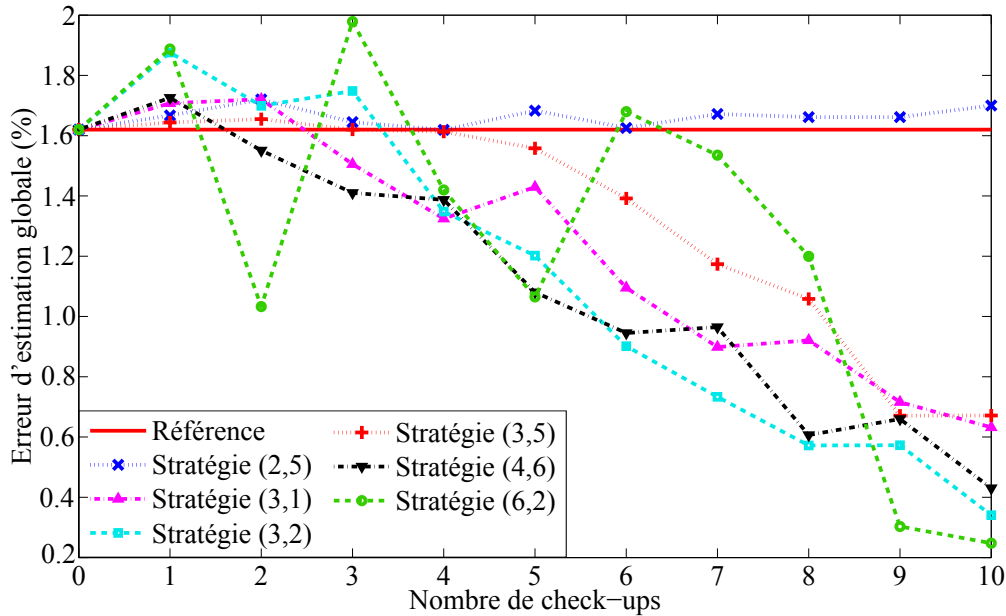


FIGURE 4.29 – Évolution des erreurs d'estimations de SOH d'une flotte, en fonction du nombre de check-ups et de la stratégie choisie

véhicule comme étant celui impactant la plus grande incertitude. Ces choix n'étant pas optimaux, il n'apparaît pas surprenant que les résultats obtenus par cette stratégie $S = (2, 5)$ soient peu précis.

Certaines stratégies, telles que la stratégie $S = (6, 2)$, présentent des dynamiques peu monotones et offrent des performances très sensibles au nombre de check-ups considéré. Ce type de résultat est expliqué par la définition empirique des paramètres intervenant dans la définition des stratégies, λ_{S6} pour la stratégie $S = (6, 2)$. Il apparaît nécessaire d'étudier une définition plus précise de ces paramètres pour rendre plus robustes de telles stratégies.

Les stratégies associées aux meilleurs résultats permettent de réduire l'erreur de référence obtenue sans check-up de 1.61 % vers une erreur d'estimation < 0.4 % en procédant à 10 check-ups. Notons que l'obtention d'une erreur d'estimation nulle sur 10 semaines nécessiterait l'application d'un check-up sur chacun des VE, à chaque semaine, c'est à dire à 100 check-ups.

Cette figure 4.29 suggère la sélection des stratégies les plus stables pour garantir une amélioration des résultats quelque soit le nombre de check-ups effectués. Ainsi, les trois meilleures stratégies semblent être $S = (3, 1)$, $S = (4, 6)$ et $S = (3, 2)$. Notons que ces stratégies ne sont pas basées sur la notion d'impact d'un check-up ($S_2 = 7$) ou sur le niveau de confiance à l'instant du check-up ($S_1 = 5$), du fait du manque de robustesse des paramètres λ choisis.

De tels résultats démontrent à la fois la performance de la stratégie d'adaptation des estimations hors ligne et en ligne aux problématiques des flottes, mais également la compatibilité de toutes les méthodologies développées.

Un exemple des effets des check-ups sur les estimations est détaillé dans la figure 4.30. Cet exemple est basé sur la répartition de 5 check-ups par le biais de la stratégie

$S = (4, 6)$.

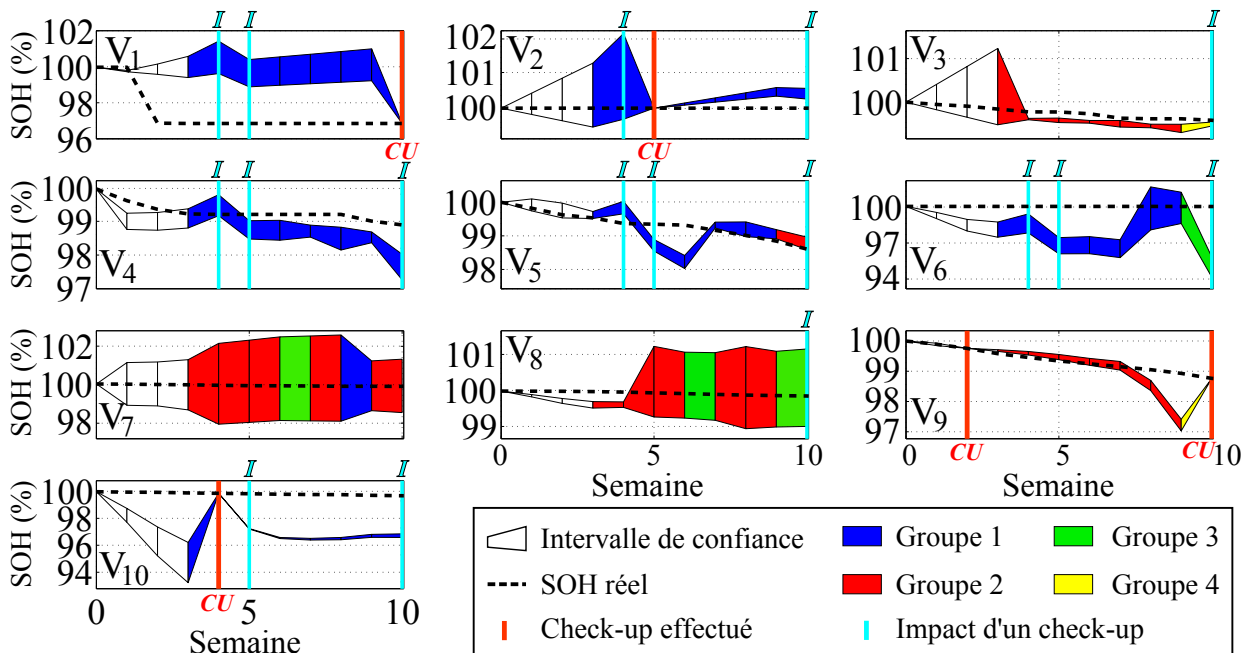


FIGURE 4.30 – Exemple de l'évolution des estimations de SOH d'une flotte avec 5 check-ups appliqués selon la stratégie $S = (4, 6)$

Cette figure 4.30 illustre les effets correcteurs des check-ups sur le véhicule choisi, mais également sur les autres véhicules de la flotte possédant une utilisation similaire. Notons que cet exemple fournit une erreur globale d'estimation des SOH de 1.06 %.

4.5.4 Perspectives

Dans cette section 4.5 nous avons développé un procédé de prise en compte de l'information obtenue par le biais de check-up. Ceux-ci permettent la fiabilisation des estimations de SOH d'une partie de la flotte, et non seulement du véhicule ayant été utilisé lors de ce check-up.

D'autre part, des stratégies de répartition de check-ups ont été détaillées. Le choix et la définition de ces stratégies sont effectués au préalable. Les résultats obtenus par ce procédé démontre ainsi la capacité de performance de l'algorithme global sans contrôle lors de l'expérimentation. De plus, le taux d'erreur optimal atteint lors de l'étude d'une flotte de 10 VE sur 10 semaines, pour 5 check-ups au maximum, est d'environ 1 %. Un tel niveau de précision démontre la fiabilité de la méthodologie globale et permet d'envisager des tests réels sur une longue période d'expérimentation.

Il est également important de noter que les stratégies de définition des check-ups sont ici basées sur des réflexions pratiques. Or, il peut être envisagé de développer un procédé d'optimisation proposant des erreurs d'estimations encore plus faibles. De la même façon, les paramètres utilisés dans cette étude sont définis de manière arbitraire par les équations 4.29. Il peut donc être possible de les rendre plus performants et

par conséquent d'obtenir un meilleur niveau de précision, par une étude de sensibilité plus approfondie sur leurs comportements.

4.6 Conclusion - Discussion

Dans ce chapitre 4 une approche des problématiques liées au contexte d'une flotte de VE a été étudiée. Le suivi d'une flotte de VE requiert de nombreux développements optimisant la gestion des plusieurs véhicules. Pour ce faire, différents outils complémentaires ont été développés.

Une approche permettant de détecter les différents types d'Usage au sein d'une flotte de VE, simplement fondée sur des variables simplement mesurables, a été explicitée (section 4.2). Les résultats de cette méthodologie démontrent la fiabilité des interprétations générées.

L'extension de ce travail d'identification des utilisations des VE se prolonge par la mise en place de préconisations d'Usage, adaptées aux comportements de chacun des VE, réduisant les impacts du vieillissement des batteries (section 4.3).

Ces outils de suivi des Usages d'une flotte de VE permettent donc de mettre en place une gestion optimale, ayant pour objectif de réduire les impacts des conditions de vies des différentes batteries, afin d'optimiser leur durée de vie.

Les estimations de SOH des batteries évoquées dans les chapitres 2 et 3 sont ici appliquées au cas d'une flotte de VE (section 4.4). Ces deux méthodes d'estimations démontrent une parfaite complémentarité, réduisant les erreurs d'estimations en combinant ces deux types d'estimations (section 4.4.2.3). L'application d'une telle stratégie d'estimation globale démontre la pertinence des deux modèles d'estimations et leur complémentarité dans le cadre d'une amélioration des résultats.

Dans la section 4.5 une approche basée sur la possibilité de répartition de check-ups est abordée. Cette méthodologie met en évidence l'importance dans le choix de stratégie de préconisation de check-ups. L'influence de ceux-ci est exploitée afin d'améliorer les estimations de SOH faites sur les batteries des VE n'ayant pas été sujets à ce check-up. Ce procédé utilise la classification des Usages des VE, ainsi que les deux types d'estimations de SOH (EHL et EEL). Cette méthodologie globale, utilisant l'ensemble des outils développées dans cette étude, engendre une réduction significative des erreurs d'estimations (section 4.5.3).

Les approches proposées dans ce chapitre 4 mettent en avant la nécessité du développement de solutions adaptées aux problématiques spécifiques des flottes de VE. En effet, la prolongation des durées de vies des batteries et l'optimisation des usages des VE sont à la base des coûts liés à la gestion d'une flotte.

Notons que les méthodologies explicitées sont toutes complémentaires et les résultats obtenus témoignent du bon niveau de confiance associé. Néanmoins, de nombreux paramètres et validations sont nécessaires dans la mise en place de ce procédé. Les paramètres ayant été dans ce chapitre fixés de manière arbitraire, il est clairement envisageable d'améliorer ces performances par le biais d'une étude plus approfondie fournissant l'estimation des paramètres optimaux. Par exemple, la synthèse automatique des variables importantes pour interpréter les usages peut être perfectionnée par l'ajustement de règles de gestion.

Les stratégies de préconisations des check-ups disponibles évoquées démontrent l'importance de cette approche. Cependant, les résultats de cette étude ne mettent pas

en évidence une performance prédominante d'une unique stratégie (section 4.5.3). Ces résultats impliquent la nécessité d'une approche d'optimisation globale permettant d'identifier une stratégie optimale, associée dans tous les cas possibles, aux meilleures estimations de SOH.

Les méthodes explicitées dans ce chapitre 4 offrent un outil de suivi et de maintenance des véhicules d'une flotte, par le biais d'une centralisation de leurs données. Ce suivi de la maintenance et de l'état de santé des véhicules d'une flotte fournit une notion de confiance aux utilisateurs importante pour réduire le *range anxiety* et les interrogations liées aux indicateurs et aux informations qui leurs sont proposés. La fiabilisation des estimations permettant ainsi la diminution de leur anxiété vis-à-vis des véhicules électriques.

Conclusion générale

Les ambitions actuelles de développement du marché des VE reposent en grande partie sur les performances des batteries utilisées. Ces batteries constituent le centre des attentes, et de nombreuses recherches tendent à les rendre plus performantes. La situation idéale ambitionne de réduire les coûts de fabrication et d'augmenter les durées de vie de ces batteries. De telles caractéristiques seraient en effet un atout majeur pour l'acceptation des VE par le plus grand nombre.

Dans l'attente des améliorations technologiques des batteries employées par les VE, il apparaît primordial d'optimiser leur durée de vie. Une telle optimisation influence la rentabilité d'un VE par rapport à son prix d'achat. De ce fait, les problématiques associées aux phénomènes de vieillissement d'une batterie en usage sur un VE deviennent des enjeux de première importance.

La compréhension des procédés de dégradations électrochimiques d'une batterie constitue un élément important dans la mise en place des méthodologies dédiées, en complément des approches physiques et électrochimiques. Ces méthodologies sont ici entièrement basées sur l'étude des données mesurées et enregistrées lors de cyclage ou de stockage de batteries embarquées sur des VE. L'apport de ce choix permet d'identifier des informations non obtenues par les diverses et communes techniques de modélisation et d'estimation du vieillissement actuelles. En effet, les bases de données importantes stockées contiennent des caractéristiques relatives aux vieillissement des batteries. La problématique est donc d'établir les bonnes approches capables d'exploiter ces données pour extraire des informations explicitant les facteurs de vieillissement.

A. Bilan des travaux

Identification des variables impactant le vieillissement

La complexité de l'identification des causes de phénomènes de vieillissement est expliquée par la diversité des variables impliquées dans ces processus. Durant la vie d'une batterie (cyclages ou stockages) des dégradations irréversibles sont provoquées selon des proportions différentes.

Les travaux développés dans cette étude ont permis, par l'analyse approfondie des données enregistrées, l'identification et la quantification de l'influence des variables dans chacun des phénomènes globaux de vieillissement se produisant sur une batterie :

- La réduction irréversible de la capacité est liée aux conditions de stockage de la batterie (SOC, température) ainsi qu'aux températures vécues lors des Usages.

- L'augmentation de la résistance globale est associée à des variables issues des conditions d'Usage de la batterie (distance, nombre de cycles, température ambiante) ou bien du comportement du conducteur (PKE, profil de courant).

L'analyse de ces données collectées a mis en évidence les conditions d'Usage contraignantes : les températures extrêmes, les roulages agressifs... Ces résultats sont en accord avec les analyses issues de l'étude du fonctionnement électrochimique des batteries, ce qui renforce la confiance obtenue par cette méthode.

Aide à l'interprétation des usages responsables du vieillissement

Les analyses produisant l'identification des principales variables responsables du vieillissement des batteries sont également employées pour l'interprétation des modes d'Utilisation des VE.

Une méthodologie de préconisation des Usages est développée sur la base de l'étude des conditions de vie des batteries. La mise en évidence des conditions d'expérimentation contraignantes, et les préconisations engendrées, pour un VE fournissent un outil de gestion et d'optimisation d'une batterie. Un utilisateur peut en effet exploiter les résultats de cette méthodologie pour restreindre les contraintes appliquées à la batterie étudiée.

L'étude approfondie des historiques de données offre ici une interprétation des causes de vieillissement d'une batterie en usage sur un VE et permet l'interprétation a-posteriori des conditions de vie d'une batterie.

Estimation de l'état de santé d'une batterie

L'exploitation de ces historiques de données est également faite dans cette étude dans le cadre d'une production d'estimation du niveau de performance d'une batterie. Notons le besoin de constituer une base de données exhaustive comprenant un maximum d'informations représentatives pour construire des modèles pertinents. Des modèles linéaires sont construits pour l'estimation des deux principaux facteurs représentatifs des performances d'une batterie, à savoir l'évolution de sa capacité et de sa résistance.

Les valeurs représentatives, et mesurables sans procédé complexe, des conditions d'Usage d'une batterie sont exploitées et offrent une estimation implémentable du niveau de capacité et de résistance d'une batterie. Les résultats démontrent la précision obtenue dans les estimations de capacité, à partir de quelques valeurs significatives des conditions vécues par la batterie durant une période donnée.

La méthodologie d'estimation des performances d'une batterie démontre la possibilité de parvenir à l'extraction d'informations relatives au vieillissement d'une batterie, uniquement à partir de l'analyse de données stockées. L'approche proposée permet le suivi régulier, hebdomadaire par exemple, des performances d'une batterie, pour un faible coût d'obtention des variables d'Usage.

Ces estimations, faites hors ligne (EHL), ne proposent cependant pas de solution au suivi fréquent et détaillé des performances d'une batterie lors de ces Usages. L'état de santé d'une batterie étant primordial dans les utilisations quotidiennes d'un VE, il devient nécessaire d'établir une méthodologie alternative fonctionnant en embarqué.

Une approche d'extraction de motifs spécifiques en ligne est développée (EEL), exploitant les modifications de comportement d'une batterie détectables au sein des signaux mesurés. Ces signaux réels offrent une plus grande précision d'information par rapport à des approches globales d'Utilisation. En effet, pour une même demande, les signaux de courant et de tension présentent des comportements différents en fonction de l'état de santé de la batterie. Le développement d'une telle méthodologie basée sur ces motifs spécifiques d'une même sollicitation, par exemple une accélération de 10 à 60 km.h⁻¹, propose une estimation de SOH lors de chacune de ces accélérations.

Les résultats de cet algorithme démontrent une erreur d'estimation très faible (< 0.5 %), permettant l'implémentation de cette approche dans un cadre embarqué. Les avantages de cette approche résident en une utilisation possible en embarqué dans un VE, une obtention des estimations sans nécessiter d'immobilisation du véhicule et le faible besoin de mesures (courant, tension et potentiellement température et SOC).

Introduction d'un outil adapté au suivi des états de santé des batteries d'une flotte

L'extension du marché des VE comprend notamment le développement de flottes de véhicules entièrement composées de VE. Cette situation engendre la création de nouvelles problématiques de gestion et d'optimisation de la durée de vie des batteries utilisées, liées aux performances des batteries, et donc à leur vieillissement. Les gestionnaires de flotte de véhicules font face à des questions de coût d'utilisation, ce qui dans le cas des VE est directement lié au vieillissement des batteries.

La recherche de l'optimisation des coûts de fonctionnement d'une flotte est synonyme de l'optimisation de la durée de vie des batteries employées. Cette durée de vie étant directement une conséquence des Usages faits de chacune des batteries, la méthodologie d'identification des conditions d'exploitations contraignantes est ici particulièrement nécessaire. De plus, la production de préconisations d'Usages adaptés à chacune des batteries permet une réduction des vieillissements engendrés et donc une augmentation de leur durée de vie.

La comparaison des Usages de chacun des véhicules d'une flotte est effectuée proposant de regrouper les véhicules par classes d'utilisation. Ce regroupement offrant une aide à l'identification des comportements contraignants appliqués aux batteries.

De la même façon que pour les utilisations individuelles, une flotte de VE est confrontée au besoin de connaissance à chaque instant des performances maximales d'une batterie, notamment en terme d'autonomie. Les algorithmes d'estimation du SOH sont ici adaptés au contexte d'une flotte de véhicules, proposant une approche de combinaison des estimations obtenues hors ligne et en embarqué. Ce procédé ayant pour objectif de rendre plus robustes les estimations de SOH.

D'autre part, l'utilisation de possibles caractérisations de batteries est étudiée, proposant des stratégies optimisant leurs impacts sur les estimations faites sur l'ensemble des VE. Cette approche introduit une réflexion sur la prise en compte de valeurs références de SOH de batteries, sur l'ensemble des autres batteries composant la flotte concernée.

Applications possibles des outils développés

Ces travaux ont pour spécificité de produire des résultats basés sur l'exploitation des données, capables d'être analysés par des experts des véhicules et des batteries susceptibles de procéder aux interprétations.

L'ensemble des méthodologies évoquées dans ces travaux ayant été guidées par des besoins issus de l'utilisateur ou du constructeur, il existe de nombreuses applications possibles dans le cadre réel d'usage de VE. Ces outils fournissent des propositions de solutions pour les utilisateurs particuliers de VE ou bien pour les gestionnaires de flottes. Le fonctionnement des méthodologies développées est illustré en figure 31.

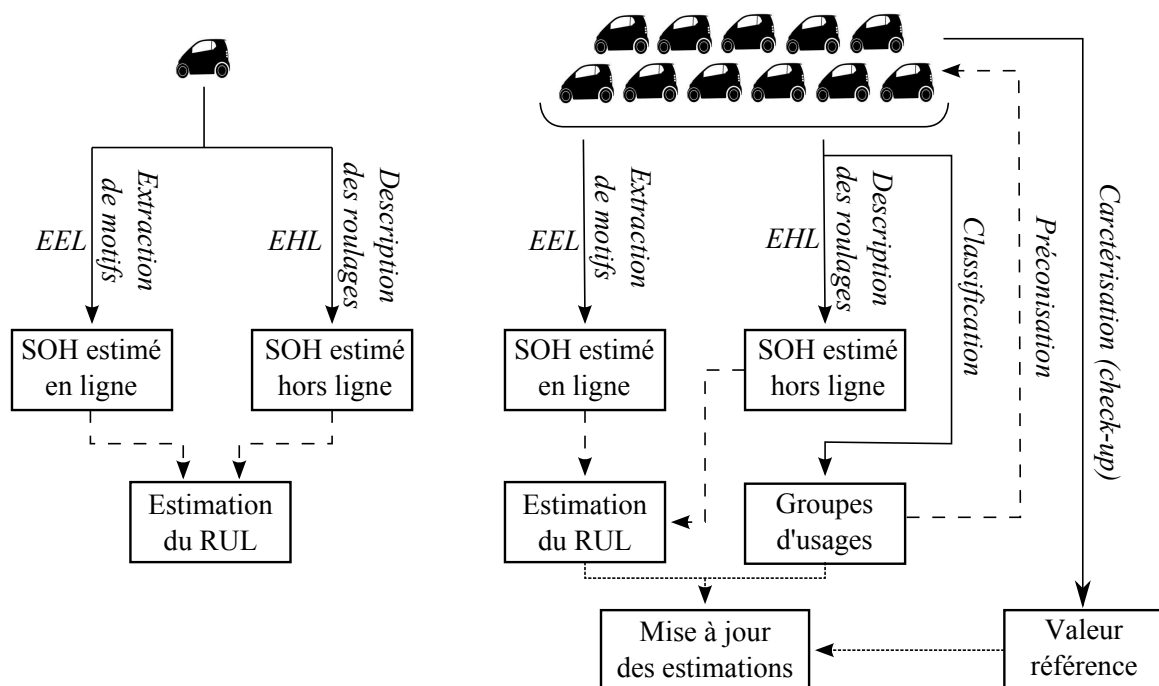


FIGURE 31 – Ensemble des méthodologies développées dans ces travaux de recherche, proposant l'estimation et la gestion de l'état de santé de batteries

Les applications proposées permettent la mise en place d'indicateurs obtenus en traitement embarqué et affichés au sein d'un véhicule. De tels indicateurs pouvant être une estimation du SOH de la batterie utilisée ou bien une préconisation d'Usage.

Une autre partie des travaux est applicable de manière débarquée au véhicule, proposant un suivi d'une flotte de VE, et offrant une stratégie d'optimisation des caractérisations faites sur les batteries. Un gestionnaire de flotte peut, par l'application des méthodologies développées dans cette étude, obtenir un tableau de bord affichant les préconisations d'Usage de chaque véhicule, les préconisations éventuelles de caractérisation d'une batterie, l'estimation du SOH de chacune des batteries employées.

B. Limites et perspectives

Limites des méthodologies évoquées

Les améliorations possibles des méthodologies explicitées dans cette étude ont bien évidemment leurs limites. Les améliorations possibles évoquées ne peuvent en effet résoudre l'ensemble des limitations identifiées.

La première limite provient de l'axe stratégique choisi par ces travaux. Une considération, rapidement faite par cette étude, porte sur l'importance primordiale du SOH sur les performances d'une batterie pour un utilisateur. Cette remarque ayant pour conséquence l'hypothèse d'une évolution de résistance peu impactante sur les performances d'une batterie pour un conducteur. Ceci est justifié par les caractéristiques techniques des batteries actuelles qui offrent des performances de puissance suffisante pour qu'une augmentation de la résistance soit moins sensible pour un utilisateur qu'une perte de capacité. Les technologies de batteries étant en constante évolution il apparaît nécessaire de proposer également une méthodologie d'estimation embarqué du niveau de résistance d'une batterie lors d'usages réels, uniquement à partir des données mesurées sans conditions spécifiques.

Il est également important de noter l'important besoin en terme de diversité de données de l'ensemble des outils développés lors de cette étude. Les principes des méthodologies *data-driven* induisent en effet la dépendance directe entre l'exhaustivité des données d'apprentissage et la précision des résultats obtenus. Le vieillissement des batteries étant un processus relativement lent (~ 10 ans), il apparaît peu réaliste d'obtenir des données réelles provenant de l'ensemble des conditions d'expérimentations possibles. Ceci forme l'argument principal des critiques de ce type d'approche et des études portant sur des tests de vieillissement accéléré.

Considérant la même réflexion que celle menée sur les besoins de données exhaustives, il est nécessaire de procéder à des expérimentations de ces méthodologies sur d'autres chimies de batteries afin de vérifier, ou non, leur adéquation et leur robustesse.

Une autre limite non abordée par ces travaux consiste à mettre à jour les modèles proposés d'estimation de SOH. Dans le cadre d'un usage individuel d'un VE, il n'a pas été envisagé de procéder à l'amélioration des modèles. En imaginant une manipulation rapide, pouvant être faite lors d'une immobilisation prévue d'un véhicule dans un garage, une valeur référence de SOH pourrait être obtenue. La prise en compte de cette information pour l'amélioration des modèles doit donc être considérée par la suite.

Améliorations des méthodologies

De nombreuses pistes d'amélioration directe des outils détaillés sont possibles, pouvant ainsi offrir de meilleurs résultats simplement à partir des données mesurées sur la batterie et sur les conditions extérieures.

La piste la plus directe pour l'amélioration consiste en l'augmentation de la diversité des données utilisées dans cette étude. En effet, les modèles reposant entièrement sur les dynamiques présentes dans les données d'apprentissage, il est très difficile pour ces algorithmes d'être fiables dans des conditions non observées dans l'ensemble d'apprentissage.

Une autre amélioration possible repose sur l'investigation d'autres métriques que la DTW ou le GA, et d'autres noyaux basés ou non sur ces métriques, pour la comparaison des motifs extraits lors de l'estimation en ligne du SOH. De nombreuses méthodes étant basées sur ces procédés, telles que les algorithmes de classification ou l'estimation en ligne du SOH (EEL), il semble possible d'améliorer la précision obtenue dans cette étude par la modification de certaines étapes de la méthodologie globale.

L'outil de suivi de flotte, et d'optimisation des Usages et des mises en places de caractérisations, offre une introduction et un aperçu des résultats possibles par ce type d'approche. Les résultats démontrent en effet l'amélioration des estimations de SOH grâce à la méthodologie développée. Ce travail préliminaire nécessite également de plus amples investigations, pour la définition des paramètres et des stratégies choisies, pour l'obtention d'une optimisation plus performante et d'une stratégie optimale dans la préconisation de caractérisations.

Perspectives globales de l'approche proposée

Les résultats obtenus de compréhension, l'estimation et la gestion des phénomènes de vieillissement d'une batterie en usage sur un VE, uniquement à partir des données mesurées, présentent de nombreuses perspectives. Cette approche étant peu utilisée dans ce domaine d'application, ces perspectives portent sur la résolution des limites évoquées de ces travaux mais également sur l'exploitation d'autres informations contenues dans ces signaux mesurés.

Concernant les limites de cette étude, nous pouvons imaginer diverses propositions de solutions. L'important besoin d'expérimentations, et donc de données, peut être pallié par une définition préalable des conditions d'expérimentations réelles nécessaires à la construction d'une base de données exhaustive. Cette problématique définit le cadre d'application des plans d'expériences, qui peut permettre une importante réduction des coûts et de la durée d'obtention des données voulues. Ce type de procédé est en effet commun dans les domaines où les expérimentations sont longues et coûteuses, par exemple pour les applications médicales.

L'outil de suivi de flotte évoqué dans ces travaux requiert une étude dédiée permettant la mise en place de méthodologies plus fines, proposant l'obtention de meilleurs résultats et par conséquent une optimisation des durées de vies de batteries d'une flotte de VE.

Les limites détaillées des méthodologies choisies illustrent également le besoin de nouveaux développements pour la mise à jour des algorithmes d'estimation du SOH. Cet axe associé aux contraintes du calcul en embarqué propose une problématique complexe qui nécessite d'être approfondie par la suite.

Comme évoqué précédemment, il est important de parvenir à une estimation fiable du niveau de résistance d'une batterie lors des usages réels, sans prise en compte de modèles externes. Une investigation approfondie des signaux mesurés et du comportement de la batterie doit être mise en place pour cela. Cette étude pouvant également mettre en évidence des particularités de comportements non identifiées par les autres types d'approches couramment déployées.

La connaissance précise des durées de vie des batteries est également un atout dans la création de modèles économiques viables. Ainsi, dans le cadre d'une utilisation bidirectionnelle des batteries, une estimation fiable du SOH permet d'optimiser son uti-

lisation sur le réseau global.

L'investigation des signaux offrent de nombreuses opportunités sur l'identification du niveau de performance d'une batterie. Ainsi, il ne fait aucun doute que l'approche proposée, basée sur la sélection des accélérations spécifiques, n'est qu'une part des modifications de comportement détectables par l'étude des signaux.

La prise en compte des erreurs de mesures dans les approches développées n'a pas été abordée. Or, dans un contexte de déploiement aux véhicules privés des méthodologies issues de ces travaux, il est important de considérer les diverses problématiques de mesures inhérentes aux applications réelles. La vérification de la qualité et de la fiabilité des mesures utilisées se doit d'être solutionnée pour garantir la robustesse des méthodologies basées sur les données.

Enfin, des problématiques liées aux phénomènes de vieillissement n'ont pas été évoqués. Ceux-ci sont de moins grande importance pour le développement du marché des VE, mais constituent les problématiques futures des études de vieillissement des batteries en usage sur des VE. Ces problématiques consistent en différents points :

- La détection spécifique de dysfonctionnement de la batterie à partir des signaux mesurés
- La mise en place d'un avertisseur fiable et quantifié, en embarqué, sur le comportement d'un utilisateur pour le vieillissement de la batterie
- La fiabilisation des garanties proposées, en fonction des usages effectués
- Le suivi de la sécurité des batteries par rapport aux phénomènes de vieillissement
- La prise en compte des applications bidirectionnelles des batteries

De ce fait, de nombreuses opportunités de recherches se présentent devant les travaux concernant le vieillissement des batteries en usages sur des VE.

Bibliographie

- [1] E. D. T. A. (EDTA) <http://www.electricdrive.org/>
- [2] E. V. I. (EVI) <http://www.evi-usa.com/>
- [3] EVWorld <http://evworld.com/>
- [4] D. Block, Electric vehicle sales and future projections, Tech. rep., Electric Vehicle Transportation Center (2014).
- [5] B. Tuttle, Falling electric car prices are saving — or destroying — the ev market, Business Time (2013).
- [6] S. Carley, R. M. Krause, B. W. Lane, J. D. Graham, Intent to purchase a plug-in electric vehicle : A survey of early impressions in large {US} cites, Transportation Research Part D : Transport and Environment 18 (0) (2013) 39 – 45.
- [7] A. Hayfield, European ev charging infrastructure – current and future projections (2013).
- [8] IHS, <http://press.ihs.com/press-release/country-industry-forecasting/electric-vehicle-charging-stations-set-soar-2020-%E2%80%93-ihs-st>
- [9] Y. Ye, L. H. Saw, Y. Shi, K. Somasundaram, A. A. Tay, Effect of thermal contact resistances on fast charging of large format lithium ion batteries, Electrochimica Acta 134 (0) (2014) 327 – 337.
- [10] Deloitte, Electric vehicle realities versus consumer expectations (2011).
- [11] U. D. of Transportation, National household transportation survey (2009).
- [12] “Range anxiety” in 2nd place on norway’s “words of the year” list. <http://insideevs.com/>
- [13] T. Nagaura, K. Tozawa, Lithium ion rechargeable battery, Progress in Batteries and Solar Cells 209 (1990) 9.
- [14] R. Brodd, Comments on the history of lithium-ion batteries (2002).
- [15] T. Kodama, H. Sakaebe, Present status and future prospect for national project on lithium batteries, Journal of Power Sources 81-82 (0) (1999) 144 – 149.
- [16] Developments in lithium-ion battery technology in the peoples republic of china, Tech. rep., Argonne National Laboratory (2008).
- [17] T. B. C. Group, Batteries for electric cars challenges, opportunities, and the outlook to 2020 (2010).
- [18] A. Zenati, P. Desprez, H. Razik, Estimation of the soc and the soh of li-ion batteries, by combining impedance measurements with the fuzzy logic inference, in : IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society, 2010, pp. 1773 –1778.

- [19] M. Augustine, Modeling li-ion battery capacity depletion in a particle filtering framework computational sustainability (2010).
- [20] J. Vetter, P. Novák, M. Wagner, C. Veit, K.-C. Möller, J. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, A. Hammouche, Ageing mechanisms in lithium-ion batteries, *Journal of Power Sources* 147 (1-2) (2005) 269 – 281.
- [21] R. Spotnitz, Advanced ev and hev batteries, *Vehicle Power and Propulsion*, 2005 IEEE Conference (2005) 4.
- [22] Plug-in hybrid electric vehicle r&d plan, Tech. rep., U.S. Department Of Energy (February 2007).
- [23] V. Pop, H. J. Bergveld, P. H. L. Notten, P. P. L. Regtien, State-of-the-art of battery state-of-charge determination, *Institute Of Physics Publishing* 16 (12) (2005) R93–R110.
- [24] F. Joho, P. Novak, M. E. Spahr, Safety aspects of graphite negative electrode materials for lithium-ion batteries, *Journal of the Electrochemical Society* 149 (2002) A1020–A1024.
- [25] H. Jannesari, M. Emami, C. Ziegler, Effect of electrolyte transport properties and variations in the morphological parameters on the variation of side reaction rate across the anode electrode and the aging of lithium ion batteries, *Journal of Power Sources* 196 (22) (2011) 9654 – 9664.
- [26] H. Buqa, A. Würsig, J. Vetter, M. Spahr, F. Krumeich, P. Novák, Sei film formation on highly crystalline graphitic materials in lithium-ion batteries, *Journal of Power Sources* 153 (2) (2006) 385 – 390, selected papers presented at the 2004 Meeting of the International Battery Association 2004 International Meeting of the International Battery Association.
- [27] P. Balakrishnan, R. Ramesh, T. P. Kumar, Safety mechanisms in lithium-ion batteries, *Journal of Power Sources* 155 (2) (2006) 401 – 414.
- [28] Q. Wang, P. Ping, X. Zhao, G. Chu, J. Sun, C. Chen, Thermal runaway caused fire and explosion of lithium ion battery, *Journal of Power Sources* 208 (0) (2012) 210 – 224.
- [29] D. Goers, M. E. Spahr, A. Leone, W. Märkle, P. Novák, The influence of the local current density on the electrochemical exfoliation of graphite in lithium-ion battery negative electrodes, *Electrochimica Acta* 56 (11) (2011) 3799 – 3808.
- [30] M. Broussely, S. Herreyre, P. Biensan, P. Kasztejna, K. Nechev, R. Staniewicz, Aging mechanism in li ion cells and calendar life predictions, *Journal of Power Sources* 97-98 (0) (2001) 13 – 21, proceedings of the 10th International Meeting on Lithium Batteries.
- [31] T. Abe, H. Fukuda, Y. Iriyama, Z. Ogumi, Solvated li-ion transfer at interface between graphite and electrolyte, *Journal of the Electrochemical Society* 151 (2004) A1120–A1123.
- [32] G.-C. Chung, H.-J. Kim, S.-I. Yu, S.-H. Jun, J. wook Choi, M.-H. Kimb, Origin of graphite exfoliation an investigation of the important role of solvent cointercalation, *Journal of the Electrochemical Society* 147 (2000) 4391–4398.
- [33] S. S. Zhang, A review on electrolyte additives for lithium-ion batteries, *Journal of Power Sources* 162 (2) (2006) 1379 – 1394, Special issue including selected

- papers from the International Power Sources Symposium 2005 together with regular papers.
- [34] S. Bourlot, P. Blanchard, S. Robert, Investigation of aging mechanisms of high power li-ion cells used for hybrid electric vehicles, *Journal of Power Sources* 196 (16) (2011) 6841 – 6846, 15th International Meeting on Lithium Batteries (IMLB).
 - [35] A. Barré, B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, D. Riu, A review on lithium-ion battery ageing mechanisms and estimations for automotive applications, *Journal of Power Sources* 241 (0) (2013) 680 – 689.
 - [36] R. N. Methekar, P. W. C. Northrop, K. Chen, R. D. Braatz, V. R. Subramaniana, Kinetic monte carlo simulation of surface heterogeneity in graphite anodes for lithium-ion batteries : Passive layer formation, *Journal of The Electrochemical Society* 158 (4) (2011) A363–A370.
 - [37] M. Koltypin, D. Aurbach, L. Nazar, B. Ellis, More on the performance of lifepo4 electrodes—the effect of synthesis route, solution composition, aging, and temperature, *Journal of Power Sources* 174 (2) (2007) 1241 – 1250, 13th International Meeting on Lithium Batteries.
 - [38] S. Zhang, K. Xu, T. Jow, Low temperature performance of li-ion cells with a libf4 based electrolyte, *Journal of Solid State Electrochemistry* 7 (2002) 147–151.
 - [39] S. Zhang, K. Xu, T. Jow, The low temperature performance of li-ion batteries, *Journal of Power Sources* 115 (1) (2003) 137 – 140.
 - [40] M. Dubarry, C. Truchot, B. Y. Liaw, Synthesize battery degradation modes via a diagnostic and prognostic model, *Journal of Power Sources* 219 (0) (2012) 204 – 216.
 - [41] P. Liu, J. Wang, J. Hicks-Garner, E. Sherman, S. Soukiazian, M. Verbrugge, H. Tataria, J. Musser, P. Finamore, Aging mechanisms of lifepo4 batteries deduced by electrochemical and structural analyses, *Journal of The Electrochemical Society* 157 (4) (2010) A499–A507.
 - [42] M. Kassem, J. Bernard, R. Revel, S. Pélissier, F. Duclaud, C. Delacourt, Calendar aging of a graphite/lifepo4 cell, *Journal of Power Sources* 208 (0) (2012) 296 – 305.
 - [43] B. L. Ellis, K. T. Lee, L. F. Nazar, Positive electrode materials for li-ion and li-batteries, *Chemistry of materials* 22 (2009) 691–714.
 - [44] K. Xu, A. von Cresce, Interfacing electrolytes with electrodes in li ion batteries, *Journal of Materials Chemistry* 21 (2011) 9849–9864.
 - [45] K. Edstrom, T. Gustafsson, J. O. Thomas, The cathode-electrolyte interface in the li-ion battery, *Electrochimica Acta* 50 (2-3) (2004) 397–403.
 - [46] M. Kerlau, M. Marcinek, V. Srinivasan, R. M. Kostecki, Studies of local degradation phenomena in composite cathodes for lithium-ion batteries, *Electrochimica Acta* 52 (17) (2007) 5422 – 5429.
 - [47] K. Amine, J. Liu, I. Belharouak, High-temperature storage and cycling of c-lifepo4/graphite li-ion cells, *Electrochemistry Communications* 7 (7) (2005) 669 – 673.

- [48] J. Zhu, K. Zeng, L. Lu, Cycling effects on surface morphology, nanomechanical and interfacial reliability of LiMn_2O_4 cathode in thin film lithium ion batteries, *Electrochimica Acta* 68 (0) (2012) 52 – 59.
- [49] M. Wohlfahrt-Mehrens, C. Vogler, J. Garche, Aging mechanisms of lithium cathode materials, *Journal of Power Sources* 127 (1-2) (2004) 58 – 64, eighth Ulmer Electrochemische Tage.
- [50] E. Meissner, G. Richter, The challenge to the automotive battery industry : the battery has to become an increasingly integrated component within the vehicle electric power system, *Journal of Power Sources* 144 (2) (2005) 438 – 460, selected papers from the Ninth European Lead Battery Conference Ninth European Lead Battery Conference.
- [51] R. Wright, C. Motloch, J. Belt, J. Christophersen, C. Ho, R. Richardson, I. Bloom, S. Jones, V. Battaglia, G. Henriksen, T. Unkelhaeuser, D. Ingersoll, H. Case, S. Rogers, R. Sutula, Calendar- and cycle-life studies of advanced technology development program generation 1 lithium-ion batteries, *Journal of Power Sources* 110 (2) (2002) 445 – 470.
- [52] R. P. Ramasamy, R. E. White, B. N. Popov, Calendar life performance of pouch lithium-ion cells, *Journal of Power Sources* 141 (2) (2005) 298 – 306.
- [53] W. Bögel, J. P. Büchel, H. Katz, Real-life ev battery cycling on the test bench, *Journal of Power Sources* 72 (1) (1998) 37 – 42.
- [54] I. Bloom, B. Cole, J. Sohn, S. Jones, E. Polzin, V. Battaglia, G. Henriksen, C. Motloch, R. Richardson, T. Unkelhaeuser, D. Ingersoll, H. Case, An accelerated calendar and cycle life study of li-ion cells, *Journal of Power Sources* 101 (2) (2001) 238 – 247.
- [55] K. Amine, C. Chen, J. Liu, M. Hammond, A. Jansen, D. Dees, I. Bloom, D. Vissers, G. Henriksen, Factors responsible for impedance rise in high power lithium ion batteries, *Journal of Power Sources* 97-98 (0) (2001) 684 – 687, proceedings of the 10th International Meeting on Lithium Batteries.
- [56] S. Zhang, K. Xu, T. Jow, Electrochemical impedance study on the low temperature of li-ion batteries, *Electrochimica Acta* 49 (7) (2004) 1057 – 1061.
- [57] M. Broussely, P. Biensan, F. Bonhomme, P. Blanchard, S. Herreyre, K. Nechev, R. Staniewicz, Main aging mechanisms in li ion batteries, *Journal of Power Sources* 146 (1-2) (2005) 90 – 96, selected papers presented at the 12th International Meeting on Lithium Batteries.
- [58] K. Nunotani, F. Yoshida, Y. Kamiya, Y. Daisho, K. Abe, M. Kono, H. Matsuo, Development and performance evaluation of lithium iron phosphate battery with superior rapid charging performance, *Vehicle Power and Propulsion Conference (VPPC), IEEE* (2011) 1–4.
- [59] J. Gnanaraj, E. Zinigrad, L. Asraf, H. Gottlieb, M. Sprecher, D. Aurbach, M. Schmidt, The use of accelerating rate calorimetry (arc) for the study of the thermal reactions of li-ion battery electrolyte solutions, *Journal of Power Sources* 119–121 (0) (2003) 794 – 798, Selected papers presented at the 11th International Meeting on Lithium Batteries.

- [60] R. Kötz, P. Ruch, D. Cericola, Aging and failure mode of electrochemical double layer capacitors during accelerated constant load tests, *Journal of Power Sources* 195 (3) (2010) 923 – 928.
- [61] H. Gong, Y. Yu, T. Li, T. Mei, Z. Xing, Y. Zhu, Y. Qian, X. Shen, Solvothermal synthesis of LiFePO_4/C nanopolyhedrons and microellipsoids and their performance in lithium-ion batteries, *Materials Letters* 66 (1) (2012) 374 – 376.
- [62] K. Asakura, M. Shimomura, T. Shodai, Study of life evaluation methods for li-ion batteries for backup applications, *Journal of Power Sources* 119-121 (0) (2003) 902 – 905, selected papers presented at the 11th International Meeting on Lithium Batteries.
- [63] D. Doerffel, S. A. Sharkh, A critical review of using the peukert equation for determining the remaining capacity of lead-acid and lithium-ion batteries, *Journal of Power Sources* 155 (2) (2006) 395 – 400.
- [64] M. Conte, F. V. Conte, I. Bloom, K. Morita, T. Ikeya, J. Belt, Ageing testing procedures on lithium batteries in an international collaboration context, 25th World Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exhibition, (2010).
- [65] E. Meissner, G. Richter, Battery monitoring and electrical energy management : Precondition for future vehicle electric power systems, *Journal of Power Sources* 116 (1-2) (2003) 79 – 98, selected Papers Presented at the Eighth European Lead Battery Conference.
- [66] M. Ecker, J. B. Gerschler, J. Vogel, S. Käbitz, F. Hust, P. Dechent, D. U. Sauer, Development of a lifetime prediction model for lithium-ion batteries based on extended accelerated aging test data, *Journal of Power Sources* (0) (2012).
- [67] M. Doyle, J. Newman, Modeling the performance of rechargeable lithium-based cells : design correlations for limiting cases, *Journal of Power Sources* 54 (1) (1995) 46–51.
- [68] M. Doyle, J. Newman, A. S. Gozdz, C. N. Schmutz, J.-M. Tarascon, Comparison of modeling predictions with experimental data from plastic lithium ion cells, *Journal of The Electrochemical Society* 143 (6) (1996) 1890–1903.
- [69] J. Newman, W. Tiedemann, Porous-electrode theory with battery applications, *AIChE Journal* 21 (1) (1975) 25–41.
- [70] R. Darling, J. Newman, Modeling side reactions in composite LiMn_2O_4 electrodes, *Journal of The Electrochemical Society* 145 (3) (1998) 990–998.
- [71] J. Christensen, J. Newman, Effect of anode film resistance on the charge/discharge capacity of a lithium-ion battery, *Journal of The Electrochemical Society* 150 (11) (2003) A1416–A1420.
- [72] J. Christensen, J. Newman, A mathematical model for the lithium-ion negative electrode solid electrolyte interphase, *Journal of The Electrochemical Society* 151 (11) (2004) A1977–A1988.
- [73] J. Christensen, J. Newman, A mathematical model of stress generation and fracture in lithium manganese oxide, *Journal of The Electrochemical Society* 153 (6) (2006) A1019–A1030.
- [74] H. J. Ploehn, P. Ramadass, R. E. White, Solvent diffusion model for aging of lithium-ion battery cells, *Journal of The Electrochemical Society* 151 (3) (2004) A456–A462.

- [75] V. Ramadesigan, V. Boovaragavan, J. C. Pirkle, V. R. Subramanian, Efficient reformulation of solid-phase diffusion in physics-based lithium-ion battery models, *Journal of The Electrochemical Society* 157 (7) (2010) A854–A860.
- [76] V. Ramadesigan, K. Chen, N. A. Burns, V. Boovaragavan, R. D. Braatz, V. R. Subramanian, Parameter estimation and capacity fade analysis of lithium-ion batteries using reformulated models, *Journal of The Electrochemical Society* 158 (9) (2011) A1048–A1054.
- [77] M. Safari, C. Delacourt, Mathematical modeling of lithium iron phosphate electrode : Galvanostatic charge/discharge and path dependence, *Journal of The Electrochemical Society* 158 (2) (2011) A62–A73.
- [78] M. Safari, C. Delacourt, Simulation-Based Analysis of Aging Phenomena in a Commercial Graphite/lifepo4 Cell, Vol. 158, Electrochemical Society, Pennington, NJ, USA, 2011.
- [79] M. Safari, C. Delacourt, Aging of a commercial graphite/lifepo4 cell, *Journal of The Electrochemical Society* 158 (10) (2011) A1123–A1135.
- [80] S. Johnson, Hierarchical clustering schemes, *Psychometrika* 32 (3) (1967) 241–254.
- [81] A. L. Dalverny, J. S. Filhol, M. L. Doublet, Interface electrochemistry in conversion materials for li-ion batteries, *Journal of Materials Chemistry* 21 (27) (2011) 10134–10142.
- [82] M. Wagemaker, D. P. Singh, W. J. H. Borghols, U. Lafont, L. Haverkate, V. K. Peterson, F. M. Mulder, Dynamic solubility limits in nanosized olivine lifepo4, *Journal of the American Chemical Society* 133 (26) (2011) 10222–10228.
- [83] K. Tasaki, A. Goldberg, J.-J. Lian, M. Walker, A. Timmons, S. J. Harris, Solubility of lithium salts formed on the lithium-ion battery negative electrode surface in organic solvents, *Journal of The Electrochemical Society* 156 (12) (2009) A1019–A1027.
- [84] K. Tasaki, S. J. Harris, Computational study on the solubility of lithium salts formed on lithium ion battery negative electrode in organic solvents, *The Journal of Physical Chemistry C* 114 (17) (2010) 8076–8083.
- [85] K. Leung, J. L. Budzien, Ab initio molecular dynamics simulations of the initial stages of solid-electrolyte interphase formation on lithium ion battery graphitic anodes, *Physical Chemistry Chemical Physics* 12 (25) (2010) 6583–6586.
- [86] S. Shi, P. Lu, Z. Liu, Y. Qi, L. G. Hector, H. Li, S. J. Harris, Direct calculation of li-ion transport in the solid electrolyte interphase, *Journal of the American Chemical Society* 134 (37) (2012) 15476–15487.
- [87] D. Haifeng, W. Xuezhe, S. Zechang, A new soh prediction concept for the power lithium-ion battery used on hevs, in : *Vehicle Power and Propulsion Conference (VPPC)*, 2009 IEEE, 2009, pp. 1649 –1653.
- [88] M. Einhorn, V. Conte, C. Kral, J. Fleig, R. Permann, Parameterization of an electrical battery model for dynamic system simulation in electric vehicles, in : *Vehicle Power and Propulsion Conference (VPPC)*, 2010 IEEE, 2010, pp. 1 –7.
- [89] C. Blanco, L. Sanchez, M. Gonzalez, J. Anton, V. Garcia, J. Viera, An equivalent circuit model with variable effective capacity for lifepo4 batteries (2014).

- [90] A. Eddahech, O. Briat, H. Henry, J.-Y. Delétage, E. Woirgard, J.-M. Vinassa, Ageing monitoring of lithium-ion cell during power cycling tests, *Microelectronics Reliability* 51 (9-11) (2011) 1968 – 1971, proceedings of the 22th European Symposium on the Reliability Of Electron Devices, Failure Physics And Analysis.
- [91] D. Andre, C. Appel, T. Soczka-Guth, D. U. Sauer, Advanced mathematical methods of soc and soh estimation for lithium-ion batteries, *Journal of Power Sources* 224 (0) (2013) 20 – 27.
- [92] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
- [93] M. E. Tipping, Sparse bayesian learning and the relevance vector machine, *J. Mach. Learn. Res.* 1 (2001) 211–244.
- [94] C. M. Bishop, *Pattern recognition and machine learning*, 1st Edition, Springer, 2006.
- [95] B. Saha, K. Goebel, S. Poll, J. Christophersen, An integrated approach to battery health monitoring using bayesian regression and state estimation, *Autotestcon, 2007 IEEE* (2007) 646–653.
- [96] B. Saha, K. Goebel, S. Poll, J. Christophersen, Prognostics methods for battery health monitoring using a bayesian framework, *IEEE Transactions on instrumentation and measurement* 58 (2) (2009) 291–297.
- [97] O. Erdinc, B. Vural, M. Uzunoglu, A dynamic lithium-ion battery model considering the effects of temperature and capacity fading (2009).
- [98] B. Y. Liaw, E. Roth, R. G. Jungst, G. Nagasubramanian, H. L. Case, D. H. Doughty, Correlation of arrhenius behaviors in power and capacity fades with cell impedance and heat generation in cylindrical lithium-ion cells, *Journal of Power Sources* 119-121 (0) (2003) 874 – 886, selected papers presented at the 11th International Meeting on Lithium Batteries.
- [99] J. Belt, V. Utgikar, I. Bloom, Calendar and phev cycle life aging of high-energy, lithium-ion cells containing blended spinel and layered-oxide cathodes, *Journal of Power Sources* 196 (23) (2011) 10213 – 10221.
- [100] K. Smith, G.-H. Kim, A. Pesaran, Modeling of nonuniform degradation in large-format li-ion batteries, presented at the 215th Electrochemical Society Meeting, 25-29 May 2009, San Francisco, CA (May 2009).
- [101] E. Thomas, I. Bloom, J. Christophersen, V. Battaglia, Statistical methodology for predicting the life of lithium-ion cells via accelerated degradation testing, *Journal of Power Sources* 184 (1) (2008) 312 – 317.
- [102] H. Wenzl, I. Baring-Gould, R. Kaiser, B. Y. Liaw, P. Lundsager, J. Manwell, A. Ruddell, V. Svoboda, Life prediction of batteries for selecting the technically most suitable and cost effective battery, *Journal of Power Sources* 144 (2) (2005) 373 – 384, selected papers from the Ninth European Lead Battery Conference.
- [103] P. Ramadass, B. Haran, R. White, B. N. Popov, Mathematical modeling of the capacity fade of li-ion cells, *Journal of Power Sources* 123 (2) (2003) 230 – 240.
- [104] J. Wang, P. Liu, J. Hicks-Garner, E. Sherman, S. Soukiazian, M. Verbrugge, H. Tataria, J. Musser, P. Finamore, Cycle-life model for graphite-lifepo4 cells, *Journal of Power Sources* 196 (8) (2011) 3942 – 3948.

- [105] S. Sankarasubramanian, B. Krishnamurthy, A capacity fade model for lithium-ion batteries including diffusion and kinetics, *Electrochimica Acta* 70 (0) (2012) 248 – 254.
- [106] W. Gu, Z. Sun, X. Wei, H. Dai, A capacity fading model of lithium-ion battery cycle life based on the kinetics of side reactions for electric vehicle applications, *Electrochimica Acta* 133 (0) (2014) 107 – 116.
- [107] S. Mishra, M. Pecht, T. Smith, I. McNee, R. Harris, Remaining life prediction of electronic products using life consumption monitoring approach, *European Microelectronics Packaging and Interconnection Symposium*.
- [108] V. Marano, S. Onori, Y. Guezennec, G. Rizzoni, N. Madella, Lithium-ion batteries life estimation for plug-in hybrid electric vehicles, *Vehicle Power and Propulsion Conference, 2009. VPPC '09. IEEE* (2009) 536–543.
- [109] M. Safari, M. Morcrette, A. Teyssot, C. Delacourt, Life-prediction methods for lithium-ion batteries derived from a fatigue approach, *Journal of The Electrochemical Society* 157 (6) (2010) A713–A720.
- [110] K. Takei, K. Kumai, Y. Kobayashi, H. Miyashiro, N. Terada, T. Iwahori, T. Tanaka, Cycle life estimation of lithium secondary battery by extrapolation method and accelerated aging test, *Journal of Power Sources* 97-98 (0) (2001) 697 – 701, proceedings of the 10th International Meeting on Lithium Batteries.
- [111] C. Chen, M. Pecht, Prognostics of lithium-ion batteries using model based and data-driven methods (2012).
- [112] K. S. Ng, C.-S. Moo, Y.-P. Chen, Y.-C. Hsieh, Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries, *Applied Energy* 86 (9) (2009) 1506 – 1511.
- [113] T. Hansen, C.-J. Wang, Support vector based battery state of charge estimator, *Journal of Power Sources* 141 (2) (2005) 351 – 358.
- [114] L. A. Zadeh, Is there a need for fuzzy logic?, *Information Sciences* 178 (13) (2008) 2751 – 2779.
- [115] K. Tsang, W. Chan, State of health detection for lithium ion batteries in photovoltaic system, *Energy Conversion and Management* 65 (0) (2013) 7 – 12.
- [116] A. J. Salkind, C. Fennie, P. Singh, T. Atwater, D. E. Reisner, Determination of state-of-charge and state-of-health of batteries by fuzzy logic methodology, *Journal of Power Sources* 80 (1-2) (1999) 293 – 300.
- [117] P. Singh, R. Vinjamuri, X. Wang, D. Reisner, Fuzzy logic modeling of eis measurements on lithium-ion batteries, *Electrochimica Acta* 51 (8-9) (2006) 1673 – 1679, selection of papers from the 6th International Symposium (EIS 2004) 16-21 May 2004, Cocoa Beach, FL, USA.
- [118] I.-S. Kim, Nonlinear state of charge estimator for hybrid electric vehicle battery, *Power Electronics, IEEE Transactions on* 23 (4) (2008) 2027 –2034.
- [119] C. Hu, B. D. Youn, J. Chung, A multiscale framework with extended kalman filter for lithium-ion battery soc and capacity estimation, *Applied Energy* (0) (2011).

- [120] G. L. Plett, Extended kalman filtering for battery management systems of lipb-based hev battery packs : Part 3. state and parameter estimation, *Journal of Power Sources* 134 (2) (2004) 277 – 292.
- [121] M. Urbain, Modelisation électrique et énergétique des accumulateurs lithium ion estimation en ligne du soc et du soh, Ph.D. thesis, INP Lorraine (2009).
- [122] D. V. Do, Diagnostic de batteries lithium ion dans des applications embarquées, Ph.D. thesis, Université de Technologie de Compiègne (2010).
- [123] J. Lampinen, A. Vehtari, Bayesian approach for neural networks—review and case studies, *Neural Networks* 14 (3) (2001) 257 – 274.
- [124] G. M. Jenkins, G. E. P. Box, *Time series analysis : forecasting and control*, Prentice Hall, 1994.
- [125] J. Kozłowski, Electrochemical cell prognostics using online impedance measurements and model-based data fusion techniques, *IEEE Aerospace Conference* 7 (2003) 3257–3270.
- [126] W. Weibull, A statistical distribution function of wide applicability, *Journal of applied mechanics* (1951) 293–297.
- [127] S.-W. Eom, M.-K. Kim, I.-J. Kim, S.-I. Moon, Y.-K. Sun, H.-S. Kim, Life prediction and reliability assessment of lithium secondary batteries, *Journal of Power Sources* 174 (2) (2007) 954 – 958, 13th International Meeting on Lithium Batteries.
- [128] Circuit Ladoux,
<http://www.michelin-engineering-and-services.com/mes/front/affich.jsp?codeRubrique=20060614155737&lang=FR>
- [129] M. André, Real-world driving cycles for measuring cars pollutant emissions – part a : The artemis european driving cycles (2004).
- [130] Z. Younes, L. Boudet, F. Suard, M. Gérard, R. Rioux, Analysis of the main factors influencing the energy consumption of electric vehicles, in : *Electric Machines Drives Conference (IEMDC), 2013 IEEE International*, 2013, pp. 247–253.
- [131] J. L. Rodgers, A. W. Nicewander, Thirteen ways to look at the correlation coefficient, *The American Statistician* 42 (1) (1988) 59–66.
- [132] S. Ihara, *Information theory for continuous systems. Vol. 2*, Singapore : World Scientific, 1993.
- [133] L. Batina, B. Gierlich, E. Prouff, M. Rivain, F. S. and N. Veyrat-Charvillon, Mutual information analysis : a comprehensive study, *Journal of Cryptology* 24 (2010) 269–291.
- [134] W. Li, Mutual information functions versus correlation functions, *Journal of Statistical Physics* 60 (1990) 823–837.
- [135] I. Jolliffe, *Principal Component Analysis*, *Encyclopedia of Statistics in Behavioral Science*, 2005.
- [136] A. B. Costello, Getting the most from your analysis, *Pan* 12 (2) (2009) 131–146.
- [137] R. H. Myers, *Classical and modern regression with applications*.
- [138] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, *The Annals of Statistics* 32 (2) (2004) 407–499.

- [139] A. E. Hoerla, R. W. Kennarda, Ridge regression : Biased estimation for nonorthogonal problems, *Technom* 12 (1) (1970) 55–67.
- [140] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B (Methodological)* 58 (1996) 267–288.
- [141] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 67 (2) (2005) 301–320.
- [142] J. Bigot, Recalage de signaux et analyse de variance fonctionnelle par ondelettes - applications au domaine biomédical, Ph.D. thesis, Université Joseph Fourier (2003).
- [143] T. Gasser, A. Kneip, Searching for structure in curve samples, *Journal of the American Statistical Association* 90 (432) (1995) 1179–1188.
- [144] H. Sakoe, S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, *Acoustics, Speech and Signal Processing, IEEE Transactions on* 26 (1) (1978) 43 – 49.
- [145] M. Cuturi, J.-P. Vert, Ø. Birkenes, T. Matsui, A kernel for time series based on global alignments, in : *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, Vol. 2, IEEE, 2007*, pp. II–413.
- [146] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection 14 (2).
- [147] R. Tibshirani, G. Walther, T. Hastie, Estimating the number of clusters in a data set via the gap statistic, *Journal of the Royal Statistical Society : Series B* 63 (2001) 411–423.
- [148] F. Suard, D. Mercier, Using kernel basis with relevance vector machine for feature selection, in : C. Alippi, M. Polycarpou, C. Panayiotou, G. Ellinas (Eds.), *Artificial Neural Networks – ICANN 2009, Vol. 5769 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009*, pp. 255–264.
- [149] C. Bahlmann, B. Haasdonk, H. Burkhardt, Online handwriting recognition with support vector machines - a kernel approach, in : *Frontiers in Handwriting Recognition, 2002. Proceedings. Eighth International Workshop on, 2002*, pp. 49–54.
- [150] B. Efron, Bootstrap methods : another look at the jackknife, *The Annals of Statistics* 7 (1) (1979) 1–26.
- [151] Toyota Fleet Management <http://www.toyotafleetmanagement.com.au/>
- [152] Verizon NetworkFleet <http://business.verizonwireless.com/content/b2b/en/solutions/technology/fleet-management.html>
- [153] Telogis Fleet http://www.telogis.com/solutions/fleet?utm_source=Capterra&utm_medium=cpc&utm_campaign=Fleet%2520Management
- [154] RTA Fleet Management Software <http://www.rtafleet.com/>
- [155] IMS Fleet Management Pro <http://www.mtcpro.com/fr/>
- [156] E. Nowak, F. Jurie, B. Triggs, Sampling Strategies for Bag-of-Features Image Classification, *Computer Vision ECCV 2006, Springer Berlin Heidelberg*.

- [157] B. W. Silverman, *Density estimation for statistics and data analysis*, School of Mathematics University of Bath, UK, 1986.
- [158] S. J. Sheather, *Density estimation*, *Statistical Science* 19 (2004) 588–597.
- [159] D. Gusfield, *Algorithms on Strings, Trees and Sequences : Computer Science and Computation Biology*, Cambridge University Press, 1997.
- [160] T. W. Liao, *Clustering of time series data—a survey*, *Pattern Recognition* 38 (2005) 1857–1874.
- [161] C. A. Ratanamahatana, E. Keogh, *Making time-series classification more accurate using learned constraints*, in : *In proc. of SDM International Conf, 2004*, pp. 11–22.
- [162] J. J. Rodríguez, C. J. Alonso, J. A. Maestro, *Support vector machines of interval-based features for time series classification*, *Knowledge-Based Systems* 18 (4) (2005) 171–178.
- [163] A. P. Shanker, A. Rajagopalan, *Off-line signature verification using dtw*, *Pattern Recognition Letters* 28 (12) (2007) 1407 – 1414.
- [164] J. Aach, G. M. Church, *Aligning gene expression time series with time warping algorithms*, *Bioinformatics* 17 (6) (2001) 495–508.
- [165] Z. Bar-Joseph, G. Gerber, D. K. Gifford, T. S. Jaakkola, I. Simon, *A new approach to analyzing gene expression time series data*, in : *Proceedings of the sixth annual international conference on Computational biology, RECOMB '02, ACM, New York, NY, USA, 2002*, pp. 39–48.
- [166] F. Petitjean, C. Kurtz, N. Passat, P. Gançarski, *Spatio-temporal reasoning for the classification of satellite image time series*, *Pattern Recognition Letters* 33 (13) (2012) 1805 – 1815.
- [167] F. Petitjean, A. Ketterlin, P. Gançarski, *A global averaging method for dynamic time warping, with applications to clustering*, *Pattern Recognition* 44 (3) (2011) 678 – 693.
- [168] D. Gavrilu, L. Davis, *3-d model-based tracking of humans in action : a multi-view approach*, in : *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on, 1996*, pp. 73 –80.
- [169] T. Rath, R. Manmatha, *Word image matching using dynamic time warping*, in : *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on, Vol. 2, 2003*, pp. II-521 – II-527 vol.2. doi : 10.1109/CVPR.2003.1211511.
- [170] H. Sakoe, S. Chiba, *A dynamic programming approach to continuous speech recognition*, in *Proceedings of the Seventh International Congress on Acoustics, Budapest, 1971*.
- [171] E. V. Ruiz, *An algorithm for finding nearest neighbours in (approximately) constant average time*, *Pattern Recognition Letters* 4 (3) (1986) 145 – 157.
- [172] P. H. W. Wong, O. C. Au, J. W. C. Wong, W. H. B. Lau, *Reducing computational complexity of dynamic time warping-based isolated word recognition with time scale modification*, *Signal Processing Proceedings, 1998. ICSP '98. 1998 Fourth International Conference on* 1 (1998) 722–725.

- [173] Q. Zhu, G. E. A. P. A. Batista, T. Rakthanmanon, E. J. Keogh, A novel approximation to dynamic time warping allows anytime clustering of massive time series datasets, in : *SDM'12*, 2012, pp. 999–1010.
- [174] C. Gouy-Pailler, H. Najmeddine, A. Mouraud, F. Suard, A. Jay, P. Maréchal, C. Spitz, Exploring incas : Multivariate data mining techniques for sensor selection in low-energy consumption buildings, *Joint Conference CIB W078 - W102* – Oct. 2011.
- [175] V. Niennattrakul, C. Ratanamahatana, Shape averaging under time warping, in : *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2009. ECTI-CON 2009. 6th International Conference on*, Vol. 02, 2009, pp. 626 –629.
- [176] L. Gupta, D. Molfese, R. Tammana, P. Simos, Nonlinear alignment and averaging for estimating the evoked potential, *IEEE Transactions on Biomedical Engineering* 4 (1996) 348–356.
- [177] V. Niennattrakul, C. A. Ratanamahatana, On clustering multimedia time series data using k-means and dynamic time warping, *International Conference on Multimedia and Ubiquitous Engineering* (2007) 733–738.
- [178] V. Niennattrakul, D. Srisai, C. A. Ratanamahatana, Shape-based template matching for time series data, *Knowledge-Based Systems* 26 (0) (2012) 1 – 8.
- [179] S. Ongwattanakul, D. Srisai, Contrast enhanced dynamic time warping distance for time series shape averaging classification, in : *Proceedings of the 2nd International Conference on Interaction Sciences : Information Technology, Culture and Human*, ICIS '09, ACM, New York, NY, USA, 2009, pp. 976–981.
- [180] F. Petitjean, P. Gançarski, Summarizing a set of time series by averaging : From steiner sequence to compact multiple alignment, *Theoretical Computer Science* 414 (1) (2012) 76 – 91.
- [181] E. Keogh, Q. Zhu, B. Hu, Y. H. X. Li, L. Wei, C. Ratanamahatana, Ucr time series classification/clustering page.
http://www.cs.ucr.edu/~eamonn/time_series_data
- [182] K. Eriksson, D. Estep, C. Johnson, *Applied Mathematics : Body and Soul*, Springer, 2005.

Valorisations liées aux travaux

3	Reuves internationales
4	Conférences internationales
1	Brevet

Reuves internationales

- [1] A. Barré, B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, and D. Riu. A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *Journal of Power Sources*, 241(0) :680 – 689, 2013.
- [2] A. Barré, F. Suard, M. Gérard, M. Montaru, and D. Riu. Statistical analysis for understanding and predicting battery degradations in real-life electric vehicle use. *Journal of Power Sources*, 245(0) :846 – 856, 2014.
- [3] A. Barré, F. Suard, M. Gérard and D. Riu. Electric vehicles performance estimation through a patterns extraction and classification methodology. *Journal of Power Sources*, 273(0) :670 – 679, 2015.

Conférences internationales

- [1] A. Barré, M. Gérard, F. Suard, M. Montaru, and D. Riu. Statistical analysis of li-ion battery degradation based on recording data of real vehicle utilization. Poster session, Istanbul, Turkey, September 2012. 5th International Conference on Advanced Lithium Batteries for Automobile Applications (ABAA5).
- [2] A. Barré, F. Suard, M. Gérard, M. Montaru, and D. Riu. Statistical method tools to analyze ageing effects on li-ion battery performances. Detroit, USA, April 2013. SAE 2013 World Congress.

- [3] A. Barré, F. Suard, M. Gérard, and D. Riu. A real-time data-driven method for battery health prognostics in electric vehicle use. Nantes, France, July 2014. 2nd European Conference of the Prognostics and Health Management Society (PHME).
- [4] A. Barré, F. Suard, M. Gérard, and D. Riu. Battery capacity estimation and health management of an electric vehicle fleet. Coimbra, Portugal, October 2014. 10th IEEE Vehicle Power and Propulsion Conference (VPPC).

Brevet

- [1] A. Barré, F. Suard and M. Gérard. Procédé, dispositif, et système d'estimation de l'état de santé d'une batterie d'un véhicule électrique ou hybride en condition d'utilisation, et procédé de construction d'un modèle pour une telle estimation. FR1450025.

Cellules A123 Systems

Les cellules A123 Systems (figure A.1) sont connues comme étant des cellules *high power*, et proclamée par le constructeur comme proposant une des meilleures durée de vie.



FIGURE A.1 – Exemple de cellule A123 Systems

La fiche technique des cellules A123 Systems est décrite dans la table A.1.

Tension nominale	3.3 V
Capacité nominale	4.5 Ah
Capacité minimale	4.3 Ah
Densité énergétique	71 Wh.kg ⁻¹
Puissance déchargée	550 Wh
Poids	0.205 kg
Température en fonctionnement	-30 à 55°C
Température en stockage	-45 à 60°C

TABLE A.1 – Spécifications des cellules A123 Systems

Chapitre B

Cellules LiFeBatt

Les cellules LiFeBatt (figure B.1) fournissent des caractéristiques bâties pour les usages en VE. Ces cellules Li-ion sont utilisées par de nombreux organismes, par exemple l'U. S. Department of Energy.



FIGURE B.1 – Exemple de cellule LiFeBatt

La table B.1 détaille l'ensemble des spécifications fournies par le constructeur pour ce type de cellules.

Fenêtre de tension	2.10 - 3.65 V
Tension maximale	3.65 V
Courant de charge maximal	30 A
Pulse de courant maximal	140 A
Résistance ohmique interne	<6 mΩ
Capacité (1C)	9 Ah
Densité énergétique	80 Wh.kg ⁻¹
Poids	0.360 kg
Températures en fonctionnement	0 à 45°C
Températures en stockage	-10 à 45°C
Durée de vie	1 500 cycles

TABLE B.1 – Spécifications des cellules LifeBatt

Relevance Vector Machines

L'approche Relevance Vector Machines (RVM), initialement introduit par [93], est basé sur une formulation Bayésienne d'un modèle linéaire avec un *prior* adapté proposant des résultats dans une représentation parcimonieuse. Considérons l'ensemble d'apprentissage $B^{train} = \{(x_1, t_1), \dots, (x_N, t_N)\}$ constitué des motifs x_i et de leur valeur cible associée t_i , et considérons que $t_i = f(x_i) + \epsilon_i$, où les bruits ϵ_i sont supposés être indépendants et de distribution Gaussienne avec une moyenne nulle et une variance σ^2 , i.e. $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, $\forall i \in \{1, \dots, N\}$.

Le but est ici d'apprendre un modèle de dépendance entre les valeurs cibles t_i et les entrées x_i . Ce qui permettrait pour des nouveaux motifs non observés x d'obtenir des prédictions précises t . Typiquement, les prédictions sont basées sur une fonction $f(x)$ définie sur l'espace des observations, et l'apprentissage peut se définir comme un processus de déduction de ces paramètres. Cette fonction prend la forme :

$$f(x) = \sum_{i=1}^N w_i K(x, x_i) + w_0 \quad (\text{C.1})$$

où $K(x, x_i)$ est une fonction noyau et $w = [w_1, \dots, w_N]^T$ sont les poids et w_0 le biais.

Par conséquent, la probabilité d'un ensemble de données peut être écrite comme :

$$p(t|w, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2} \|t - \phi w\|^2\right\} \quad (\text{C.2})$$

où $\phi = [\phi(x_1), \dots, \phi(x_N)]^T$ est appelée la matrice de *design*, avec :

$$\phi(x_i) = [1, K(x_i, x_1), K(x_i, x_2), \dots, K(x_i, x_N)]^T \quad (\text{C.3})$$

Dans le but d'apprendre la relation entre x et t , nous souhaitons contraindre la complexité et par conséquent la croissance des poids w en définissant une distribution de probabilité préalable des w . La préférence, pour les fonctions régulières et donc moins complexes, est codée en utilisant une distribution des w comme étant une Gaussienne de moyenne nulle. Cela donne alors :

$$p(w|\alpha) = \sum_{i=1}^N \mathcal{N}(0, \alpha_i^{-1}) \quad (\text{C.4})$$

où α_i est utilisé pour décrire l'inverse de la variance pour chaque w_i . Cela revient alors à introduire un hyperparamètre α_i associé à chaque poids, modifiant la pertinence

du *prior*. Pour compléter la spécification de ce *prior* hiérarchique, nous définissons les *hyperpriors* en fonction de α ; ainsi que la variance du bruit σ^2 .

Ayant défini le *prior*, l'inférence Bayésienne procède en calculant le *posterior* pour chaque inconnue compte tenu des données, à partir de la règle de Bayes :

$$p(w, \alpha, \sigma^2 | t) = \frac{p(t | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2)}{p(t)} \quad (\text{C.5})$$

Assumant que l'objectif t_* d'un nouvel individu x_* soit utilisé pour faire une prédiction, on a :

$$p(t_* | t) = \int p(t_* | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2 | t) dw d\alpha d\sigma^2 \quad (\text{C.6})$$

On peut décomposer le *posterior* $p(w, \alpha, \sigma^2 | t)$ tel que :

$$p(w, \alpha, \sigma^2 | t) = p(w | t, \alpha, \sigma^2) p(\alpha, \sigma^2 | t) \quad (\text{C.7})$$

Et donc, la distribution a posteriori sur les poids est la suivante :

$$p(w | t, \alpha, \sigma^2) = \frac{p(t | w, \alpha, \sigma^2) p(w | \alpha)}{p(t | \alpha, \sigma^2)} \sim \mathcal{N}(w | \mu, \Sigma) \quad (\text{C.8})$$

où la covariance a posteriori et la moyenne s'écrivent :

$$\Sigma = (\sigma^{-2} \phi^T \phi + A)^{-1} \quad (\text{C.9})$$

$$\mu = \sigma^{-2} \Sigma \phi^T t \quad (\text{C.10})$$

avec $A = \text{diag}(\alpha_0, \dots, \alpha_N)$. Notons que σ^2 est également traité comme un hyperparamètre, qui peut être estimé à partir des données.

Ainsi, l'apprentissage statistique consiste à rechercher les plus probables hyperparamètres *posteriors* α_{MP} et σ_{MP}^2 .

Pour un nouvel objet b_* , la prédiction est faite selon l'intégration des poids pour obtenir la vraisemblance maximale des hyperparamètres :

$$p(t | \alpha_{MP}, \sigma_{MP}^2) = \int p(t_* | w, \sigma_{MP}^2) p(w | \alpha_{MP}, \sigma_{MP}^2) dw p(t | \alpha_{MP}, \sigma_{MP}^2) \quad (\text{C.11})$$

$$= \mathcal{N}(t_* | x_*, \sigma_*^2) \quad (\text{C.12})$$

avec :

$$h_* = \mu^T \phi(b_*) \quad (\text{C.13})$$

$$\sigma_*^2 = \sigma_{MP}^2 + \phi(x_*)^T \Sigma \phi(x_*) \quad (\text{C.14})$$

Filtered Shape Averaging

D.1 Introduction

L'augmentation des séries temporelles, et leurs investigations, ont conduit de nombreuses études à développer des mesures de similarités afin de comparer ces signaux. Cet intérêt intervient dans divers domaines d'application tels que le biomédical, l'imagerie, la géoscience, la finance... Avec le développement des bases de données, résumer des séries temporelles est devenu une issue majeure [159]. Les principales applications des moyennes de séries temporelles sont les méthodes de classification et de clustering [160, 161, 162]. Les moyennes de séquences sont particulièrement important, étant inclus dans de nombreux algorithmes. Basé sur le clustering, l'analyse de groupes d'individus, et leur interprétation, amène au développement de méthodes capables de résumer l'information d'un groupe en un unique signal.

Ces algorithmes requièrent l'utilisation d'une métrique de similarité pour comparer les signaux. L'actuelle mesure référence dans ce domaine est la Dynamic Time Warping (DTW). La DTW a été introduite dans les années 1970 pour la reconnaissance vocale et demeure la référence des mesures de similarité entre séquences [163, 164, 165, 166]. Ainsi, les plus récentes études portant sur la moyenne de signaux sont basées sur la DTW [167].

Dans ce travail, nous proposons une nouvelle méthodologie pour l'obtention d'une moyenne de séquences à partir de la DTW. Le principal objectif de cette approche est de parvenir à construire un prototype moyen représentatif des dynamiques des signaux initiaux. Ce prototype est généré artificiellement, un des signaux initiaux ne pouvant représenter l'ensemble de son groupe. Cette méthode est construite pour fonctionner avec des signaux réels, de diverses longueurs. L'objectif est d'obtenir un prototype moyen facilitant l'interprétation du comportement des signaux initiaux.

D.2 Contexte

Dans cette section nous fournissons les définitions de la distance euclidienne et de la DTW. Certaines méthodes proposant la moyenne de signaux sont également présentées.

D.2.1 Distance euclidienne

La distance euclidienne est la manière la plus simple de comparer la similarité entre des signaux. Pour deux signaux S_1 et S_2 , tous les deux de longueur n , cette métrique est définie par :

$$d_{ED}(S_1, S_2) = \sqrt{\sum_{i=1}^n |S_1(i) - S_2(i)|^2} \quad (\text{D.1})$$

La distance d_{ED} présente des inconvénients, elle ne fonctionne par exemple pas pour des signaux de longueurs différentes. De plus, cette métrique ne quantifie pas la similarité entre des signaux, mais effectue une distance point à point. Cette propriété implique une distance ne considérant pas les dynamiques des signaux. Par exemple, la distance euclidienne entre deux séries temporelles identiques, seulement différenciées par une légère translation temporelle, produit une distance élevée, ce qui est en contradiction avec la nature similaire des deux dynamiques observées.

D.2.2 Dynamic Time Warping

Le concept de la DTW a initialement été créé pour obtenir une fonction d'alignement, associée à un coût [144].

L'objectif premier de la DTW est de trouver l'alignement optimal entre des séquences, par distorsion temporelle non linéaire. Depuis son apparition, la DTW est largement employée dans les problématiques de similarité de signaux, pour diverses applications : reconnaissance vocale, analyse de données, imagerie ... [168, 169].

La méthode est basée sur la distance de Levenshtein [170]. L'alignement optimal entre deux signaux détecté par la DTW produit une fonction d'alignement, associée à un coût de transformation. Formellement, pour deux séquences S_1 et S_2 de longueurs respectives n_1 et n_2 , leur distance est calculée suivant l'équation D.2.

$$D(S_1, S_2) = d(S_1(n_1), S_2(n_2)) + \min \begin{cases} D(S_1(1...n_1 - 1), S_2(1...n_2 - 1)) \\ D(S_1(1...n_1), S_2(1...n_2 - 1)) \\ D(S_1(1...n_1 - 1), S_2(1...n_2)) \end{cases} \quad (\text{D.2})$$

avec $D(\emptyset, \emptyset) = 0$, $D(S_{i...n_1}, \emptyset) = D(\emptyset, R_{j...n_2}) = \infty$. Notons que toutes les métriques peuvent être employées pour définir $d(s_i, r_i)$, y compris la norme L_1 : $d(s_i, r_i) = |s_i - r_i|$, et la norme L_2 : $d(S_1(i), S_2(i)) = \sqrt{(S_1(i) - S_2(i))^2}$. La distance $d(S_1(i), S_2(i))$ est ici considérée comme étant la norme L_2 .

Cet algorithme récursif génère une complexité importante, ainsi qu'un grand besoin en espace de mémoire. Ce temps de calcul élevé est une limitation majeure de la DTW. Partant de cette remarque, des études ont essayé de réduire cette complexité initialement de $O(|S_1| \cdot |S_2|) = O(n_1 \cdot n_2)$ [171, 172, 173].

Un exemple de la procédure d'alignement par DTW est donné en figure D.1a. Le résultat numérique de la distance associée à la DTW est le poids cumulé associé au chemin optimal. La définition de chacun des poids est basée sur la distance entre deux coordonnées, généralement définie comme la distance euclidienne entre deux points [174]. Le chemin optimal illustre les avantages de la DTW pour la comparaison de dynamiques (figure D.1b).

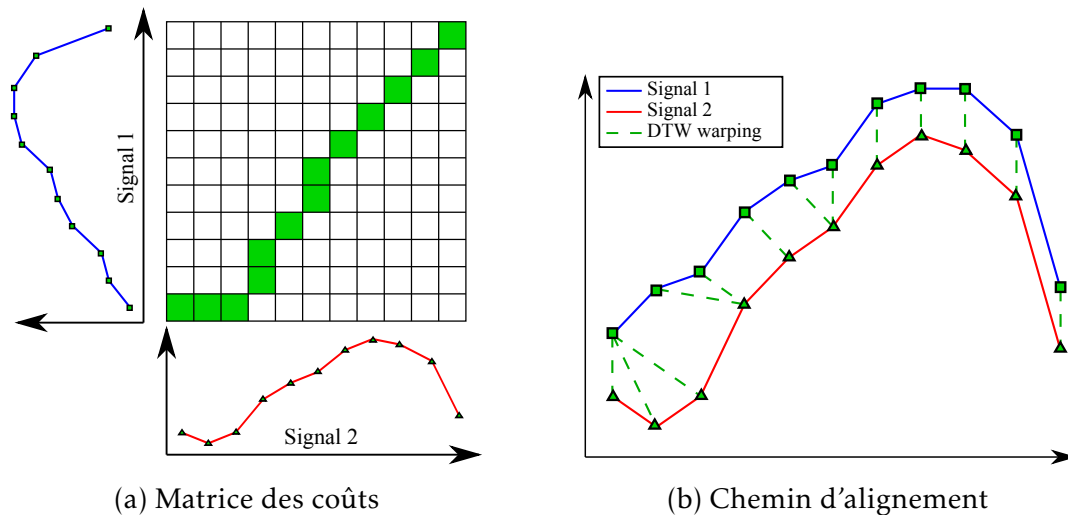


FIGURE D.1 – Définition de la Dynamic Time Warping

D.2.3 Méthodes de moyenne

Le but des méthodes de moyenne de séquences est de produire un nouveau prototype supposé être représentatif du comportement des signaux étudiés. Cependant, le concept de moyenne représentative est sujette à discussion car il existe de nombreux critères de comparaison selon l'objectif désiré. Il est néanmoins intéressant d'observer le fait que la majorité des méthodes utilisent la mesure de similarité DTW.

Les études actuelles se focalisent sur la création d'une moyenne de séquences, suivant un critère de minimisation de la variance des distances DTW entre le prototype moyen et les signaux initiaux [167, 175, 176]. La réduction de ce critère est guidé par les objectifs de performances de classification. Les principales méthodes fournissant une moyenne de séquences sont présentées ici.

Gupta et al. [176] propose une méthode nommée Nonlinear Alignment and Averaging Filters (NLAAF), considérant un schéma heuristique basé sur les dynamiques moyennes [177]. Cet algorithme forme un signal moyen à partir des coordonnées des associations de deux signaux par DTW. Deux versions de cette méthode existent NLAAF₁ et NLAAF₂, suivant différents schémas d'itérations, permettant d'éviter la nécessité de prise en compte d'un nombre de signaux étant d'une puissance de deux. Cette approche est largement critiquée car elle génère un prototype moyen de longueur plus grande que les signaux originels, impliquant une haute complexité. De plus, il s'agit d'une méthode non déterministe [167].

Une autre méthode, appelée Prioritized Shape Averaging (PSA), a été introduit afin de solutionner les désavantages évoqués de l'algorithme NLAAF [178, 179]. La méthode PSA utilise un schéma ascendant hiérarchique, combiné avec un procédé de moyenne effectué par sélection d'un élément par composante connectée. Même si cette méthode réduit la propagation d'erreurs grâce à son approche hiérarchique, elle ne solutionne pas l'augmentation de la longueur du prototype moyen. La PSA connaît donc également une augmentation importante de sa complexité au fil des itérations nécessaires.

Des études plus récentes de Petitjean et al. [167] ont développé l'algorithme DTW Barycenter Averaging (DBA). Cet algorithme considère une stratégie heuristique con-

struite comme une méthode de moyenne globale. Ainsi, la DBA ne nécessite pas d'ordonnement des signaux et préserve une propriété déterministe. Les autres avantages de la DBA sont sa conservation d'une longueur stable du prototype moyen, sa faible complexité et sa capacité à traiter des signaux de longueurs différentes.

D.2.4 Motivation

Les méthodes de moyenne basée sur la moyenne de couple de signaux sont sensibles à l'ordre initial des signaux, car la DTW n'est pas associative. La NLAFF et la PSA ne garantissent donc pas la génération de résultats identiques en fonction de l'ordre initial des signaux. Ce point établit l'avantage principal de la DBA, basée sur une approche globale de moyenne. Il faut noter que toutes ces méthodes ne sont que rarement testées sur des données réelles, même si cela est considéré comme techniquement possible avec la DBA [180].

Le désavantage des méthodes existantes réside dans leur mode de construction. Leur structure est guidée par le critère de minimisation de la distance DTW intra-group [175]. Avec ce critère, les méthodes actuelles ne considèrent pas les dynamiques réelles des signaux initiaux mais tentent d'obtenir les meilleurs résultats possibles sur les problèmes de classification. Ces méthodes sont donc sujettes au surapprentissage car elles ne considèrent pas de terme de pénalisation. Cette inconvénient est important pour diverses applications qui requièrent une interprétation du prototype moyen afin d'analyser le comportement de signaux, telles que les approches de diagnostic ou de maintenance. Ce travail cherche à répondre à ce besoin par la méthode Filtered Shape Averaging (FSA), ayant pour objectif de proposer un prototype moyen offrant des dynamiques similaires aux signaux initiaux et donc permettant son interprétation.

D.3 Méthode proposée

Cette section détaille la méthodologie du Filtered Shape Averaging (FSA). La méthode est divisible en deux parties distinctes : une moyenne faite à partir de la DTW, puis l'application d'une fonction de lissage Gaussien.

D.3.1 Moyenne basée sur la Dynamic Time Warping

Cette méthode permet de moyenniser deux signaux, basée sur la DTW, suivant l'idée d'utiliser le chemin optimal entre ces deux signaux. Intuitivement, l'alignement optimal de la DTW propose une dynamique moyenne, construite à partir des deux séquences. Ce signal moyen est composé d'un ensemble de coordonnées issus de chaque association du chemin optimal (figure D.2). Cette approche propose un bon compromis car elle respecte les dynamiques initiales des signaux et car elle minimise la distance moyenne entre le prototype et les signaux considérés.

Une telle stratégie ne peut cependant fonctionner que pour deux séquences. L'extension naturelle pour un large nombre de signaux poursuit cette construction avec un processus itératif. À chaque itération, l'idée est de diviser le nombre de séquences traitées par deux, jusqu'à obtenir un unique prototype moyen. Pour L signaux, cet algorithme consiste à faire L moyennes successivement. Ce procédé itératif correspond

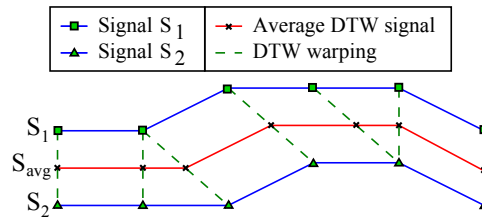


FIGURE D.2 – Moyenne de deux séquences, un coordonnée pour chaque association

au $NLAAF_1$ [176] et fonctionne idéalement avec un nombre de séquences étant une puissance de deux.

Le but est alors d’adapter cette stratégie pour le cas général. Un algorithme séquentiel consiste à traiter uniquement un nombre pair de signaux à chaque étape, ajoutant un poids à chaque signal moyen. Ces poids étant définis comme le nombre de signaux résumés dans ce prototype moyen (figure D.3).

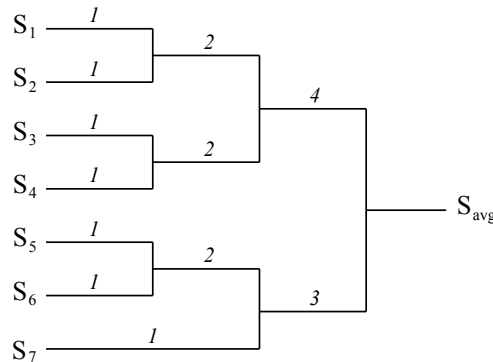


FIGURE D.3 – Schéma d’un processus itératif de moyenne - Le poids associé à chaque signal est défini sur chaque branche

Le prototype moyen entre deux séquences est composé un ensemble de coordonnées pour chaque association du chemin optimal. Les coordonnées ne sont pas situés au milieu de chaque association mais correspondent au barycentre de ces associations, considérant les poids respectifs des signaux moyennés. Cette approche réduit la dépendance vis à vis de l’ordre initial des signaux, ce qui évite des propagations d’erreurs possiblement induits par de signaux atypiques.

La longueur du prototype moyen S_{avg} de deux séries temporeles S_1 et S_2 , de longueurs respectives n_1 et n_2 , est minorée par $\min(n_1, n_2)$ et majorée par $n_1 + n_2 - 2$:

$$\min(n_1, n_2) \leq |S_{avg}| \leq n_1 + n_2 - 2 \tag{D.3}$$

Dans le processus complet, la propriété D.3 implique une augmentation de la complexité, car la longueur du signal moyen augmente considérablement au fil des itérations. Ainsi, pour permettre l’usage de la méthode développée dans un contexte de données réelles, nous proposons une méthode d’échantillonnage. Cette étape permet à la FSA de considérer des signaux synchrones ou asynchrones.

D.3.2 Fonction de lissage par noyau Gaussien

Pour réduire la complexité de l'algorithme proposé, nous introduisons l'idée d'appliquer un lissage par noyau Gaussien à chaque prototype moyen. Cela réduit également la présence du bruit dans ces prototypes. Pour éviter l'augmentation de la longueur des prototypes moyens, la longueur du signal moyen lissé est égale au maximum des longueurs des deux signaux traités. Ce choix arbitraire permet de majorer la longueur du prototype final par la longueur maximale des signaux initiaux. Ainsi, pour deux signaux S_1 et S_2 de longueurs n_1 et n_2 , la FSA produit un prototype moyen S_{avg} de longueur n_{avg} égale à $\max(n_1, n_2)$:

$$n_{avg} = |S_{avg}| = \max(n_1, n_2) \quad (D.4)$$

Le point critique est alors de définir les n_{avg} abscisses des coordonnées X^{avg} du signal moyen lissé. Pour respecter les dynamiques des signaux initiaux et pour conserver leurs propriétés, la méthodologie consiste en une sélection au sein de la combinaison des coordonnées $\{X_1, X_2\}$ des deux signaux initiaux S_1 et S_2 . Cette sélection s'effectue par un processus basé sur les percentiles. Ainsi, n_{avg} valeurs sont sélectionnées parmi $\{X_1, X_2\}$, afin de générer les coordonnées des abscisses du prototype moyen.

Formellement, pour un signal $S = \{(x_1, y_1), \dots, (x_n, y_n)\}$, résultant d'une moyenne basée sur la DTW explicitée en section D.3.1, son signal régularisé est défini par la fonction \hat{f} présentée en équation D.5 :

$$\hat{f}(x_0) = \frac{\sum_{i=1}^n K_\lambda(x_0, x_i) y_i}{\sum_{i=1}^n K_\lambda(x_0, x_i)} \quad (D.5)$$

où K_λ est la noyau Gaussien :

$$K_\lambda(x_0, x) = \frac{1}{\lambda} \exp\left(-\frac{\|x - x_0\|^2}{2\lambda}\right) \quad (D.6)$$

Le paramètre λ contrôle la largeur du voisinage, il est donc primordial de le définir pour obtenir une bonne régularisation. Dans ce travail nous utilisons le paramètre λ_{SROT} (Silverman's rule of thumb), décrit par l'équation D.7 [157, 158], avec IQR l'espace interquartile et n la longueur du signal S . Le noyau Gaussien est choisi ici pour sa capacité à produire une régularisation performante avec un paramètre déterministe. Il peut être envisagé de considérer d'autres noyaux en fonction de l'application visée.

$$\lambda_{SROT} = \left(\frac{4}{3}\right)^{1/5} \min\left(\sqrt{\sigma}, \frac{IQR(S)}{1.349}\right) \cdot n^{-1/5} \quad (D.7)$$

La figure D.4 illustre la réduction de la longueur par FSA, ainsi que son procédé de lissage permettant de réduire la propagation des erreurs.

D.3.3 Complexité

L'étude de la complexité est important car les méthodes actuelles permettant la moyenne de signaux possèdent des temps de calcul élevés. La complexité de la DTW, pour deux signaux S_1 et S_2 , respectivement de longueur n_1 et n_2 , est en $O(n_1 \cdot n_2)$. Pour détailler la complexité de la méthodologie présentée ici, nous considérons L séries

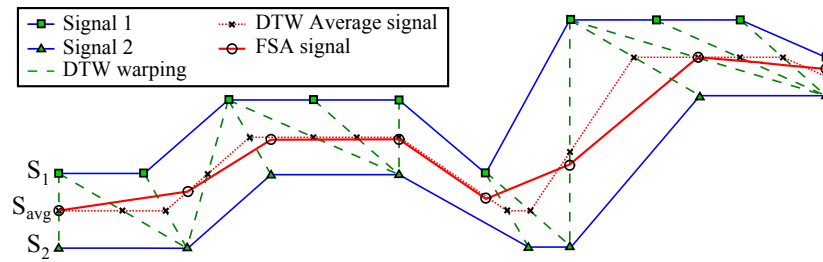


FIGURE D.4 – Moyenne de deux signaux, avec lissage Gaussien

temporelles de longueur inférieures ou égales à n , chaque itération est divisée en deux parties jusqu'à ce que $L = 1$:

1. Remplacer les L séquences par leurs $L/2$ moyennes.
2. Utiliser une fonction Gaussienne pour lisser les signaux moyens et réduire leurs longueurs.

De ce fait, la complexité algorithmique est séparée en deux parties distinctes :

- Le calcul de la moyenne de deux séquences utilise uniquement l'algorithme DTW. Chaque moyenne de deux séquences nécessite au plus $O(n^2)$.
- Le lissage par noyau Gaussien d'un signal requiert la sélection de n coordonnées d'abscisses suivant un processus basé sur les percentiles, et le calcul du paramètre λ_{SROT} . Ces deux calculs nécessitent au plus $O(n \log n)$ opérations. Ensuite, l'équation D.5 est appliquée pour obtenir un signal moyen lissé. Cette étape a une complexité de $O(n^2)$. L'étape de lissage Gaussien a donc une complexité globale de $O(n^2)$.

Comme nous l'avons détaillé, l'algorithme global effectue environ L moyennes pour résumer L signaux. La complexité totale de l'algorithme FSA est alors de $O(L.(n^2+n^2)) = O(L.n^2)$.

La table D.1 affiche un comparatif des complexités des principales méthodes de moyenne de signaux. Notons I le nombre d'itérations de l'algorithme DBA. Ce résultat illustre l'avantage en terme de temps de calcul de la méthodologie proposée.

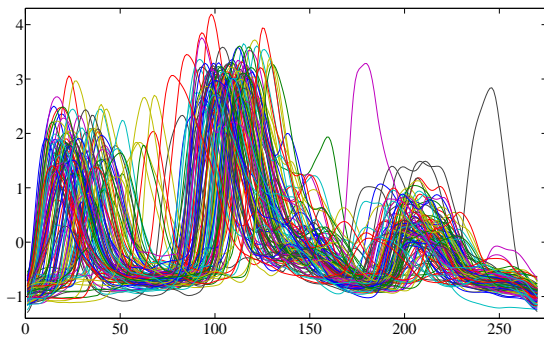
	NLAAF	PSA	DBA	FSA
Complexité	$O(L.n^3)$	$O(L^2.n^2 + L.n^3)$	$O(I.L.n^2)$	$O(L.n^2)$

TABLE D.1 – Comparatif de la complexité des principaux algorithmes fournissant une moyenne de signaux

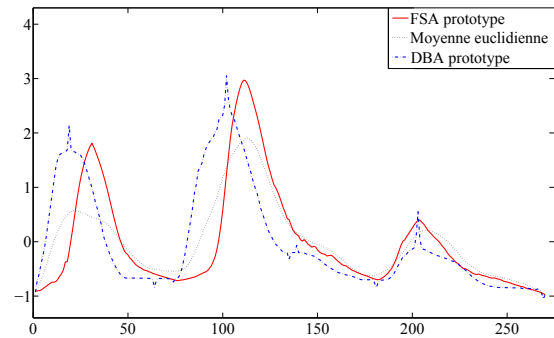
D.4 Expérimentations sur la base de données UCR

Cette section présente les résultats obtenus avec la méthode FSA sur la base de données *UCR classification/clustering* [181]. Ces données sont fréquemment utilisées pour comparer les performances des méthodes, spécialement dans les problématiques de moyenne de signaux.

La figure D.5 et la figure D.6 proposent un comparatif des prototypes moyens d'un ensemble de signaux, obtenus par moyenne euclidienne, par DBA et par FSA. Les résultats obtenus démontrent la bonne représentation visuelle du prototype construit par FSA, par rapport aux signaux initiaux. Ceci n'est pas forcément le cas des autres méthodes, car établies pour des problématiques de classification. L'interprétation du prototype moyen est un point crucial pour la précision de possibles applications (physique, chimique, santé...).

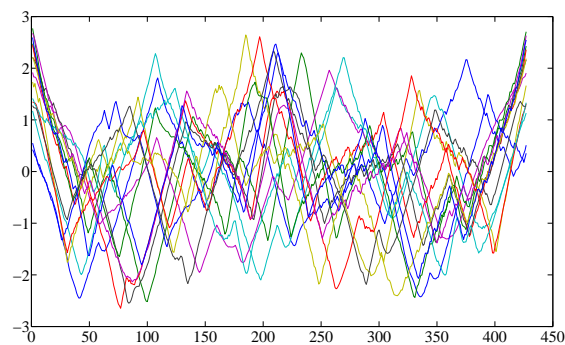


(a) Une classe des données 50words

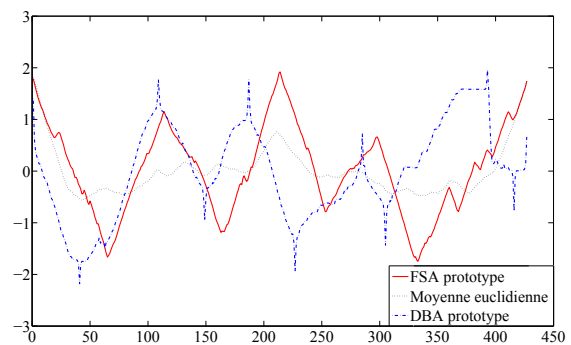


(b) Les séquences moyennes de cette classe

FIGURE D.5 – Exemple de résultat de la méthode FSA sur une classe des données 50words, comparativement à la moyenne euclidienne et au prototype moyen obtenu par DBA



(a) Une classe des données OSULeaf



(b) Les séquences moyennes de cette classe

FIGURE D.6 – Exemple de résultat de la méthode FSA sur une classe des données OSU-Leaf, comparativement à la moyenne euclidienne et au prototype moyen obtenu par DBA