

UNIVERSITÉ DU QUÉBEC

THÈSE PRÉSENTÉE À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DU DOCTORAT EN GÉNIE ÉLECTRIQUE

PAR

NILSON HENAO

DÉSAGRÉGATION NON-INTRUSIVE DE CHARGES
ÉLECTRIQUES RÉSIDENTIELLES

SEPTEMBRE 2018

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

DOCTORAT EN GÉNIE ÉLECTRIQUE (PH.D.)

Programme offert par l'Université du Québec à Trois-Rivières

DÉSAGRÉGATION NON-INTRUSIVE DE CHARGES ÉLECTRIQUES RÉSIDENIELLES

PAR

NILSON HENAO

Dr. Kodjo AGBOSSOU, directeur de recherche Université du Québec à Trois-Rivières

Dr. François NOUGAROU, président du jury Université du Québec à Trois-Rivières

Dr. Souso KELOUWANI, codirecteur de recherche Université du Québec à Trois-Rivières

Dr. Roland MALHAME, évaluateur externe École Polytechnique de Montréal

Dr. Brayima DAKYO, évaluateur externe Université Le Havre Normandie

Thèse soutenue le 20 juin 2018

Résumé

La participation active du client résidentiel au réseau intelligent par le biais d'un gestionnaire d'énergie peut tirer avantage de plusieurs techniques, dont la désagrégation non intrusive des charges. En effet, connaître les types de charges en opération et leur consommation énergétique est essentiel pour (i) prévoir la consommation globale de la résidence ; (ii) gérer efficacement les modes d'opérations de ces appareils et (iii) optimiser les coûts énergétiques. Si l'approche de sous-mesurage procure des avantages au niveau de la précision, son caractère intrusif (matérialisé par l'installation de plusieurs capteurs) représente une contrainte limitant son application dans le domaine résidentiel. C'est la raison pour laquelle plusieurs techniques d'identification non intrusive des charges et très peu supervisées sont en cours de développement et les premiers résultats font l'objet de diffusion scientifique.

Les profils des charges électriques résidentielles peuvent varier considérablement en fonction de facteurs environnementaux, des habitudes de l'utilisateur et d'autres paramètres non déterministes. Dans ce contexte, une nouvelle approche de séparation de profils de charges basée sur des modèles stochastiques est envisagée. Celle-ci considère les particularités des usages typiques au Canada, où les charges thermostatiques telles que le chauffage d'espace habitable et le chauffage de l'eau sanitaire prennent un grand pourcentage de l'énergie électrique consommée pendant les périodes climatiques froides. Ces usages sont normalement composés par des contrôleurs thermostatiques ainsi que d'un nombre fini d'éléments résistifs.

Afin de détecter les charges couramment connues dans le domaine, la littérature scientifique propose une grande variété de techniques basées sur l'analyse des caractéristiques

microscopiques. Toutefois, la détection basée sur l'analyse de signatures microscopiques est peu applicable dans la désagrégation de charges résistives à cause de leur linéarité. En revanche, les profils de puissance active de ces charges sont fondés sur des structures temporelles pouvant être caractérisées par l'analyse séquentielle et événementielle à l'échelle des cycles d'opération. Dans cette perspective et à l'issue de l'analyse bayésienne, les modèles markoviens ont émergé en tant qu'outil capable de suivre le comportement des automates stochastiques en modélisant la structure temporelle des données. Néanmoins, des difficultés liées aux fréquences thermostatiques de commutation relativement élevées, ainsi qu'aux similarités entre les puissances des éléments individuels génèrent un scénario complexe d'application.

Cette thèse propose le développement d'une approche probabiliste parallèle capable de gérer les similarités et le chevauchement d'événements entre les éléments qui composent la courbe de charge agrégée. En effet, les durées des états thermostatiques ont été considérées dans le calcul de probabilités de transition afin d'améliorer le pouvoir discriminant des modèles markoviens. Cette approche combine de façon cohérente les modèles individuels en construisant un modèle à états finis combinés facilitant la surveillance parallèle et optimale des charges d'intérêt. L'analyse des résultats démontre l'applicabilité et la robustesse de la méthode proposée pour la détection des charges cibles dans des configurations complexes et incertaines du profil agrégé. De plus, une analyse comparative avec des approches avancées de l'état de l'art a été réalisée.

Remerciements

Je tiens à remercier sincèrement mon directeur de recherche Monsieur Kodjo Agbossou, professeur du département de génie électrique et génie informatique de l'UQTR et directeur du Laboratoire d'Innovation et de Recherche en Énergie Intelligente (LIREI), pour m'avoir confié ce projet, pour la qualité de son encadrement, pour ses conseils, son soutien et ses encouragements. J'aimerais remercier aussi mon codirecteur de recherche Monsieur Souso Kélouwani, professeur du département de génie mécanique de l'UQTR, pour son aide, son appui scientifique, pour ses précieuses contributions et sa disponibilité tout au long de ce travail de thèse.

J'exprime ma gratitude à Monsieur Alben Cárdenas, agent de recherche du LIREI, pour l'appui technique et sa participation dans l'élaboration du projet de recherche. Je remercie également Michael Fournier, chercheur du Laboratoire de Technologies de l'Énergie (LTE) d'Hydro-Québec, pour sa participation scientifique, ainsi que pour le temps qu'il a consacré à nos échanges, réunions et révisions de documents relatifs au travail de recherche.

Ce travail n'aurait pas été possible sans l'aide de différentes instituts et organismes qui, au travers de leur soutien financier ou technique, ont reconnu mon travail et m'ont fait confiance : le LTE d'Hydro-Québec, le Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG), l'Intitut de Recherche sur l'Hydrogène (IRH) et la Fondation de l'UQTR.

Table des matières

Résumé	i
Remerciements	iii
Table des matières	iv
Table des figures	vii
Liste des tableaux	x
Liste des symboles	x
Liste des acronymes	xii
Chapitre 1 - Introduction	1
1.1 Contexte général et motivation	1
1.2 Bénéfices de la surveillance de charges dans le secteur résidentiel	3
1.3 Problématique de thèse	6
1.4 Objectifs et contributions	9
1.5 Méthodologie	10
1.6 Plan du manuscrit	12
Chapitre 2 - NILM - État de l'art	14
2.1 Généralités sur la surveillance non-intrusive	14
2.1.1 Catégorisation des charges	16
2.1.2 Type de caractéristiques	18
2.2 Désagrégation de mélanges additifs	19
2.2.1 Apprentissage automatique	22

2.2.2	Optimisation combinatoire	23
2.2.3	Analyse séquentielle déterministe	24
2.2.4	Analyse séquentielle probabiliste	25
2.3	Approches orientées aux charges thermostatiques	28
2.4	Synthèse des revues et critères pour la sélection du modèle	29
Chapitre 3 - Modélisation de multiples éléments thermostatiques		32
3.1	Introduction	32
3.2	Modèle probabiliste à deux états	34
3.2.1	Les états cachés et les observations associées	35
3.2.2	Analyse probabiliste de la dynamique des états	37
3.2.3	Modèle de transition incluant la durée du séjour	43
3.2.3.1	Le temps de séjour et l'analyse de survie conditionnelle	45
3.2.3.2	Risque instantané et probabilité de transition	46
3.2.4	Reconnaissance probabiliste des profils à deux états	50
3.2.4.1	Maximisation du logarithme de la vraisemblance	53
3.2.4.2	Programmation dynamique pour l'inférence des états	54
3.2.4.3	Étiquetage de séquences	57
3.2.4.4	Probabilités stationnaires des états thermostatiques	58
3.2.4.5	Puissance et énergie moyenne d'un profil thermostatique	60
3.3	Agrégation de modèles et détection parallèle des états	61
3.3.1	Transitions entre super-états	63
3.3.2	Émissions agrégées	64
3.3.3	Pré-allocation des probabilités de transition avec les opérations de Kronecker	67
3.3.4	Algorithme de Viterbi	72
3.4	Synthèse du chapitre	73
Chapitre 4 - Évaluation de la méthode		75
4.1	Cadre comparatif et scénarios d'évaluation	75
4.2	Apprentissage supervisé des approches NILM étudiées	76
4.2.1	Estimation des états à partir des profils individuels	77

4.2.2	Entraînement des modèles à durée explicite	77
4.2.3	Entraînement des modèles FHMM	79
4.2.4	Entraînement des modèles RNN	80
4.3	Métriques et indicateurs de performance	81
4.3.1	Métriques pour la détection de l'état	82
4.3.2	Métriques pour la reconstruction du profil	83
4.4	Résultats préliminaires	83
4.4.1	Désagrégation des charges conventionnelles à deux états	84
4.4.2	Effet des charges inconnues	88
4.5	Désagrégation de charges thermostatiques actives à des puissances similaires	89
4.6	Cas des charges de chauffage contrôlées par des thermostats électroniques	91
4.6.1	Modèle MLI avec rapport cyclique stochastique	92
4.6.2	Initialisation et mise à jour des paramètres	94
4.6.3	Résultats	95
4.7	Discussion sur les résultats de la validation	98
Chapitre 5 - Conclusions et recommandations		101
5.1	Synthèse des travaux	101
5.2	Limitations et difficultés	104
5.2.1	Complexité de la méthode factorielle	104
5.2.2	Maximums locaux dans l'apprentissage de Viterbi	105
5.2.3	L'hypothèse d'indépendance entre les éléments	105
5.3	Recommandations	106
5.3.1	Lois paramétriques alternatives pour la durée de l'état	106
5.3.2	Distributions priori sur les paramètres de la loi de Weibull	106
5.3.3	Estimation non supervisée des paramètres thermostatiques	107
5.3.4	Profit des ressources NILM disponibles	108
Bibliographie		111

Table des figures

1-1	Flux de consommation de l'énergie primaire au Canada au 2015. Source . . .	2
1-2	Le suivi de la consommation en temps réel et de l'économie de l'énergie dans le secteur résidentiel. Diagramme adapté à partir de [5].	4
1-3	Taux de pénétration des charges électrothermiques dans les régions Canadiennes [13]	6
1-4	Méthodologie de recherche suivie.	11
2-1	Structure générale d'un système NILM avec les possibles types de configuration.	15
2-2	Carte de signatures intemporelles utilisées pour les transitions des automates déterministes. Source : [60].	25
2-3	Réseau bayésien d'un FHMM [37].	27
3-1	Schéma du processus de surveillance proposé.	33
3-2	Profil de puissance d'un convecteur électrique	38
3-3	Profil et distributions des puissances associées aux états d'un convecteur électrique contrôlé par un thermostat électronique.	38
3-4	Réseaux bayésiens des processus markovien, semi-markovien et à durée explicite d'une séquence discrète à deux états.	42
3-5	Profil de puissance du réfrigérateur	49
3-6	Distributions sur la durée et la fonction de hasard de l'état "ON" du profil thermostatique de la Figure 3-5	50

3-7	Réseau bayésien proposé pour le HMM à durée explicite.	51
3-8	Log-vraisemblance des profils individuels de la puissance de deux charges thermostatiques.	59
3-9	Profils individuels de la puissance de quatre convecteurs électriques contrôlés par de thermostats électroniques.	65
3-10	Profil FSM résultant de l'agrégation des quatre profils thermostatiques affichés à la Figure 3-9.	65
3-11	Distributions conditionnelles de la puissance agrégée de quatre éléments thermostatiques.	67
3-12	Résultats du produit de Kronecker entre quatre distributions de Bernoulli . . .	69
4-1	Schéma du processus d'entraînement du RNN appliqué à la désagrégation supervisée des charges thermostatiques.	81
4-2	Profil agrégé des cinq charges thermostatiques.	85
4-3	Indicateur MCC pour la détection des cinq charges thermostatiques de la base de données ECO.	86
4-4	Puissance du congélateur désagrégée par la méthode RNN.	87
4-5	Puissance du congélateur désagrégée par la méthode EDHMM.	87
4-6	Indicateur MCC pour la détection des deux charges thermostatiques en pré- sence de charges inconnues.	89
4-7	Système de mesure de la puissance pour des charges thermostatiques résistives.	90
4-8	Distribution de Weibull pour la durée des états thermostatiques des convecteurs électriques.	91
4-9	Profil d'état ON-OFF d'un thermostat électronique.	93
4-10	Distribution de Weibull initiale pour les durées des états thermostatiques. . .	95
4-11	Structure du système de désagrégation non supervisée des profils thermosta- tiques.	96
4-12	Estimation en ligne de la distribution sur la durée de l'état thermostatique . . .	97
4-13	Profil reconstruit des quatre profils thermostatiques de la Figure 4-14.	97

4-14	Profils individuels reconstruits de quatre convecteurs électriques contrôlés par de thermostats électroniques.	98
5-1	STFFT de la puissance agrégée contenant multiple charges thermostatiques. .	108
5-2	Analyse de la densité des transitions en puissance de charges thermostatiques.	108

Liste des tableaux

2-1	Type de charges électriques résidentielles selon la morphologie de leurs profils de puissance.	17
3-1	Distributions utilisées pour modéliser les durées des états	49
4-1	Paramètres associés à l'état "ON" des charges de la base ECO.	85
4-2	Indicateur MCC pour la désagrégation charges thermostatiques actives à des puissances similaires.	91
4-3	Indicateur PRA pour la désagrégation charges thermostatiques actives à des puissances similaires.	91

Liste des symboles

Terme	Description
A	Matrice de transition.
δ	Delta de Kronecker.
Δt	Temps d'échantillonnage.
d_k	Durée de l'état à l'instant k .
\mathbf{d}	Vecteur de durées.
E	Énergie.
K	Longueur d'une séquence.
k	Indice du temps discret.
M	Nombre d'instances (modèles ou charges).
\mathcal{M}	Modèle d'une charge.
m	Indice des instances (modèles ou charges).
ϕ_s	Paramètres de la distribution de la durée de l'état.
π_s	Probabilité stationnaire de l'état s .
ψ_k^s	Probabilité de l'état s à l'instant k .
S	Ensemble discret d'états.
s	État $s \in S$.
τ_s	Durée moyenne de l'état s .
τ	Durée moyenne d'un cycle thermostatique.
θ_s	Paramètres pour la distribution sur la puissance.

Terme	Description
μ	Puissance moyenne.
\mathbf{u}	Vecteurs des puissance moyennes associées aux états d'une charge.
v_s	Variance de la puissance dans l'état s .
X	Variable aléatoire représentant la puissance de la charge.
x_k	Puissance individuelle.
$\bar{\mathbf{x}}$	Séquence de mesures individuelles.
Y	Variable aléatoire représentant la puissance agrégée.
y_k	Signal agrégé.
$\bar{\mathbf{y}}$	Séquence de mesures agrégées.
Δy_k	Différences consécutives de la puissance active.
Z	Variable aléatoire représentant l'état discret de la charge.
z_k	Variable d'état discret au temps k .
\mathbf{z}_k	Vecteur des variables d'état.

Liste des acronymes

Terme	Description
ANN	Artificial Neural Network.
DHMM	HMM à Différences.
ECO	Electricity Consumption and Occupancy.
ECS	Eau Chaude Sanitaire.
EDHMM	HMM à Durée Explicite.
EM	Expectation-maximization.
FHMM	HMM Factoriel.
FSM	Finit State Machine.
GPRA	PRA Global.
HMM	Hidden Markov Model.
HSMM	Hidden Semi-Markov Model.
IMA	Infrastructure de Mesurage Avancé.
MAP	Maximum a posteriori.
MC	Monte-Carlo.
MCC	Coefficient de Corrélation de Matthews.
MCMC	Monte-Carlo par Chaînes de Markov.
ML	Machine Learning.
MLE	Maximum of likelihood.
MLI	Modulation à Largeur d'Impulsion.

Terme	Description
NILM	Non-Intrusive Load Monitoring.
PRA	Précision de la Reconstruction en Puissance.
REI	Réseaux Électriques Intelligents.
RNCan	Ressources Naturelles Canada.
RNN	Réseau de Neurones Récurrents.

Chapitre 1 - Introduction

1.1 Contexte général et motivation

L'approvisionnement énergétique et l'atténuation des problèmes environnementaux font partie des défis connexes majeurs auxquels la planète est confrontée aujourd'hui. Conséquemment, des politiques prioritaires doivent être concrétisées afin de combler la demande croissante d'énergie et en même temps atteindre les objectifs de réduction de l'empreinte écologique. À cet égard, la mise en œuvre des stratégies promouvant une consommation rationnelle et intelligente de l'énergie, les efforts pour l'amélioration de l'efficacité des systèmes énergétiques actuels ainsi que l'intégration massive des sources non polluantes font partie des actions valides visant le développement énergétique durable [1].

Compte tenu des tendances du développement technologique actuel, il est quasiment certain que la substitution des sources d'énergies fossiles par des sources à faible impact écologique aboutit à un processus d'électrification des différents secteurs de consommation. Dans cette perspective, la production non polluante et la gestion efficace du stockage, du transport et de la distribution de l'électricité deviendront des activités stratégiques dans le marché énergétique mondial. Par ailleurs, dans le contexte canadien, en tant que pays grand producteur (cinquième) et grand consommateur (septième), les enjeux environnementaux et économiques de l'électrification sont considérables à cause de la forte demande actuelle des ressources fossiles. Cette situation est illustrée dans le diagramme de la Figure 1-1. Il est montré que plus de 70% de l'approvisionnement en énergie primaire provient des sources polluantes. Conséquemment, pour respecter les engagements dans la réduction de l'empreinte carbone, un processus d'électrification doit être planifié et soutenu par une augmentation de la production d'électricité à partir des sources énergétiques propres.

En regardant la consommation finale par secteur, le résidentiel représente 17% de l'énergie

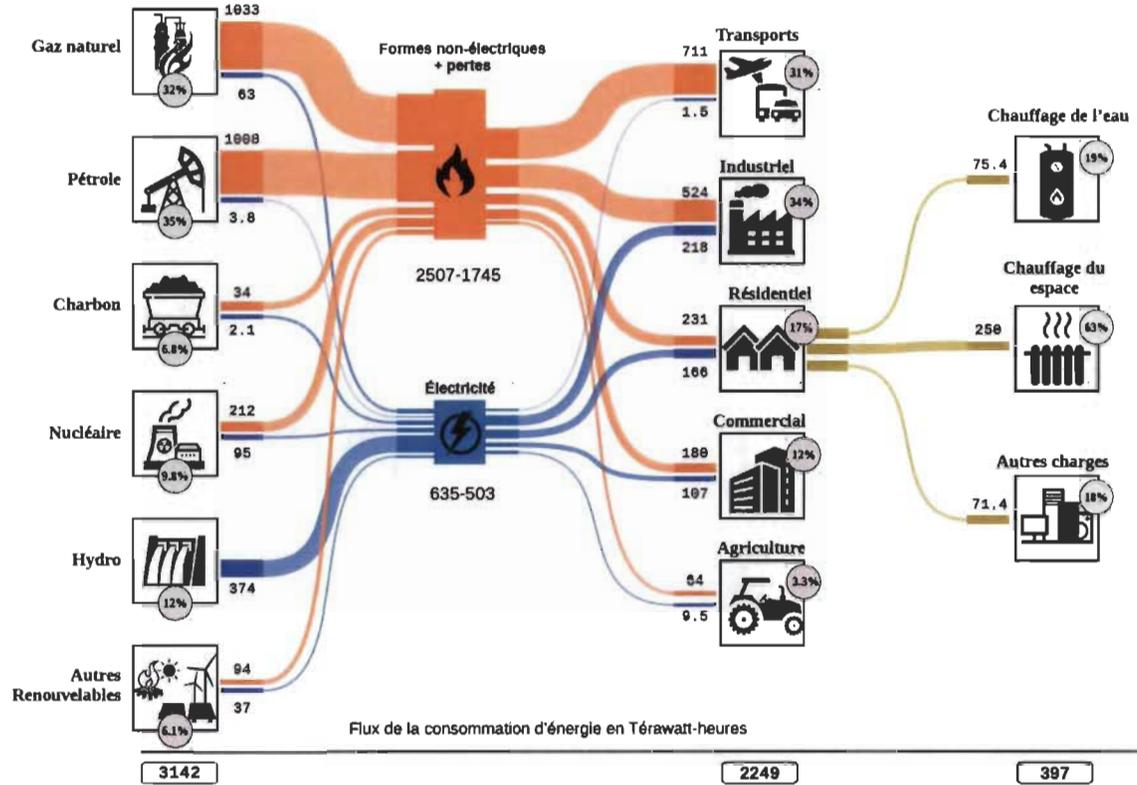


FIGURE 1-1 Flux de consommation de l'énergie primaire au Canada en 2015. Source : Agence internationale de l'énergie et Ressources naturelles du Canada.

primaire et 31% de la demande d'électricité en tant qu'énergie secondaire. Cela est dû notamment aux conditions climatiques spéciales entraînant une forte consommation pour des usages tels que le chauffage des pièces ainsi que le chauffage de l'eau sanitaire. En effet, en 2014, ces deux usages représentaient respectivement 64% et 19% de l'énergie totale demandée par les bâtiments résidentiels au Canada [2]. Cette situation combinée à la croissance constante du coût et de l'énergie¹ requiert des actions spécifiques en faveur de la réduction significative de la consommation et à la maximisation de l'efficacité énergétique des systèmes thermiques résidentiels.

En outre, les initiatives priorisant les innovations technologiques dans la gestion de l'énergie sont de plus en plus bien encadrées dans le concept des Réseaux Électriques Intelligents (REI) [4]. Les REI ciblent le développement de mécanismes de gestion adaptés aux nouvelles

1. Entre 2000 et 2010 le prix de l'électricité a augmenté au Canada jusqu'à 79% tandis que, la consommation d'énergie par habitant a augmenté à 20% [3]

réalités du marché de l'énergie électrique où le consommateur devient un acteur majeur. Justement, un des objectifs du modèle REI consiste à établir un flux d'information bidirectionnel en temps réel afin d'améliorer l'engagement de l'utilisateur dans l'interaction avec le réseau électrique. Dans ce scénario, le développement des outils facilitant la génération des informations pertinentes et de qualité est fondamental. Ces informations vont aider les consommateurs dans la prise de décisions plus avisées. De ce fait, les technologies conçues pour le mesurage détaillé et pour l'analyse automatisée des habitudes de consommation de l'énergie jouent un rôle capital dans la génération de telles informations.

1.2 Bénéfices de la surveillance de charges dans le secteur résidentiel

Des études montrent que les échanges d'information en temps réel entre les fournisseurs et les consommateurs permettent d'économiser entre 5 et 20 % de l'énergie consommée dans le secteur résidentiel [5], [6]. À cet égard, la surveillance de charges résidentielles a été identifiée comme une source d'information précieuse tant pour le gestionnaire du réseau que pour les consommateurs. En effet, les systèmes de surveillance mesurent des variables relatives à la consommation des appareils ou des usages² électriques installés dans un bâtiment. Tel qu'illustré à la Figure 1-2, plusieurs niveaux d'économie sont envisagés en fonction de la granularité et le temps de retour d'information vers le client. Selon la littérature [5], il est attendu que le retour en temps réel d'information sur la consommation par usage permet d'atteindre des économies au-delà de 10%. Dans le contexte résidentiel, le temps réel est associé à une granularité de l'information, soit la fréquence de la mise à jour, qui est de l'ordre de quelques minutes [7]. Il est important de noter que le temps réel est défini par rapport à la rapidité et à la pertinence de l'information dans une stratégie d'auto-gestion de la consommation. Dans celle-ci, les actions sont normalement prises par des personnes sans nécessairement compter sur un mécanisme de décision automatisé.

Les informations recueillies par les systèmes de surveillance peuvent potentiellement servir dans la gestion énergétique adaptative de la demande et au développement des stratégies

2. L'usage électrique est une catégorisation des charges électriques selon leur utilisation, par exemple le froid alimentaire, le chauffage électrique de l'espace habitable et le chauffage de l'eau sanitaire.

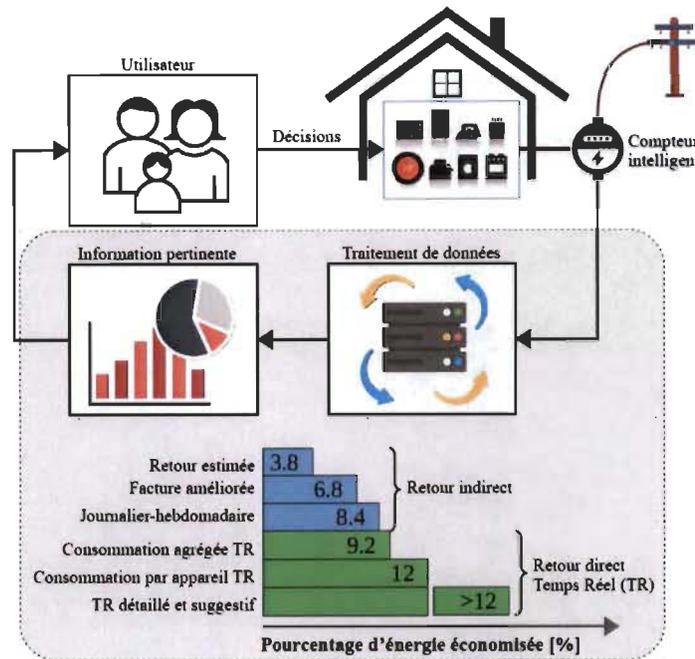


FIGURE 1-2 Le suivi de la consommation en temps réel et de l'économie de l'énergie dans le secteur résidentiel. Diagramme adapté à partir de [5].

pour réduire les pics de consommation [8], [9]. Toutefois, du côté résidentiel, ces dernières peuvent aussi offrir de nombreux bénéfices tels que :

- Obtenir un rapport détaillé sur la demande d'énergie électrique de la résidence permettant l'identification des facteurs causant des consommations élevées. Sur ce point, les analyses statistiques sur des profils individuels pourront servir au client à déterminer les moments et les activités qui sont à la base de certaines surconsommations.
- Détecter des anomalies dans l'opération des charges électrothermiques résidentielles. En effet, certaines applications pourront se servir des analyses non intrusives pour détecter en temps réel des fuites de chaleurs et d'autres pertes d'énergie considérables [10].
- Lancer des applications pour détecter des appareils défectueux ou avec faible efficacité énergétique [11].
- Proposer des corrections ou des changements liés aux habitudes des clients afin d'optimiser la consommation. Par exemple, un rapport détaillé sur la consommation d'un chauffe-eau électrique pourra aider à détecter une consommation d'eau chaude excessive.

- Améliorer les modèles de prévision de la consommation locale. En effet, la présence de patrons d'utilisation périodique de certaines charges peut augmenter la qualité de l'information utilisée dans la prédiction de la consommation. De plus, l'analyse de la corrélation entre les profils désagrégés, les variables météorologiques et les données calendaires pourra enrichir les modèles de prévision.

En ce qui concerne la stratégie de surveillance, certaines approches proposent l'utilisation de capteurs dédiés à chaque charge d'intérêt. Bien que le mesurage individuel procure des avantages au niveau de la précision, le caractère intrusif défavorise son application à cause de la faible acceptabilité du client résidentiel. De plus, le nombre des capteurs, le matériel d'acquisition ainsi que les procédures d'installation entraînent de coûts relativement élevés [12]. Afin de pallier à ces difficultés, la surveillance non intrusive de charges (Non-Intrusive Load Monitoring (NILM)) a émergé en tant que technique capable d'estimer la consommation par usage à partir d'une mesure unique. Celle-ci comprend des variables telles que la tension et le courant échantillonnés au niveau du compteur ou du panneau principal. Idéalement, un système NILM doit être capable de : séparer de façon automatique les différents profils agrégés dans la courbe de charge au niveau du panneau électrique, reconnaître et classifier les appareils en fonction de leurs caractéristiques communes, détecter l'installation des nouveaux appareils et les variations du profil de ceux qui sont déjà installés, estimer le pourcentage d'énergie consommée par appareil et intégrer de nouvelles technologies et données, si ces dernières sont disponibles, afin de faciliter la désagrégation.

Dans le contexte Canadien, l'utilisation majoritaire des charges électriques thermostatiques concerne particulièrement les régions dans lesquelles le chauffage et le chauffe-eau électriques sont très répandus. Cette situation est illustrée à la Figure 1-3a, où les taux de pénétration par région montrent l'importance de ces deux usages dans l'intensité de la demande de l'énergie électrique résidentielle (plus de %80 du total). Par conséquent, une approche orientée vers la détection des charges électrothermiques est pertinente afin de mieux comprendre les facteurs qui déterminent les consommations élevées et faciliter postérieurement la prise de décisions ciblant l'économie. C'est la raison pour laquelle cette thèse propose de développer une approche de surveillance non intrusive orientée vers la détection des charges contrôlées par des thermostats.

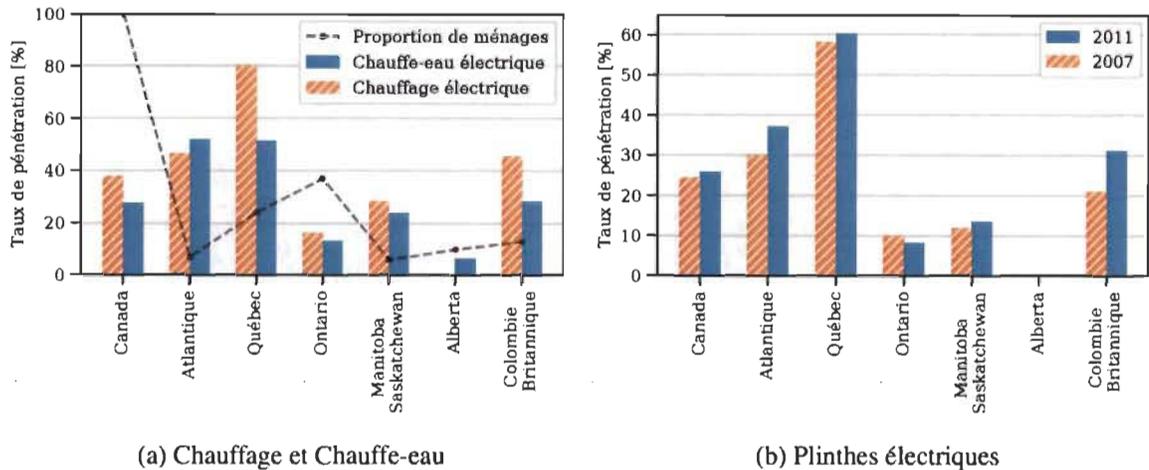


FIGURE 1-3 Taux de pénétration des charges électrothermiques dans les régions Canadiennes [13]

1.3 Problématique de thèse

À ce jour, plusieurs approches de surveillance non intrusive sont proposées dans la littérature scientifique, mais la viabilité de ces systèmes dans les marchés de masse est entravée par des contraintes technologiques et économiques telles que : des exigences de matériel spécialisé et coûteux pour l'acquisition et le traitement des données, un déploiement demandant une configuration intrusive et lente et une faible adaptabilité aux changements des profils d'intérêt.

Contexte du problème selon la complexité de la signature : La littérature scientifique concernant le domaine NILM a identifié quelques difficultés majeures associées à la détection de charges. Ces difficultés sont établies selon les caractéristiques de la signature électrique et elles sont considérées indépendantes de la méthode utilisée pour traiter le problème de désagrégation [14]. En effet, la complexité de la désagrégation est étroitement liée à la façon dont les caractéristiques individuelles sont mélangées dans l'espace de signatures agrégées. Sur ce point, l'étude publiée dans [14] énonce cinq causes nuisant aux performances des algorithmes de détection :

1. Un nombre relativement élevé de charges : cette situation augmente la probabilité d'avoir des ambiguïtés dans l'espace d'analyse de la courbe agrégée.

2. Les fréquences de commutation élevées de certains appareils : ce problème rend difficile l'analyse événementielle car plusieurs charges pourront commuter sur la même période d'analyse.
3. La présence de charges à multiples états augmente considérablement les configurations possibles de la courbe agrégée.
4. Les similarités de signatures entre les charges : il s'agit du problème le plus complexe à gérer. En effet, pour certaines charges thermostatiques, en plus d'avoir une consommation semblable, leur dynamique à travers le temps peut être similaire.
5. Le bruit et l'incertitude en général : bien que le bruit des mesures peut être négligeable, les incertitudes introduites par le fonctionnement de charges inconnues augmentent significativement le nombre possible d'interprétations d'une mesure agrégée.

A priori, la dynamique thermostatique engendre une signature simple. Cependant, le problème de surveillance se complexifie à cause des similarités entre les éléments chauffants et du taux élevé d'événements introduits par un nombre considérable de convecteurs électriques. En effet, les informations sur les séquences de changement d'état sont fréquemment utilisées dans l'élaboration de l'espace de signatures dans les approches NILM conventionnelles. Toutefois, le nombre d'appareils élevé, combiné aux commutations rapides engendre une quantité non négligeable d'informations à traiter³. Cette situation est quasiment négligée par la littérature NILM existante, tout simplement parce que le chauffage à convecteurs électriques a un taux de pénétration faible dans les pays autres que le Canada. De plus, tel que rapporté par Ressources Naturelles Canada (RNC) et illustré à la Figure 1-3b, le taux de pénétration des plinthes électriques augmente avec le temps dans presque toutes les provinces canadiennes. Donc, il est attendu que les futurs profils agrégés soient de plus en plus saturés d'évènements thermostatiques. En outre, l'absence d'études fiables et complètes traitant le cas énoncé rend la généralisation des approches NILM conventionnelles plutôt difficile.

Par ailleurs, il est connu que le cycle de fonctionnement des charges thermostatiques crée une signature de puissance qui se répète d'une manière plus ou moins similaire à chaque utilisation. L'analyse des récurrences et des périodicités des cycles d'opération peut aboutir à

3. Un thermostat électronique commercial a une période de commutation autour de 16 secondes.

la séparation et à l'identification statistique des profils ciblés. Cependant, l'application des approches existantes se complique à cause des difficultés suivantes :

- Étant donnée la nature linéaire des éléments chauffants, leur signature est réduite à une information à l'échelle d'onde [15]. C'est à dire, les mécanismes de détection basés sur les déformations de l'onde du courant (ondes non-sinusoïdales), en incluant l'analyse harmonique, sont peu applicables sur des charges purement résistives. Cette particularité engendre une signature toujours agrégée dans le régime permanent ; la détection devient un problème de calcul combinatoire. À cet égard, la solution la plus fréquemment utilisée dans le domaine NILM consiste à évaluer les événements ou les amplitudes des transitions de régime, soit du courant ou de la puissance active.
- Dû au grand nombre d'éléments présents dans le mélange, les possibilités de chevauchement spatial (amplitudes des signaux de puissances similaires) et temporel augmentent de façon significative. En effet, le chevauchement temporel est une difficulté de l'analyse événementielle. Cette situation aura lieu quand deux ou plusieurs charges changent d'état sur la même période d'échantillonnage. Quand cela arrive, les signatures événementielles seront agrégées, ce qui amène à considérer le calcul combinatoire. La plupart des approches NILM adoptent l'hypothèse que seule une charge change d'état à la fois. Cependant, à faibles fréquences d'échantillonnage et avec la présence de plusieurs thermostats électroniques, cette hypothèse est peu réaliste.
- La variation temporelle de la signature électrique peut avoir lieu en fonction de multiples facteurs non déterministes tels que les configurations et demandes de l'utilisateur, le nombre des résidents et les conditions environnementales [16]. En effet, les charges thermostatiques tels que le chauffage électrique sont influencées par des variables très stochastiques telle que la température extérieure. Dans ces scénarios, les allures des profils de puissance peuvent changer significativement avec le temps. Même si les charges thermostatiques gardent une puissance quasiment constante pour l'état d'activation, les variations significatives dans la demande d'énergie thermique sont transférées aux durées des états thermostatiques.
- Les modèles NILM actuels décrivant la dynamique des charges à deux états (On-Off ou thermostatiques), sont peu adaptés au comportement des charges électrothermiques

canadiennes. En ce qui concerne les durées des états thermostatiques, peu d'études sur leur distributions sont disponibles dans la littérature.

1.4 Objectifs et contributions

L'objectif général de cette thèse consiste de développer une approche pour l'identification non intrusive et très peu supervisée d'appareils domestiques thermostatiques à consommation élevée. De ce fait, les objectifs suivants sont définis :

- Développement d'un modèle à deux états basé sur l'analyse statistique et probabiliste des profils individuels des charges thermostatiques stationnaires à partir de la proposition d'une loi de transition adaptée à l'analyse permettant la caractérisation des durées des états d'opération d'une machine à deux états.
- Conception d'un algorithme pour l'agrégation de multiple éléments à deux états. Cet algorithme doit permettre l'ajout ou le retrait d'appareils avec une structure qui intègre facilement tous les aspects des modèles individuels.
- Évaluation des performances des approches proposées et comparaison avec des méthodes de l'état de l'art.

Les études réalisées et les approches proposées dans le cadre de cette thèse ont abouti à l'amélioration des connaissances approfondies dans le domaine de la surveillance non intrusive des charges électriques résidentielles canadiennes. Nous énumérons les contributions suivantes :

1. Développement d'une approche probabiliste dédiée à l'identification et la reconstruction de profils de puissance des charges thermostatiques canadiennes. Cette proposition s'appuie sur l'adaptation des réseaux bayésiens utilisés dans la modélisation de processus markoviens à durée explicite.
2. Conception d'une méthode basée sur les opérations de Kronecker permettant l'agrégation d'un nombre fini d'automates probabilistes à deux états ; ceci est compatible avec les algorithmes de programmation dynamique existants et capable d'intégrer les modèles de transitions markoviens et non markoviens.

3. Intégration des informations *a priori* concernant le fonctionnement des charges thermostatiques dans la phase d'initialisation des modèles de reconnaissance semi supervisés.

La première problématique que nous traitons dans ce sujet de thèse est la désagrégation des profils ayant une structure basée sur l'agrégation de processus stochastiques discrets à deux états. Le système de désagrégation est structuré à la fois à partir de modules développés dans le cadre de ce projet et en s'appuyant sur des techniques proposées dans la littérature. Toutefois, la nouveauté dans cette étude porte sur une formulation probabiliste du problème d'identification et une classification de multiple profils thermostatiques ayant des signatures en puissance similaires. Cette formulation est basée sur l'analyse statistique des informations macroscopiques (cycle d'opération) des charges ciblées. L'approche de désagrégation est conçue en suivant des critères favorisant le déploiement résidentiel du système NILM. Parmi ces critères nous considérons principalement : la reconnaissance à partir de mesurage de la puissance active échantillonnée à basse fréquence et la possibilité d'introduire les connaissances générales sur les de charges d'intérêt. Cette dernière capacité est à la base de la conception d'un système d'apprentissage très peu supervisé. C'est à dire, dans certains cas, on dispose seulement des informations de haut niveau plutôt que des signatures ou des données brutes des charges. Pour ce cas, on préfère une approche capable d'apprendre des données mais aussi capable de considérer les paramètres généraux décrivant le fonctionnement des charges.

Outre les contributions originales liées à l'analyse spécifique des charges thermostatiques, une contribution supplémentaire qu'apporte cette thèse est l'étude comparative entre plusieurs modèles de reconnaissance parmi des méthodes avancées de l'état de l'art. Cette étude a été focalisée sur l'analyse de signatures macroscopiques basées sur la puissance active des charges à états finis.

1.5 Méthodologie

Afin de traiter la problématique discutée, une approche du type statistique-probabiliste est proposée. Celle-ci considère des aspects négligés ou peu connus des profils agrégés des multiples charges thermostatiques. Pour ce faire, ce travail de recherche se divise en trois phases. La première phase concerne l'exploration du problème posé à travers la recherche bibliographique et la maîtrise approfondie des notions reliées au domaine d'étude. Dans une

deuxième phase, une philosophie de résolution est adoptée tout en répondant aux critères, aux exigences et aux particularités du problème abordé. D'abord, les approches existantes sont priorisées, leurs limitations et difficultés sont identifiées, et à partir d'une idée de recherche continue, des améliorations, des ajustements et des nouvelles méthodes sont introduites afin d'atteindre une solution appropriée. Dans la troisième phase, l'efficacité des propositions est validée en simulation et expérimentalement. Qui plus est, des études comparatives avec des techniques de l'état de l'art ont été réalisées. Un schéma résumant la méthodologie de recherche suivie dans ce projet est montré à la Figure 1-4.

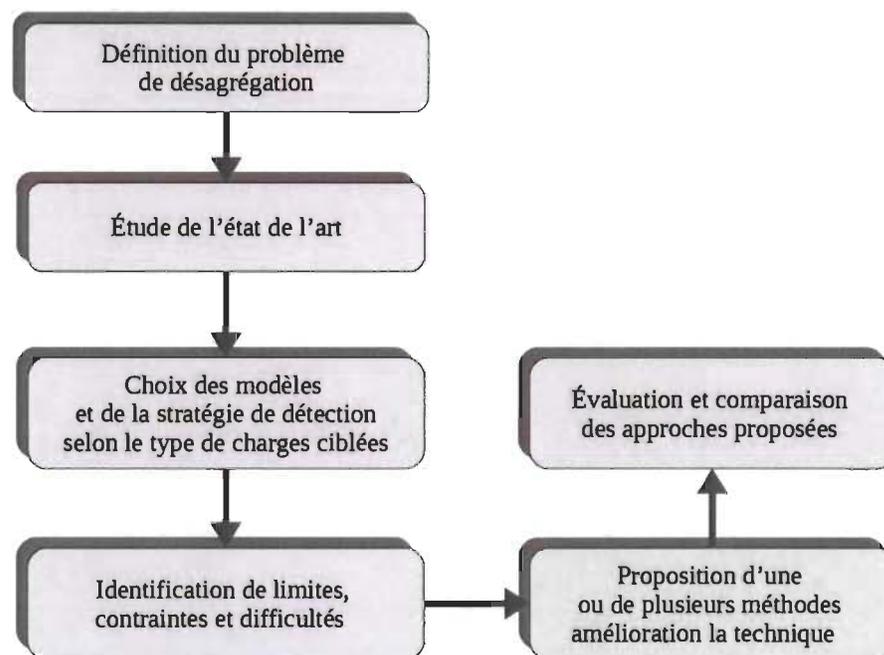


FIGURE 1-4 Méthodologie de recherche suivie.

L'étude exploratoire nous a permis d'identifier différents systèmes et méthodologies de résolution proposées actuellement pour la surveillance non intrusive. En effet, dans la bibliographie sur NILM plusieurs approches orientées vers l'analyse à états finis s'adaptent à la morphologie des profils thermostatiques. Parmi les techniques les plus complètes et générales, nous avons retenu les modèles Markoviens et certaines variantes connexes. Ces méthodes issues de l'analyse Bayésienne permettent, dans un seul cadre méthodologique, d'attaquer un nombre considérable de sous-problèmes associés à la surveillance non intrusive

dans un environnement incertain. Parmi les avantages de la méthode Bayésienne dans l'optique de la résolution du problème de la gestion et de la détection des signatures de charges, on souligne :

- l'intégration de l'information initiale et de l'apprentissage en utilisant de données réelles ;
- la prise en compte de l'incertitude sur les caractéristiques électriques et comportementales des signatures ;
- la mise en correspondance des données observées avec les modèles ;
- l'analyse de la structure temporelle du profil et l'inférence optimale des états non observés ;
- la classification statistique et le rejet des profils ;
- la modélisation des données agrégées à partir de la combinaison de paramètres et modèles individuels.

Il est clair que les modèles Markoviens ont été largement étudiés et appliqués dans le domaine NILM. Cependant, certaines hypothèses adoptées dans l'analyse de charges conventionnelles limitent leur application dans l'analyse d'un nombre considérable de charges. Sur ce point, le développement algorithmique de cette thèse cible l'extensibilité des modèles stochastiques discrets dans l'analyse de multiples instances thermostatiques. Justement, ce développement entraîne des améliorations dans le calcul de probabilités de transition en considérant des modèles sur les durées des états thermostatiques plus adaptés au comportement réel des charges. De plus, on présente une stratégie d'agrégation des modèles individuels qui augmente les possibilités de reconnaissance optimale parallèle. Par ailleurs, les modèles probabilistes utilisés se distinguent par leur pouvoir explicatif, c'est à dire, qu'ils sont en mesure d'identifier le rôle des éléments du modèle dans le fonctionnement du système physique.

1.6 Plan du manuscrit

Ce document de thèse est composé de six chapitres structurés de la façon suivante :

- Chapitre 1 : Introduction générale au sujet de thèse. La problématique, les objectifs et la méthodologie de recherche sont clairement définis dans ce chapitre.

- Chapitre 2 : Dans ce chapitre, un tour sur l'état de l'art général du problème de détection de charges est réalisé. D'abord, on décrit les aspects généraux de la surveillance. La façon dont l'apprentissage statistique aborde plusieurs problèmes reliés à la modélisation, la détection et l'étiquetage automatique de profils sera traitée. À la fin de ce chapitre, une synthèse sur les propositions NILM attaquant spécifiquement le problème des charges thermostatiques est abordée.
- Chapitre 3 : Ce chapitre est divisé en deux grandes sections décrivant de façon approfondie les approches proposées. La première section est consacrée à la description statistique des profils à deux états. Des formulations visant une proposition unifiée des modèles sur les durées des états et les modèles de transition du type Markovien et semi-Markovien sont énoncés afin d'établir un support mathématique cohérent avec l'analyse Bayésienne. La deuxième section est axée sur la proposition d'une approche probabiliste capable d'agréger de façon structurée un nombre fini de modèles à deux états. Celle-ci aboutit à un type particulier de machines probabilistes à états finis.
- Chapitre 4 : Évaluation de la méthode proposée. Dans ce chapitre, une évaluation visant à déterminer les performances des méthodes proposées dans plusieurs scénarios est réalisée. De plus, des analyses supplémentaires sont effectuées afin de comparer la performance des approches avec celle d'autres techniques NILM bien connues dans la littérature.
- Chapitre 5 : Conclusions et recommandations. Une synthèse des travaux réalisés est présentée, suite à quoi quelques limitations et difficultés retrouvées sont énoncées. Le chapitre finit avec des recommandations permettant la future amélioration de la méthode proposée.

Chapitre 2 - NILM - État de l'art

2.1 Généralités sur la surveillance non-intrusive

Les premières manifestations d'intérêt pour la surveillance non intrusive des charges datent des années 80 avec les travaux du Dr G.W. Hart dans les laboratoires de l'institut d'énergie du MIT (Massachusetts Institute of Technology) aux États-Unis [17]. Dans les premières études théoriques, Hart a défini le système de surveillance non intrusive comme celui qui est capable d'identifier les différents usages à l'intérieur d'un bâtiment en utilisant seulement la courbe de puissance agrégée mesurée au niveau du compteur électrique. Hart a aussi introduit le concept de signature non intrusive, comme celle qui peut être observée sans modifier l'opération normale d'un appareil. Ces concepts sont amplement utilisés et acceptés dans le domaine et donnent un point de départ pour la conception de systèmes NILM.

Selon la littérature, la détection d'un appareil implique que le système est capable de reconnaître sa signature définie dans l'espace de variables électriques mesurées. Quant à la désagrégation, elle implique que le système est capable de dissocier les profils individuels de la courbe de puissance agrégée. Ceci est connu comme un problème de décomposition d'un mélange additif en utilisant une mesure unique. De façon générale, la détection d'une charge est réalisée à partir d'une mise en correspondance entre les données mesurées et les informations préalables que le système dispose de telle charge. La Figure 2-1 montre l'architecture classique d'un système NILM. Dans celle-ci, les connaissances nécessaires pour l'identification sont répertoriées dans une base de données remplie pendant une phase de configuration ou d'apprentissage hors ligne du système. Toutefois, cette base de connaissances peut être construite ou améliorée à partir de processus d'apprentissage en ligne. De ce point de vue, deux types d'informations relatives aux charges sont stockées :

— **Signatures** : Ce sont des échantillons réels du profil de la charge d'intérêt, normalement

sous la forme de séries temporelles. Les signatures de la base de données sont utilisées pour détecter la charge en faisant une comparaison avec les signatures observées ou mesurées.

- **Descripteurs et modèles** : Ce sont des informations de haut niveau sur le comportement et la structure du profil d'une charge. Les modèles requièrent des paramètres relatifs à la charge d'intérêt à partir desquels des signatures peuvent être extraites pendant la phase d'apprentissage. Les descripteurs et les modèles peuvent être aussi introduits par un expert ou adaptés à partir des informations provenant du fabricant.

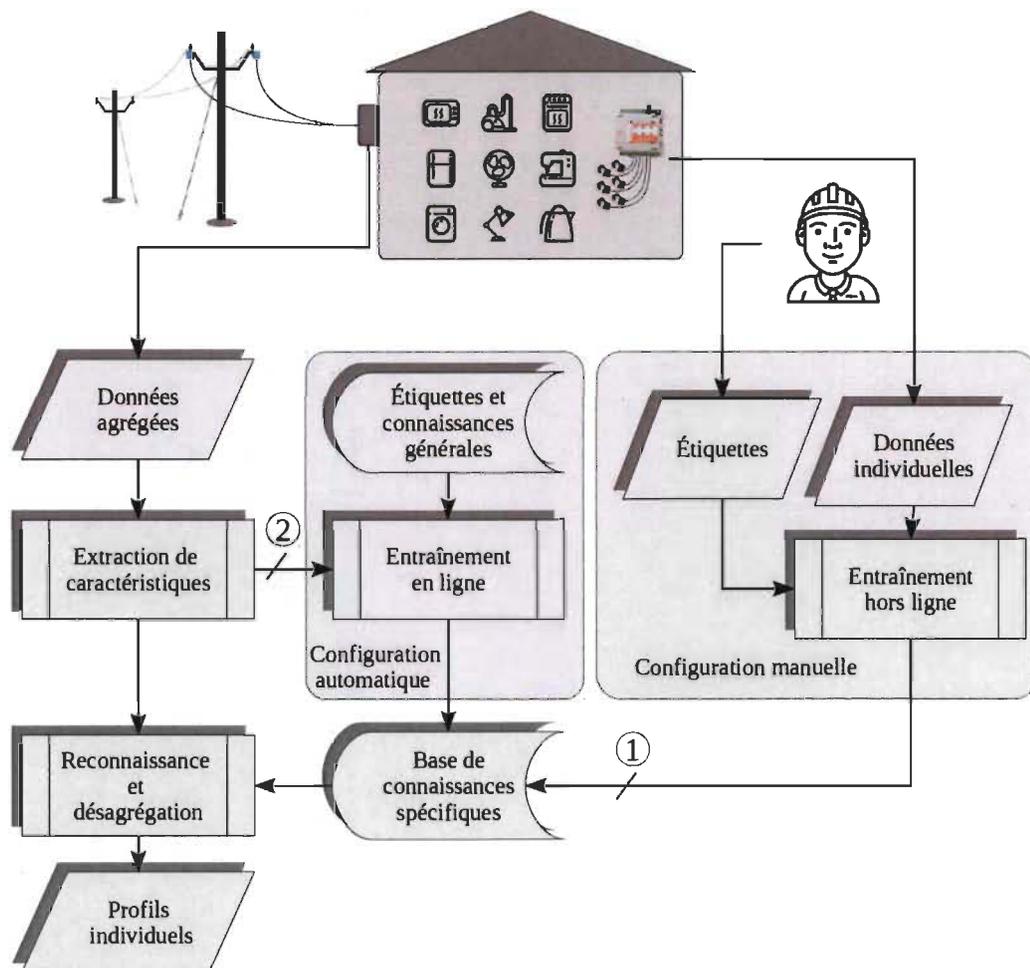


FIGURE 2-1 Structure générale d'un système NILM avec les possibles types de configuration.

De plus, la construction d'une base de connaissances est associée à une phase d'apprentissage qui peut être réalisée de façon hors-ligne ou en ligne. Dans la stratégie hors-ligne, les informations *a priori* sont introduites avant l'installation du dispositif. Pour la stratégie en

ligne le répertoire de connaissances est bâti à partir de données mesurées et étiquetées localement, ou bien transmises par des systèmes distants. Par ailleurs, la stratégie de configuration initiale du système NILM doit définir la façon dont les données étiquetées nécessaires pour l'apprentissage sont introduites dans le système. À ce propos, deux types de configuration sont normalement utilisés [17]–[19] :

1. **Configuration manuelle** : Cette stratégie comprend une phase de mesurage dédiée à chaque charge, celle-ci peut être faite en utilisant des capteurs pertinents et un système d'acquisition de données ou des compteurs divisionnaires (sub-metering). En utilisant cette stratégie, le système final peut demander moins des ressources de calcul que dans les approches de configuration automatique, mais il sera plus sensible aux variations dans les modes d'opération et aux changements des appareils à l'intérieur de la maison. De plus, pour le déploiement en masse, cette méthode est peu pratique, coûteuse, et elle peut être catégorisée comme semi-intrusive [20].
2. **Configuration automatique** : Les méthodes automatiques évitent l'intrusion temporelle à cause de l'installation de capteurs ou de la modification de l'architecture de l'installation électrique de l'immeuble. La configuration automatique vise un système "plug-and-play" dans lequel la minimisation de l'intervention humaine est privilégiée. Cette stratégie cherche la flexibilité du système de surveillance pouvant supporter les changements des profils et l'installation des nouveaux appareils, situations fortement probables dans un contexte de réseau électrique intelligent.

Les techniques basées sur l'entraînement à partir de sous-mesurage sur place sont peu pratiques dans les applications réelles. Cela à cause de la faible collaboration des clients et des coûts associés à l'installation des capteurs. Afin d'éviter le sous-mesurage, certaines approches utilisent des profils génériques mesurés dans des laboratoires ou dans des maisons expérimentales [21]. Cependant, cette solution demeure valide pour des charges ayant un profil peu variable d'une maison à une autre.

2.1.1 Catégorisation des charges

Hart a proposé une catégorisation des charges résidentielles en observant la morphologie des profils de puissance. Cette classification comporte trois grands groupes [17], [22] tels que

montrés à la Figure 2-1. Le premier (type I) regroupe les charges qui ont deux niveaux de puissance «on-off». Le deuxième groupe (type II) comprend les charges ayant un comportement défini par un nombre fini d'états d'opération, mais supérieur à deux. Dans le troisième groupe (type III) se trouvent les charges continuellement variables. Dans le contexte canadien, le type I et le II sont les plus courants dans les résidences et consomment un grand pourcentage d'énergie du total. De plus, les charges de type I et II sont aussi catégorisées comme des machines à états finis (Finite State Machine (FSM) ¹, en anglais). Cette dernière classification est à la base des techniques générales d'analyse séquentielle conçues pour modéliser le comportement des automates finis [23].

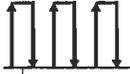
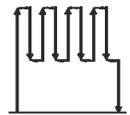
Type de Charge	Profil de puissance	Caractéristique	Exemple
I		Deux états de d'opération à puissance constante (On-Off).	-Réfrigérateur -Ampoule -Chauffe-eau (un élément) -Convecteur électrique
II		Plusieurs états d'opération à puissance constante.	-Aspirateur -Machine à laver -Usages composés de plusieurs charges de type I : → Chauffage → Chauffe-eau (plusieurs éléments)
III		Puissance d'opération continue.	-Air climatisée -Ampoule avec gradateur

TABLE 2-1 Type de charges électriques résidentielles selon la morphologie de leurs profils de puissance.

De façon générale, une FSM ou automate fini est un modèle séquentiel représentant un système caractérisé pour avoir un nombre déterminé de modes d'opération. Dans une FSM chaque état discret représente un mode d'opération lequel permet de lier le comportement de la machine à certaines variables observées ou manifestations de l'état de la machine. L'évolution

1. Une machine à nombre d'états fini sert à modéliser le comportement séquentiel d'un objet. Elle comporte un nombre limité et défini d'états.

d'une FSM est définie par des lois de transition entre les états discrets. Selon la nature du processus modélisé, ces lois peuvent être déterministes, stochastiques ou mixtes. Dans le contexte NILM, l'état de la charge est une variable cachée, ce sont alors les analyses sur les mesures du courant, de la tension et du temps qui permettront d'avoir une idée du mode d'opération dans lequel elle se trouve. Plusieurs variables dérivées de ces premières telles que la puissance active et réactive peuvent être utilisées pour la tâche de surveillance.

2.1.2 *Type de caractéristiques*

La sélection de caractéristiques ou de descripteurs des charges consiste à proposer un vecteur de mesures adapté à la méthode de résolution du problème particulier de désagrégation. En utilisant cette information, l'algorithme de désagrégation pourra discriminer les différents appareils. Cette discrimination est faite à partir de la mise en correspondance entre les modèles et les caractéristiques mesurées. Selon le domaine d'analyse deux types de caractéristiques sont fréquemment traitées dans NILM, les microscopiques et les macroscopiques [24].

Les caractéristiques microscopiques sont reliées à l'analyse des formes d'onde, principalement celle du courant [25], [26]. Normalement, on s'attend à ce que chaque état et chaque appareil émette des caractéristiques micro portant des différences significatives. Les descripteurs micro sont aussi souvent appelés grandeurs du régime permanent. Quelques descripteurs assez populaires rapportés dans la littérature sont : la puissance active et réactive [17], le facteur de puissance [27], les formes d'onde du courant [19], [28]–[30], les harmoniques de courant [17] et les courbes I-V [19], [31]. Toutefois, le mesurage de certains de ces descripteurs peut demander du matériel spécialisé à des fréquences d'échantillonnage élevées [31]. C'est pour cela que la recherche sur NILM vise l'utilisation de l'infrastructure de mesurage standard et déjà déployée telle que les compteurs intelligents [32]. Par ailleurs, il est connu que les charges thermostatiques (principalement du chauffage de l'espace avec convecteurs et le chauffe-eau) sont normalement résistives. Dans ces conditions, peu de descripteurs micro autre que la puissance active pourront être utilisés. En effet, une résistance ne déforme pas l'onde du courant, conséquemment, ce qui va différencier une charge résistive d'une autre c'est l'amplitude de la puissance et son comportement dans le temps.

En outre, l'analyse du comportement des caractéristiques micro pendant les phases de changement d'état des charges donne lieu à l'analyse transitoire. Celle-ci capture des informations pendant les phases de changement d'état de machines. Les transitions ont normalement des durées courtes et les probabilités de chevauchement sont plus faibles par rapport au régime permanent. Dans la littérature, l'analyse transitoire est associée au traitement des caractéristiques décrivant la phase de transition entre les états incluant la détection des événements [21], [33], [34]. En effet, un événement est un changement considérable de l'amplitude de la variable observée sur une durée courte. La détection basée sur événement est faite normalement en surveillant un seuil de déclenchement. Quelques exemples de ces descripteurs transitoires sont : les pointes du courant et de la puissance, la forme du transitoire, l'enveloppe spectrale [35], l'analyse par ondelettes [36], le temps de réponse [36], l'analyse de variations de la tension.

Par ailleurs, les caractéristiques macroscopiques intègrent des informations à des horizons plus larges. En effet, dans l'analyse macroscopique les caractéristiques micro et transitoires sont repérées à travers des fenêtres de temps pour augmenter la dimension et le pouvoir discriminant des classificateurs. Parmi les méthodes qui traitent les caractéristiques macro, on trouve les modèles sur les durées des états, les analyses de séries temporelles, les modèles des triangles et des carrés [27] et les techniques probabilistes incluant les durées des cycles d'opération [37].

2.2 Désagrégation de mélanges additifs

La désagrégation consiste à dissocier les différents profils composant la courbe de charge agrégée, laquelle est considérée comme un mélange additif. C'est à dire, une somme des profils individuels. La nature de ce mélange dépend entièrement des grandeurs électriques mesurées. En considérant les variables mesurables au niveau du panneau électrique, l'agrégation est une combinaison additive du courant, de la puissance et de l'admittance des charges connectées au réseau.

Soit $\vec{y} = \{y_k\}_{k=1, \dots, K}$ une séquence de mesures agrégées comprenant les résultats d'un mélange linéaire instantané², où k représente l'indice du temps discret et K la longueur

2. Le signal agrégé ne dépend que des combinaisons linéaires des puissances individuelles aux mêmes instants.

de la séquence. De la même façon, notons avec $\bar{\mathbf{x}}_m = \{x_k^m\}_{k=1,\dots,K}$ pour $m = 1, \dots, M$, les composants individuels de M charges en opération dans une résidence. Le profil agrégé $\bar{\mathbf{y}}$ est donc défini par la somme suivante :

$$\bar{\mathbf{y}} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 + \dots + \bar{\mathbf{x}}_M \quad (2-1)$$

Du point de vue algorithmique, la désagrégation consiste à déterminer un, plusieurs ou tous les termes à droite de l'équation (2-1) à partir de l'observation $\bar{\mathbf{y}}$. Ceci configure un problème inverse *mal posé*, où les mesures $\bar{\mathbf{y}}$ ne suffisent pas pour déterminer les composants $\bar{\mathbf{x}}_m$ individuels [38]. Afin d'aboutir à une séparation des profils, il faut introduire des contraintes ou des informations *a priori* permettant une solution unique. Cependant, l'individualisation d'un profil ou d'un ensemble des profils en considérant l'observation $\bar{\mathbf{y}}$ ainsi que la gestion de l'information préalable générale ou spécifique sur les charges demeure la difficulté principale du processus de désagrégation. En effet, il est parfois difficile de connaître en détail les modèles ou profils individuels. De plus, les variations temporelles des modes d'opération des certaines charges compliquent l'utilisation de techniques où les bases de connaissances restent statiques. Sur ce point, les méthodes adaptatives ou ayant des capacités d'apprentissage en ligne commencent à être de plus en plus utilisées et étudiées dans le domaine.

La formulation du problème de détection est des fois associée à l'estimation des signaux d'activation indiquant l'utilisation ou non d'une charge. Dans cet ordre d'idées, les connaissances *a priori* sont établies à partir de paramètres ou même des signatures reliées à l'état d'activation. La consommation moyenne en puissance et sa variance sont des exemples d'informations permettant la paramétrisation d'une charge. Les charges ayant plusieurs modes d'opération peuvent également être modélisées par des états discrets que le système de détection doit estimer. Mathématiquement, l'état d'une machine peut être représenté par une variable discrète z_k prenant les valeurs d'un ensemble aussi discret $S = \{1, \dots, M\}$ où M est le nombre d'états possibles associés à la machine. Cette formulation permet de regrouper toutes les machines de type I et II dans le concept de FSM. La puissance active estimée d'un appareil

à l'état z_k est définie par l'équation suivante :

$$\hat{x}_k = \mathbf{u} [z_k^m] \quad (2-2)$$

Où $\mathbf{u} = [\mu_1, \mu_2, \dots, \mu_M]$ est un vecteur définissant les consommations moyennes en puissance de la charge pendant chacun de ses états d'opération.

Conséquemment, l'estimation de la consommation d'une charge cible passe dans la plupart des approches NILM par la détection de ses états d'opération z_k^m tout au long de la période d'analyse. Cette stratégie a comme conséquence la manipulation de l'état d'opération en tant que variable non observée ou cachée. La stratégie adoptée pour la détection des états peut varier significativement selon le type d'information que l'on mesure, le modèle adopté et de la nature des charges cibles.

De façon générale, une approche NILM peut détecter les charges individuelles à partir des stratégies suivantes :

1. Détecter les états d'opération individuels en utilisant des techniques de reconnaissance de forme. Ceci est possible à travers une mise en correspondance entre les caractéristiques extraites de la courbe de charge agrégée et les bases de connaissances. Par exemple, la distribution harmonique d'une charge non-linéaire peut être suffisante pour savoir si elle est en opération dans une période spécifique [39]. Cette stratégie considère que les charges émettent certains types de signatures dans des régions différentes de l'espace de signatures.
2. Dissocier des profils génériques à partir du partitionnement de la courbe agrégée et les classifier selon une mesure de similitude ou de vraisemblance avec les connaissances ou les modèles individuels. Par exemple, une séparation aveugle de profils de type I peut retrouver ceux qui se répètent fréquemment sur les données agrégées telles que proposées dans [40] ou [41]. Une fois que les profils sont dissociés, une technique de classification pourra assigner des étiquettes selon les informations des charges candidates.

Généralement, les processus de désagrégation fréquemment implémentés font partie des approches plus générales de l'apprentissage automatique, l'analyse combinatoire et le

traitement de données séquentielles [22], [25].

2.2.1 Apprentissage automatique

L'apprentissage automatique est un domaine dont l'intérêt majeur est le développement des algorithmes permettant à une machine d'apprendre à partir d'un ensemble de données. Dans le domaine de surveillance de charges, ces méthodes visent la reconnaissance des structures temporelles dérivées des mesures brutes (courant, tension et temps). En d'autres termes, le module de reconnaissance cherche à identifier les états des charges z_k en mettant en correspondance les caractéristiques observées dans le signal agrégé avec les caractéristiques connues sur chaque appareil. La mise en œuvre du système de surveillance est basée sur trois phases bien établies dans l'apprentissage automatique : la sélection de caractéristiques, la conception du classificateur et la stratégie d'apprentissage. Dans l'apprentissage automatique, le travail consiste principalement en la création d'une fonction de prédiction de classes. Ces fonctions sont réglées à partir d'une procédure d'apprentissage. L'utilisation de règles ou lois de discrimination pour donner une étiquette à chaque instance d'entrée configure un processus de classification, et la sélection d'un type de classificateur donné va déterminer aussi le type d'apprentissage nécessaire. En termes généraux, l'apprentissage des systèmes de reconnaissance peut se faire à partir de données provenant de mesures faites sur les appareils d'intérêt ou bien en profitant des connaissances empiriques sur les charges permettant la configuration d'un système expert ³[42]–[44].

Selon la disponibilité de l'information et des données étiquetées, trois types d'apprentissages sont fréquemment utilisés dans le domaine NILM. Ce sont les apprentissages supervisés, non supervisés et semi-supervisés. Dans l'apprentissage supervisé, un ensemble d'échantillons de données correspondant aux mesures sur un ensemble de variables est disponible avec les étiquettes respectives associées. Ces données sont utilisées comme exemple dans la conception de modèles de régression. Toutefois, l'obtention des étiquettes pour les signatures est généralement effectuée manuellement, par un utilisateur ou en utilisant des compteurs installés sur chaque appareil dans la maison. Une façon de contourner ce défaut consiste à faire un

3. Un système expert est un outil capable de reproduire les mécanismes cognitifs d'un expert, dans un domaine particulier. Il s'agit de l'une des voies tentant d'aboutir à l'intelligence artificielle.

apprentissage supervisé à partir de données de tiers (base de données publique) [37]. Les modèles ou règles créés sont améliorés en utilisant une phase d'apprentissage semi-supervisé qui considère les données non étiquetées (normalement agrégées) pour compléter l'apprentissage [45], [46]. Les algorithmes les plus populaires basés sur l'apprentissage supervisé sont : la méthode des k plus proches voisins (algorithme KNN pour « k-nearest-neighbor ») [47], [48], les méthodes basées sur les réseaux de neurones (Artificial Neural Network (ANN), en anglais) [49]–[51], les séparateurs à vastes marges (SVM pour « Support Vector Machine ») [47], [52] et les classificateurs bayésiens [53], [54]. D'autre part, l'apprentissage non supervisé est implémenté en utilisant des méthodes de partitionnement automatique, tels que l'algorithme des k-moyens qui permet de regrouper les données par paquets homogènes par rapport à une mesure de similarité [24], [55], [56].

2.2.2 Optimisation combinatoire

L'optimisation combinatoire et la programmation linéaire sous certaines contraintes sont utilisées dans plusieurs approches NILM afin de faire une estimation des états z_k^n . En effet, à partir d'une connaissance suffisante sur les profils composant la courbe de charge, l'ensemble des vecteurs d'états $z_k = \{z_k^1, z_k^2, \dots, z_k^M\}$ peut être estimé par le biais d'une programmation en nombres entiers (problème d'estimation de l'état d'opération). Dans sa forme générale, un problème d'optimisation combinatoire (on dit aussi d'optimisation discrète) consiste à trouver dans un ensemble de solutions réalisables. La notion de meilleure solution est définie par une fonction objectif basée normalement sur l'erreur d'estimation. Pour le cas de l'erreur quadratique, cette fonction prend la forme suivante [57] :

$$\hat{z}_k = \arg \min_{z_k} \left(y_k - \sum_{m=1}^M \mathbf{u}_m [z_k^m] \right)^2 \quad (2-3)$$

Quelques inconvénients sont inhérents à l'approche combinatoire [58]. D'abord, l'un des problèmes reconnus est d'avoir une complexité de calcul élevée et une solution sensible aux perturbations (profils non modélisés ou inconnus). De plus, certaines hypothèses considérées sont difficiles à respecter dans une application pratique. Précisément, la rigidité du problème combinatoire et la dépendance d'une connaissance approfondie de tous les profils composant

la courbe de charge sont peu compatibles avec les méthodes non supervisées ou automatisées. De plus, la solution du problème défini par l'équation (2-3) ne considère pas les dynamiques temporelles des profils de puissance. C'est à dire, on trouve une solution pour l'instant k indépendamment de solutions trouvées pour les autres instants du temps. Cette dernière limitation fait de l'optimisation combinatoire une méthode peu applicable pour les profils agrégés composés d'éléments à puissances similaires. En effet, de multiples combinaisons peuvent expliquer le même profil agrégé rendant une solution manquant d'unicité. Par exemple, dans un profil composé de trois charges, deux premières ayant une puissance nominale de 100 watts et une troisième avec une puissance de 200 watts, si l'on observe une puissance totale de 300 Watts, on aura deux solutions avec le même coût. Cette difficulté de manque d'unicité est solutionnée par les méthodes d'analyse séquentielle dont la structure des données au cours du temps devient importante dans la solution du problème combinatoire.

2.2.3 *Analyse séquentielle déterministe*

La plupart des méthodes d'apprentissage automatique sont orientées à la détection de l'état de façon ponctuelle et certaines difficultés sont retrouvées pour l'analyse de l'information enfermées dans de longues séquences. En effet, l'analyse structurelle du profil en tant que série chronologique est un sujet de recherche très actif dans le domaine NILM. Les méthodes orientées vers la modélisation des longues séquences permettent d'élargir l'horizon d'analyse qui est normalement réduit dans l'analyse des caractéristiques statiques ou de transitoires courts. Cependant, quelquefois les techniques d'analyse à court terme sont combinées avec les méthodes séquentielles afin de couvrir une plage plus large dans la description du profil. Dans cette perspective, une approche assez complète est proposée dans [23], celle-ci est composée d'un algorithme partitionnement flou, un algorithme génétique (GA) et un module basé sur la programmation dynamique.

Par ailleurs, une des difficultés dans l'analyse des longues séquences se trouve dans leur dimension élevée. Ceci parce que le système doit trouver la correspondance entre les données tout au long du temps d'analyse. Cette difficulté est normalement évitée en supposant une indépendance temporelle réduite entre les mesures. En effet, le modèle de machine à états finis proposés par Hart [17] considère seulement la relation séquentielle entre les événements

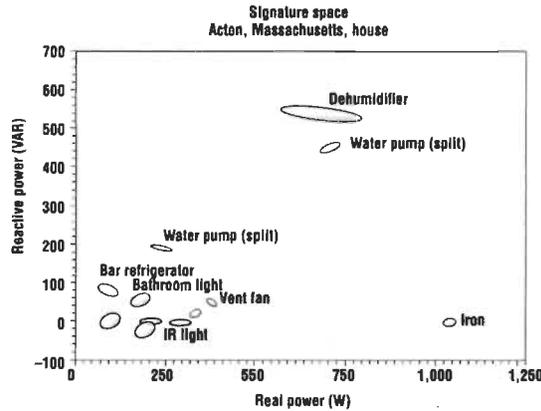


FIGURE 2-2 Carte de signatures intemporelles utilisées pour les transitions des automates déterministes. Source : [60].

positifs et négatifs à amplitude similaire afin de faire évoluer des automates virtuels simples. Dans le modèle de Hart, les événements sont associés aux variations de la puissance active et de la puissance réactive que les charges génèrent au démarrage. Il est nécessaire, dans cette méthode, de construire une carte d'événements où chaque charge doit occuper une position différente dans l'espace. Un exemple d'une carte de caractéristiques pour différentes charges est illustré à la Figure 2-2. En outre, les variables considérées dans cette analyse peuvent être la puissance active, réactive ou autre grandeur électrique. La technique de Hart considère qu'un événement détecté dans la courbe de charge agrégée correspond à un événement généré par un usage individuel. Toutefois, cette hypothèse est juste pour des profils agrégés ayant un faible taux d'événements. Pour le cas de profils ayant des charges avec des fréquences de commutation élevées, les événements qui se présentent de façon simultanée compliquent la désagrégation et demandent des techniques complémentaires pour la discrimination [59]. Ceci rend difficile l'application des méthodes basées sur l'analyse événementielle classique dans le contexte canadien.

2.2.4 Analyse séquentielle probabiliste

Les techniques basées sur la statistique et les modèles stochastiques permettent aussi de gérer les différents types d'apprentissages dans un cadre théorique bien défini. Les méthodes bayésiennes ont pris beaucoup d'intérêt dans les dernières années dans le domaine NILM, parmi ces techniques, nous soulignons les modèles de Markov cachés (Hidden Markov Model

(HMM), en anglais) [37], [46], [61]. En effet, les HMM offrent une formulation générale pour les machines ayant un nombre fini d'états (type I et II). Les états z_k sont de variables latentes, ou inobservables, le décodage est possible en utilisant des techniques de programmation dynamique. Cet algorithme intègre de façon cohérente la probabilité d'avoir une séquence d'états cachés et celle d'avoir la séquence observée. En utilisant les HMM standard il est possible d'estimer les états cachés de machines à partir des séquences observées non agrégées, mais pour le cas de signaux agrégés il est nécessaire d'utiliser de variantes des HMM telles que les modèles HMM à différences et factoriels.

Le modèle HMM à Différences (DHMM) a été introduit dans les travaux [37], [46], [61]. La particularité des DHMM se trouve dans le fait qu'ils analysent les changements produits par la charge sur le signal observé pendant les transitions des états. Cette stratégie permet d'unifier l'analyse événementielle avec les méthodes séquentielles probabilistes. De plus, afin de rejeter des transitions inconnues et de pénaliser les surestimations, des calculs de probabilité supplémentaires sont ajoutés dans l'algorithme de décodage. Toutefois, l'analyse des transitions à l'égard de chaque modèle individuel considère une opération non conjointe entre les différents modèles du mélange. Ceci entraîne deux problèmes spécifiques : 1.) la solution est sensible à l'ordre de désagrégation ; 2.) les effets combinés ou chevauchés ne sont pas prédits par les modèles, par conséquent, ils seront catégorisés rares ou inconnus. Une solution pour éviter les deux problèmes antérieurs consiste à faire une évaluation parallèle de l'ensemble des modèles. Dans cette perspective, on calcule la solution qui maximise la vraisemblance conjointe du mélange des modèles individuels. C'est à dire, on évalue l'optimalité de l'inférence des transitions combinées tout en prédisant les effets agrégés des processus markoviens. Cette dernière stratégie est connue comme l'analyse factorielle débouchant dans les HMM Factoriel (FHMM) [37], [61].

En raison de leur capacité d'analyse des effets combinés, les modèles FHMM conduisent à un formalisme plus adéquat pour représenter des signaux agrégés et pour le décodage des états de plusieurs machines fonctionnant simultanément. Comme illustré à la Figure 2-3, cette approche considère de multiples chaînes de Markov évoluant de façon indépendante et parallèle. La variable observée est modélisée par une loi gaussienne dont la moyenne est la combinaison linéaire de moyennes des HMM en question [62].

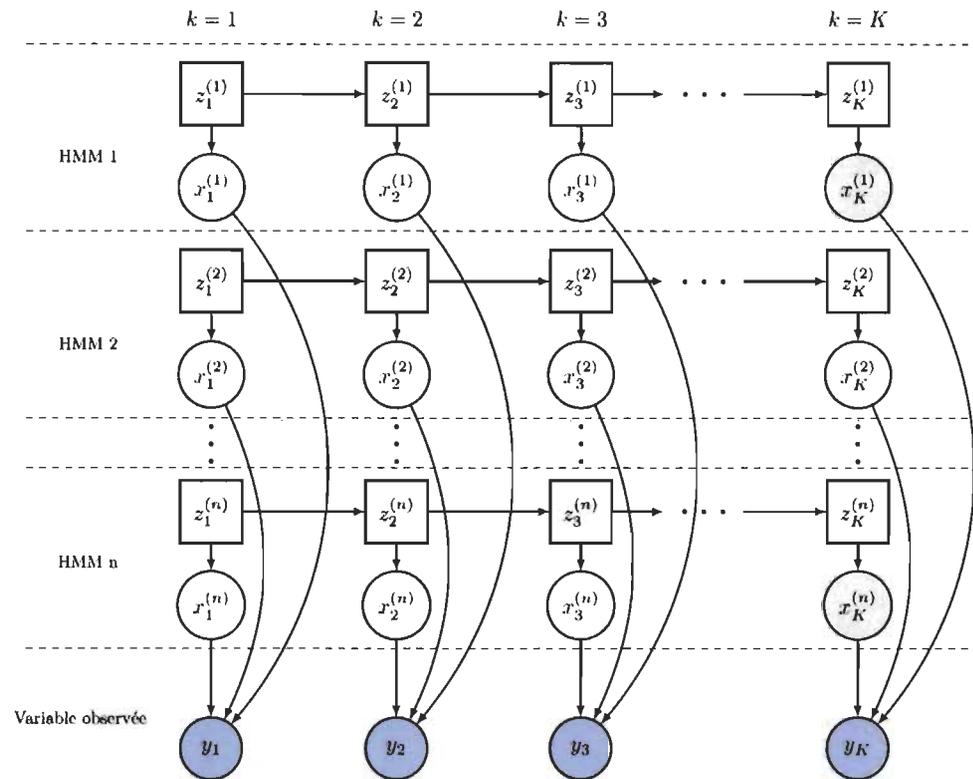


FIGURE 2-3 Réseau bayésien d'un FHMM [37].

Par ailleurs, il est connu dans la littérature que les HMM ont des limitations dans la profondeur d'analyse sur l'historique des séquences [63]. Ces limitations sont liées à la complexité de l'analyse de dépendances entre les états à l'instant k et les états antérieurs autres que celui de l'instant $k - 1$. En effet, pour gérer les dépendances à moyen et long terme dans la structure temporelle des données, souvent, les modèles de Markov d'ordre supérieur sont utilisés [64]. Toutefois, ces modèles engendrent des lois de transition ayant une dimension proportionnelle à l'ordre d'analyse (profondeur de l'historique), ce qui est peu efficace en termes du coût de calcul. Ces difficultés ont motivé le développement des modèles dont les durées des états apportent une information simplifiée de l'historique. Ces approches sont encadrées dans la méthodologie générale définie par les approches semi-markoviennes (Hidden Semi-Markov Model (HSMM), en anglais) [38], [65]. Le modèle semi-markovien a comme caractéristique principale l'utilisation de distributions adaptées pour les durées des états. Cette caractéristique ajoute de la complexité dans l'analyse due aux besoins dans la gestion de la durée et l'espace de possibilités dans le calcul de probabilité de transition. En revanche, cette méthode fournit une modélisation plus précise en récupérant le contrôle sur la durée de

chaque état. De plus, leur capacité de généraliser les HMM rendent les HSMM une technique intéressante dans le développement des approches NILM orientées aux machines à états finis stochastiques.

2.3 Approches orientées aux charges thermostatiques

Les principaux travaux réalisés dans la littérature considérant de façon très détaillée le chauffage par convecteurs et le chauffage de l'eau sanitaire ont été faits en France [66]. Une première approche proposée comporte une représentation temps-fréquence fondée sur l'utilisation conjointe d'un détecteur de non-stationnarités dans le plan temps-fréquence (spectrogramme) et d'une méthode d'extraction des motifs fréquents [67]. La reconnaissance des profils thermostatiques stationnaires est faite en évaluant la densité des événements. Pour détecter le régime stationnaire, un test d'hypothèses est fait en utilisant une fonction de contraste, la décision sur la reconnaissance ou non est prise à partir d'une règle de seuillage.

Une deuxième approche de type probabiliste est proposée pour la décomposition de la courbe agrégée contenant seulement des charges pour le chauffage [68]. Celle-ci considère que la courbe agrégée de puissance est constituée uniquement par un mélange d'un nombre inconnu de créneaux quasi périodiques. La période de régulation et la puissance nominale de chaque élément sont de caractère déterministe, en contrepartie, le nombre de cycles, le rapport cyclique de chacun et le moment de début sont considérés comme des variables aléatoires. La solution du problème d'estimation de cycles thermostatiques est faite en utilisant un algorithme d'optimisation basé sur les Monte-Carlo par Chaînes de Markov (MCMC) à Sauts-Réversibles (RJMC) [69]. Cette méthode s'inscrit dans le cadre du formalisme bayésien et elle génère des données permettant une solution au sens du maximum *a posteriori*.

Cependant, le comportement de l'algorithme dans un scénario ayant plus de charges autres que le chauffage et le problème de similitude de signatures restent à étudier. Nonobstant, ces études dévoilent des pistes importantes dans la construction d'une approche générale pour les charges thermostatiques. D'abord, le caractère non déterministe de la plupart des paramètres reliés à la détection de profils amène à considérer les modèles stochastiques ayant un cadre de résolution bien défini tel que le formalisme bayésien. D'autre part, les problèmes reliés au

chevauchement d'événements et de signatures ne sont pas assez étudiés dans le domaine, un obstacle qui rend difficile la comparaison des méthodes existantes.

2.4 Synthèse des revues et critères pour la sélection du modèle

Généralement, les différentes méthodes Machine Learning (ML) utilisées pour la solution du problème NILM peuvent être regroupées dans deux grandes familles bien définies : les modèles génératifs et les modèles discriminants [54]. Nous allons présenter certains critères qui nous permettront de choisir, parmi ces familles, une méthode bien adaptée à la problématique de thèse.

Dans la famille de modèles génératifs, nous trouverons les méthodes modélisant les relations entre les charges et les observations à partir de lois jointes. Ces approches, issues du théorème de Bayes, développent généralement toute la structure de la solution à partir des hypothèses faites sur les relations de dépendance entre données, modèles et paramètres. Comme leur nom l'indique, les modèles génératifs sont capables de générer des données synthétiques à partir de l'échantillonnage des distributions. Cette capacité facilite énormément la création de scénarios de simulation, et dans certains cas, cette qualité est exploitée pour résoudre des problèmes d'inférence complexes. Dans le domaine NILM, les modèles génératifs sont présentes dans un grand nombre d'approches modélisant des automates probabilistes à états finis. Ceci à travers les modèles de Markov et leurs variantes connexes. Une autre caractéristique importante des modèles génératifs est reliée à leur capacité intrinsèque de gérer les incertitudes. En effet, à partir des définitions des intervalles de confiance et de crédibilité le système peut prendre des décisions adéquates face aux événements rares. Un événement rare ou inconnu est très probable dans les profils de consommation agrégée. Ils sont normalement introduits par des charges non modélisées et d'utilisation peu fréquente. Ils sont aussi induits par des variations significatives dans le mode d'opération d'un appareil, par exemple, si un utilisateur oublie d'éteindre le four électrique le matin et celui-ci reste en fonctionnement toute la journée.

Par ailleurs, dans la famille de modèles discriminants nous retrouvons les approches ciblant la classification à partir de règles qui vont tracer les frontières de décision. Un vaste et varié nombre d'approches NILM peuvent entrer dans cette famille [70]. Parmi les plus populaires,

nous soulignons, les réseaux de neurones artificiels (ANN) [50], les Machines à vecteurs de support (SVM) [26], la régression logistique et les k plus proches voisins (KNN) [71]. La famille discriminante est très utilisée dans les problèmes d'apprentissage supervisé et elle présente de bonnes performances dans la classification de caractéristiques non chevauchées. L'hybridation avec des modèles des automates déterministes permet leur application dans l'analyse séquentielle [42].

Les descriptions faites préalablement sur les deux familles peuvent être approfondies et probablement d'autres familles pourront être proposées. Cependant, pour le domaine NILM, nous pourrions classifier presque toutes les approches existantes dans ces deux catégories. Par ailleurs, afin de déterminer la méthode plus adaptée au problème de surveillance concerné dans cette thèse, nous allons présenter par la suite des critères de sélection selon les besoins et les exigences actuelles pour les systèmes NILM. En effet, tel que discuté dans la méthodologie, nous avons basé l'analyse sur des approches non-déterministe. Toutefois, nous représenterons quelques arguments justifiant ce choix. D'abord, il faut signaler qu'il existe un grand nombre de critères qui pourront être ajoutés pour avoir une vision plus large sur la qualité des modèles disponibles. Nous pouvons citer la robustesse, la convergence, l'apprentissage non supervisé, le temps réel, l'optimalité, l'efficacité, entre autres. Ces critères sont bien étudiés et démontrés pour le cas des modèles génératifs. Toutefois, nous voulons énoncer quelques exigences complémentaires dans le développement des approches NILM. Celles-ci font du formalisme bayésien une méthodologie assez appropriée pour le problème formulé :

1. Analyse séquentielle : possibilité de modéliser la structure temporelle des données et d'évaluer le profil sur des périodes de temps ajustés aux durées des cycles d'opération de charges. Sur ce point, le succès des méthodes Markoviennes dans plusieurs domaines d'application de l'intelligence artificielle se trouve précisément dans la gestion cohérente des données séquentielles discrètes. En effet, à partir du modèle génératif, on peut calculer la vraisemblance d'une séquence d'états et évaluer en termes des probabilités s'il elle a été générée par un modèle spécifique [72].
2. Gestion des connaissances *a priori* : cette capacité permet au système d'intégrer de façon cohérente les informations de haut niveau préalables. Ceci est important quand on dispose de l'information statistique générale ou de paramètres sur le mode d'opérateur

d'une charge plutôt que des échantillons de leurs signatures électriques. Par exemple, la valeur de la puissance moyenne demandée par des réfrigérateurs standards peut être incluse en tant qu'hyperparamètre [73].

3. Extensibilité : l'approche doit permettre une croissance progressive des connaissances en s'appuyant sur la capacité d'intégrer de nouveaux modèles sans affecter les structures existantes. En effet, l'agrégation probabiliste encadrée dans les techniques FHMM permet de prédire les combinaisons d'observations à partir des modèles préalables et les nouveaux sans re-entraîner et sans redéfinir les paramètres de l'algorithme de désagrégation. À cet égard, on privilégie les techniques capables de modéliser la distribution marginale du profil agrégé par ses causes (profils individuels). De cette façon, au moment d'inclure une nouvelle charge dans l'analyse, l'approche sera capable de prédire la nouvelle distribution du profil agrégé plutôt que l'estimer directement à partir des données.
4. Gestion de l'incertitude : la méthode est capable de modéliser les aspects non déterministes des profils d'intérêt, de plus, elle doit être en mesure de dissocier les événements rares ou inconnus. En effet, les variations non déterministes des signatures macroscopiques demeurent une difficulté importante à gérer dans les approches NILM existantes. Ces variations sont normalement provoquées par des changements dans les habitudes d'utilisation des charges et des facteurs environnementaux non mesurés. Par exemple, la variation du nombre de personnes dans une maison peut changer la signature macroscopique de la cuisinière ou de réfrigérateur, ainsi que la variation de la température extérieure change significativement la signature de l'ensemble des éléments de chauffage.
5. Unicité méthodologique : tous les processus et méthodes associées découlent de l'analyse statistique et probabiliste. Par conséquent, peu de paramètres ou de procédures non liés aux modèles individuels sont requis.

Chapitre 3 - Modélisation de multiples éléments thermostatiques

3.1 Introduction

Comme discuté préalablement dans le Chapitre 1, l'utilisation d'un grand nombre des charges à états finis configure un cas particulier d'agrégation de la courbe de consommation. En ce qui concerne la complexité de l'analyse des charges thermostatiques, la littérature NILM a identifié le problème de chevauchement ou similarité de signatures comme un défi majeur dans le développement et le déploiement de systèmes adressés à la surveillance de ce type de charges. Afin d'attaquer ces problèmes, nous développons dans ce chapitre une approche basée sur l'analyse des profils stochastiques discrets. Cette proposition permet d'exploiter efficacement la structure temporelle des processus discrets afin de mieux expliquer la consommation des charges individuelles, même quand leur puissance est très semblable en amplitude. Celle-ci est une méthode d'analyse basée sur la caractérisation macroscopique de la puissance active. Elle permet aussi de modéliser la composition des profils générés par un nombre fini d'éléments. Cette composition de charges donne lieu à un cas particulier des machines à états finis (FSM) dans lequel les modèles de transition entre les états agrégés peuvent être construits à partir de la combinaison de modèles élémentaires.

En outre, pour la surveillance des charges à états finis non déterministes, les HMM avec les variantes connexes se présentent comme l'état de l'art dans le domaine [21]. Les HMM s'inscrivent dans une méthodologie assez générale qui permet d'accomplir dans le même cadre théorique l'apprentissage, la détection, la classification et la décomposition des profils influencés par des incertitudes. Ces modèles comprennent l'analyse stochastique de séquences d'états discrets, ce qui donne une perspective plus large dans la modélisation et

la reconnaissance des profils. Les HMM sont décrits par une structure composée d'états discrets et de lois gouvernant les transitions. À cet égard, nous examinons l'application de variantes des HMM à partir de la manipulation des descripteurs statistiques d'une charge de type I, tout en arrivant à une description du comportement non déterministe des charges composées de multiples sous-profils thermostatiques. Une illustration schématisant la structure générale du processus de modélisation et désagrégation est montrée à la Figure 3-1. Cette stratégie permet de construire des modèles factoriels à partir de la combinaison des paramètres identifiables à partir de données expérimentales par une approche d'apprentissage supervisé ou bien semi-supervisé.

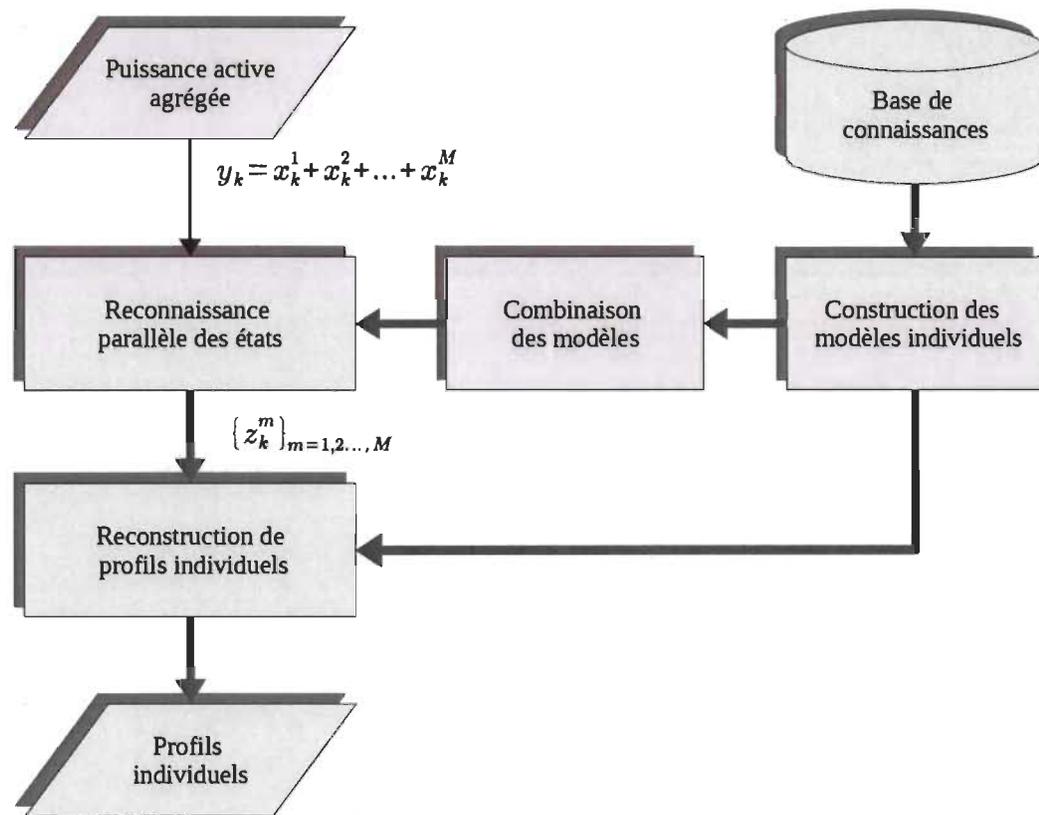


FIGURE 3-1 Schéma du processus de surveillance proposé.

Afin d'établir les fondements théoriques de l'approche factorielle utilisée, nous étudierons dans un premier temps l'aspect aléatoire du comportement des profils à deux états. La contribution principale dans cette première étude réside dans l'incorporation des distributions sur les durées des états pour le calcul de matrices de transitions tout en gardant une complexité

réduite dans les procédures d'estimation des états cachés. Une fois que les modèles subjacents sont déterminés, nous proposons une approche capable de combiner les paramètres individuels vers une structure à multiples états. Cette dernière possibilité simplifie de façon considérable l'analyse des lois de transition pour les charges composées de plusieurs éléments résistifs.

3.2 Modèle probabiliste à deux états

Bien que le caractère thermostatique est relié à l'existence de deux types d'opérations, ce concept peut être étendu à des machines à deux états d'opération même si la consommation à l'état "ON" n'est pas strictement constante. En effet, il n'est pratiquement pas possible de trouver un appareil avec une consommation strictement constante. Le fait d'avoir des charges constituées de résistances fixes n'assure pas une demande en puissance invariante dans le temps. En pratique, les variations de la puissance consommée par des charges résistives sont causées par des fluctuations dans la tension du réseau ou par les changements de la température interne des éléments chauffants. Par ailleurs, d'incertitudes plus considérables, mais moins étudiées dans le domaine, sont introduites par des facteurs gouvernant la durée des états [37]. En fait, il est très difficile sinon impossible de déterminer avec exactitude le temps d'utilisation d'une charge résidentielle. Dans ce contexte, les méthodes à effets aléatoires constituent un moyen plus élaboré d'étudier les variabilités et incertitudes.

De façon générale, le comportement de puissance des profils à deux états est lié à l'existence d'une variable permettant la régulation des charges de type tout-ou-rien (ON-OFF). Dans le processus de modélisation que nous développons pour la surveillance de charges de type I, nous relierons la variable de contrôle thermostatique au concept d'état de la charge. Le fonctionnement du contrôle thermostatique est relativement simple, le contrôleur habilite l'écoulement de puissance pendant une période de temps vers une charge ayant une impédance à valeur nominale constante, il interrompt l'écoulement par la suite, pendant une autre période de temps. Ce cycle peut se répéter plusieurs fois de façon régulière ou jamais. Le temps moyen dans l'état "ON" dépend des besoins énergétiques du système et de la nature de la charge. Les convecteurs pour le chauffage de l'espace, gérés par les thermostats électroniques ou bimétalliques, sont des exemples concrets de ce type de comportement. Ce groupe inclut

également d'autres charges telles que les chauffe-eau électriques, les résistances (serpentins) des cuisinières et la plupart des réfrigérateurs.

L'information non mesurée ou cachée comprend toutes les grandeurs physiques négligées par le système d'acquisition de données. Des exemples de ces grandeurs sont : la température de l'eau dans le chauffe-eau électrique, la température interne de la charge, ou bien le rapport cyclique d'un thermostat pour le chauffage de pièces. Cependant, pour le contrôle thermostatique en général, les effets de la plupart des variables non observées sont transmis de façon directe ou indirecte à la variable d'état z_k . Cette dernière s'avère aussi non mesurable, mais éventuellement estimable à partir des modèles.

3.2.1 Les états cachés et les observations associées

L'objectif de cette analyse consiste à introduire l'aspect comportemental des profils à deux états dans un cadre probabiliste. Plus précisément, nous adoptons un modèle permettant d'inférer les états cachés de charges thermostatiques à partir des variables mesurables lesquelles sont des manifestations des états. Pour ce faire, notons Z une variable aléatoire discrète indiquant l'état d'une charge de type I. Dénotons par z_k une réalisation de Z à l'instant k . Étant donné le caractère thermostatique du profil d'intérêt, nous désignons l'espace de possibilité pour Z l'ensemble discret $S = \{0, 1\}$, de telle sorte que $z_k = 0$ quand la machine se trouve à l'état "OFF" et $z_k = 1$ quand la machine séjourne dans l'état "ON".

Afin de modéliser la consommation, on dénote X une variable aléatoire représentante de la puissance demandée par la charge et x_k une réalisation de X à l'instant k . Étant donné la dépendance directe entre les processus X et Z , il est alors convenable d'exprimer leur interaction à partir de lois conditionnelles [74]. Dans cette perspective, on considère le théorème de Bayes comme point de départ dans l'analyse conditionnelle de variables. En effet, ce théorème exprime la probabilité *a posteriori* de la variable manquante (à estimer) sachant la variable observée (la puissance). Il est applicable à des observations ponctuelles, mais aussi sur des séries temporelles ou chronologiques. De ce point de vue, la probabilité postérieure d'une séquence d'états $\vec{z} = \{z_k\}_{k=1, \dots, K}$ de longueur K , sachant une séquence d'observations

$\vec{x} = \{x_k\}_{k=1,\dots,K}$, est définie par [22], [75] :

$$P(\vec{z}|\vec{x}) = \frac{P(\vec{z})P(\vec{x}|\vec{z})}{P(\vec{x})} \quad (3-1)$$

Dans l'équation (3-1), le lien entre la variable d'état Z et l'observation X est défini par la vraisemblance $P(\vec{x}|\vec{z})$, tandis que la dynamique du profil repose entièrement sur la probabilité $P(\vec{z})$. Cette dernière permet d'introduire les paramètres décrivant la structure temporelle du profil. Par ailleurs, le terme $P(\vec{x})$ désigne la probabilité marginale de la séquence d'observations. Pour l'inférence des états, le calcul de cette probabilité n'est pas nécessaire, car elle est une constante de normalisation. Conséquemment, la probabilité postérieure est une quantité proportionnelle à la loi jointe des états et observations : $P(\vec{z}|\vec{x}) \propto P(\vec{z})P(\vec{x}|\vec{z})$, le symbole \propto dénote proportionnalité.

Si les termes à droite de l'équation (3-1) sont calculables, l'estimation de la séquence d'états \vec{z} peut s'exprimer comme la solution du problème du maximum a posteriori (MAP). C'est à dire, parmi les séquences d'états possibles, on a choisi celle qui maximise la loi *a posteriori*. Les possibilités de résolution de ce problème reposent sur la capacité de calculer de manière appropriée les termes correspondants à la vraisemblance du modèle $P(\vec{x}|\vec{z})$ et la probabilité de la séquence d'états $P(\vec{z})$. De façon générale, l'inférence d'états consiste à trouver la séquence d'états qui maximise :

$$\vec{z}^* = \arg \max_{\vec{z}} P(\vec{z})P(\vec{x}|\vec{z}) \quad (3-2)$$

Sous l'hypothèse que la dynamique du système repose entièrement sur la variable d'état, une indépendance conditionnelle entre les valeurs de la séquence d'observations permet sa factorisation. Dans ce sens, la probabilité conditionnelle $P(\vec{x}|\vec{z})$ exprimée par le produit des vraisemblances individuelles tout au long de la séquence est donnée par :

$$P(\vec{x}|\vec{z}) = \prod_{k=1}^K P(x_k|z_k) \quad (3-3)$$

Afin de définir la vraisemblance en fonction de distributions et de paramètres, il est nécessaire d'associer à chaque état $s \in \{0, 1\}$ une loi conditionnelle $f(x|\theta_s)$, où θ_s est un

ensemble de paramètres associés à l'état s . L'analyse de ces distributions va permettre l'implémentation des techniques pour l'estimation des paramètres (ou apprentissage) dans un cadre supervisé. Généralement, les méthodes statistiques comprennent des méthodes assez connues et robustes telles que l'estimateur par maximum de vraisemblance (MLE), l'algorithme Espérance-Maximisation (EM), l'estimateur du Maximum *a posteriori* (MAP), et les méthodes Monte-Carlo. En appliquant les mêmes principes, les modèles statistiques vont faciliter l'estimation des états z_k quand ils ne seront pas observables. Pour le cas des charges à états finis, l'inférence des paramètres à partir des observations étiquetées $\mathbf{x}_s = \{x_k | z_k = s\}_{k=1,2,\dots,K}$ est accomplie en résolvant le problème suivant :

$$\hat{\theta}_s = \arg \max_{\theta_s} (f(\mathbf{x}_s | \theta_s) P(\theta_s)) \quad (3-4)$$

Où $P(\theta_s)$ est une loi *a priori* sur les paramètres.

Afin d'illustrer les distributions $f(x|\theta_s)$ en utilisant les données expérimentales, nous avons traité le profil de puissance présenté dans la Figure 3-2. Ceci correspond au profil de puissance d'un convecteur électrique contrôlé par un thermostat. Pour ce profil, deux états ont été détectés à partir d'une procédure de seuillage ciblant la détection de l'état d'opération du thermostat. Les distributions respectives pour chacun des états sont affichées à la Figure 3-3. Pour l'estimation de ces distributions, deux méthodes non paramétriques ont été utilisées. La première technique consiste à l'estimation de la distribution empirique sur un espace discret (histogramme). La deuxième correspond à l'estimateur de densité à noyau gaussien (G-KDE). Cet exemple permet de mettre en évidence le caractère stochastique des mesures de consommation d'une charge considérée fréquemment comme une charge à consommation constante.

3.2.2 Analyse probabiliste de la dynamique des états

L'analyse statistique de la durée des états a pris de l'intérêt dans les approches NILM, après avoir conclu que, les descripteurs spatiaux de la charge ne sont plus suffisants pour accomplir la reconnaissance des profils ayant des similarités en puissance. Par exemple, une approche basée sur la détection des événements de démarrage et d'arrêt d'un thermostat peut reconstruire un

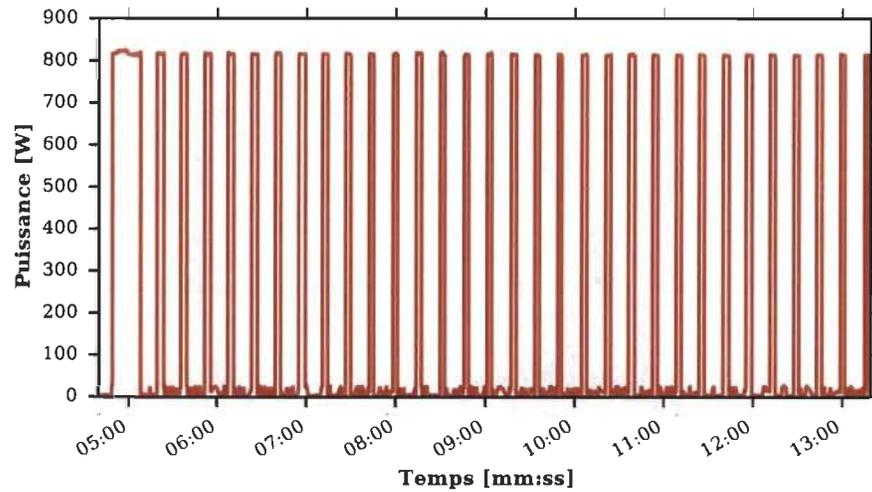


FIGURE 3-2 Profil de puissance d'un convecteur électrique. Données mesurées au laboratoire de l'IRH à l'UQTR et enregistrées à 2 échantillons par seconde.

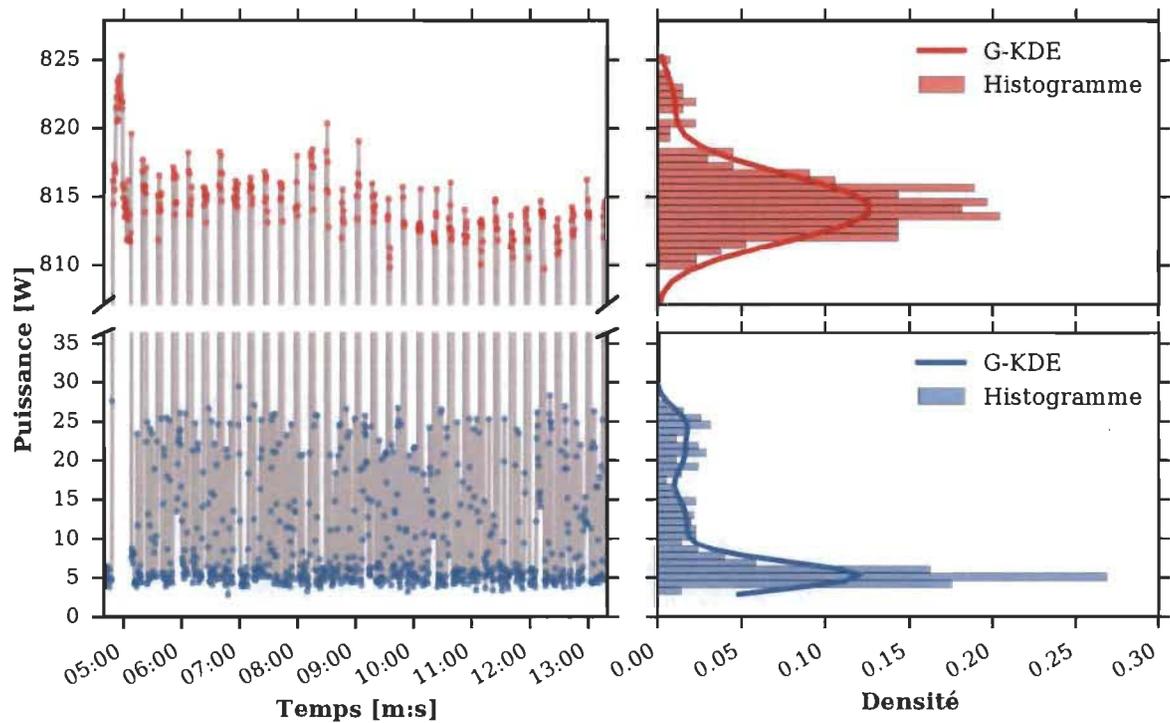


FIGURE 3-3 Profil et distributions des puissances associées aux états d'un convecteur électrique contrôlé par un thermostat électronique.

profil s'il n'y a pas de confusion avec les événements similaires provenant des autres charges. Cependant, l'analyse du temps écoulé entre les événements peut faire augmenter la capacité du système à dissocier les cycles caractérisés par les descripteurs fréquentiels ou temporels [76].

Pour les charges thermostatiques, la description de la dynamique du profil repose entièrement sur le comportement de la variable Z tout au long du temps d'analyse. À cet égard, pour caractériser statistiquement la structure temporelle du profil, il est donc nécessaire d'analyser le temps que la machine passe dans chacun de ses états. Dans un premier temps, l'information reliée aux durées des états sera utile pour évaluer la consommation énergétique espérée d'une machine quelconque de type I. Postérieurement, les paramètres statistiques résultants permettront de compléter un modèle discret stochastique qui s'avérera approprié dans le processus de reconnaissance et de classification. Étant donné la nature aléatoire des durées des états, certaines lois de probabilités doivent être proposées pour décrire ce comportement incertain.

Le terme $P(\vec{z})$ indiquant la probabilité de la séquence d'états a une importance primordiale dans la description de la structure temporelle des données. En effet, une grande partie de la modélisation des charges thermostatiques consiste à établir le lien entre cette probabilité et la dynamique des contrôleurs. Selon la règle du conditionnement en chaîne, la probabilité d'une séquence $P(\vec{z})$ peut être factorisée par des termes conditionnels selon l'équation suivante :

$$P(\vec{z}) = P(z_0)P(z_1|z_0)P(z_2|z_1, z_0) \dots P(z_K|z_{K-1}, z_{K-2}, \dots, z_1, z_0) \quad (3-5)$$

Étant donné que cette expression englobe toute la dynamique du profil FSM, son calcul relève un défi très compliqué si l'état présent de la machine a une dépendance à long terme avec l'état passé. Par conséquent, la dimension des lois conditionnelles explose quand la séquence devient longue. Cependant, le développement des approximations et l'adoption des hypothèses réduisant l'historique de la séquence d'états ramènent le problème à une dimension abordable. Ceci en tenant compte des hypothèses plus ou moins fortes pour réduire significativement la complexité de calcul. Dans le domaine d'analyse du profil de charges électriques à états finis, deux hypothèses d'indépendance à court et à moyen terme sont introduites afin de tronquer le calcul de $P(\vec{z})$:

1. le processus est Markovien : celle-ci considère que la probabilité de transition au présent ne dépend que de la connaissance de l'état immédiatement antérieur ;
2. le processus est semi-Markovien ou à durée explicite de l'état : cette hypothèse implique que la probabilité de transition au présent dépend du temps de séjour dans l'état actuel.

La première hypothèse est fréquemment utilisée dans la littérature sur NILM en ce qui concerne l'utilisation des modèles stochastiques des machines à états finis [21], [77]. Elle simplifie le calcul à partir de l'expression suivante :

$$P(z_k | z_{k-1}, z_{k-2}, \dots, z_1, z_0) = P(z_k | z_{k-1}) \quad (3-6)$$

Cette hypothèse réduit la mémoire du processus, car le fait de connaître les valeurs du passé lointain de la séquence n'apporte aucune information pour déterminer l'état futur de la machine.

La deuxième hypothèse correspond mieux aux comportements de certains appareils munis de contrôleurs thermostatiques. Dans ce type de charges, la persistance d'un état peut être reliée à certains phénomènes comme la durée qui catégorise une variable importante, puisqu'elle détermine l'énergie transférée à la charge. Pour ce cas, la continuité du système est associée à une probabilité calculée en fonction de la durée de l'état [78]–[80] :

$$P(z_k | z_{k-1}, z_{k-2}, \dots, z_1, z_0) = P(z_k | z_{k-1}, d_{k-1}) \quad (3-7)$$

Où d_{k-1} correspond au temps de séjour de la machine dans l'état spécifié par z_{k-1} .

L'équation (3-7) configure un modèle markovien à durée explicite (ou semi-markovien). Dans celui-ci l'information concernant l'historique de la séquence est réduite à la connaissance de l'état précédent et sa durée. Ceci permet d'estimer la probabilité du changement d'état sachant que la charge a demeuré pendant un certain temps d_k dans un état s . Ce calcul est avéré juste si l'information relative à la tranche de séquence $\{z_{k-2}, \dots, z_1, z_0\}$ est contenue dans le couple $\{z_{k-1}, d_{k-1}\}$. Il faut néanmoins souligner que, la variable d_{k-1} représente explicitement la valeur du compteur du temps de séjour à chaque instant de temps k . En effet, dans les approches semi-markoviennes les probabilités de transition sont calculées à partir des hypothèses sur la durée d'un état spécifique, laquelle correspond à la valeur de d_{k-1} au moment de changement d'état [38], [65]. Ces méthodes cherchent la construction d'une loi postérieure favorisant l'utilisation des distributions sur le temps de séjour. En effet, l'approche semi-markovienne calcule une postérieure jointe pour le couple état-durée conditionnée à l'état précédent. Cependant, une loi postérieure incluant la durée est considérée comme

naïve, car elle impose un espace de possibilités avec dimension élevée [38]. Afin de gérer la complexité dans l'inférence, ces approches s'adaptent aux techniques simulées ou méthodes de Monte-Carlo [65].

Selon la proposition de MURPHY [78], la complexité de l'analyse semi-markovienne naïve peut se réduire à partir du comportement semi-déterministe de la durée d_k . Cela implique une évolution de d_k déterminée par les états z_k et z_{k-1} . Cette caractéristique permet de garder une loi postérieure à espace réduit tel que proposé par le modèle (3-12). Toutefois, l'approche [78] utilise des compteurs décrémentationaux dont leur valeur initiale suit la distribution de la durée de l'état. Ceci force à une analyse par segments qui complexifie la manipulation des probabilités au moment d'inclure les séquences d'observations. De plus, cette technique ajoute des restrictions sur la plage des durées en les limitant à des valeurs maximales, ce qu'introduit de paramètres supplémentaires. Par ailleurs, une stratégie de calcul basée sur des compteurs incrémentaux est proposée dans [79]. Cette méthode proportionne un processus de calcul plus souple en basant l'analyse de transition sur la fonction cumulative de la distribution de la durée. Elle évite aussi les restrictions sur les compteurs. Cependant, la durée d_k est considérée inconnue, ceci oblige à une estimation approximative du temps de séjour.

La Figure 3-4 montre trois réseaux bayésiens utilisés dans la modélisation d'une séquence d'états. Ces réseaux décrivent le comportement d'une machine à deux états à travers le temps discret. Les diagrammes sont valables pour des machines à états finis en général. Le premier réseau correspond au processus markovien classique dont l'état présent est seulement affecté par l'état précédent. Dans celui-ci, la durée de l'état n'est pas considérée dans le calcul des probabilités conditionnelles associées à chaque nœud. Le deuxième réseau illustre la représentation d'un processus semi-markovien segmenté. Les états sont conditionnés par son précédent à partir d'une transition markovienne et la durée est conditionnée seulement à l'état. Le réseau modélise le processus semi-markovien en considérant explicitement la durée de l'état dans le calcul de probabilités de chaque nœud associé au temps. Ceci est basé sur un compteur décremental dont sa valeur initiale est générée à partir de la distribution du temps de séjour. À chaque fois que le compteur arrive à "1" la transition d'état est forcée est une nouvelle durée générée pour le nouvel état (processus de renouvellement markovien ou processus de Markov déterministes par morceaux). Dans cette procédure, le compteur est aléatoire pour

$d_{k-1} = 1$, le reste du temps, la durée et l'état même évolue de façon déterministe. Le troisième réseau est similaire en comportement par rapport au deuxième, nonobstant, ceci considère un compteur incrémental lequel est déterministe, car sa valeur est estimable à partir de l'état. Cette dernière caractéristique simplifie le calcul des probabilités conditionnelles associées aux nœuds de la durée. Cependant, l'état ne sera plus déterministe dans des morceaux du temps comme prévu dans le réseau deux. Ceci n'est pas un problème majeur, car le support de l'état a une dimension très faible (nombre d'états) comparée avec le support du temps (peut aller à l'infini). Le fait d'utiliser un compteur incrémental empêche d'utiliser explicitement la distribution de la durée. Toutefois, tel qu'il sera montré ultérieurement, la probabilité de l'état en fonction de la durée incrémentale peut se définir par des lois conditionnelles de survie.

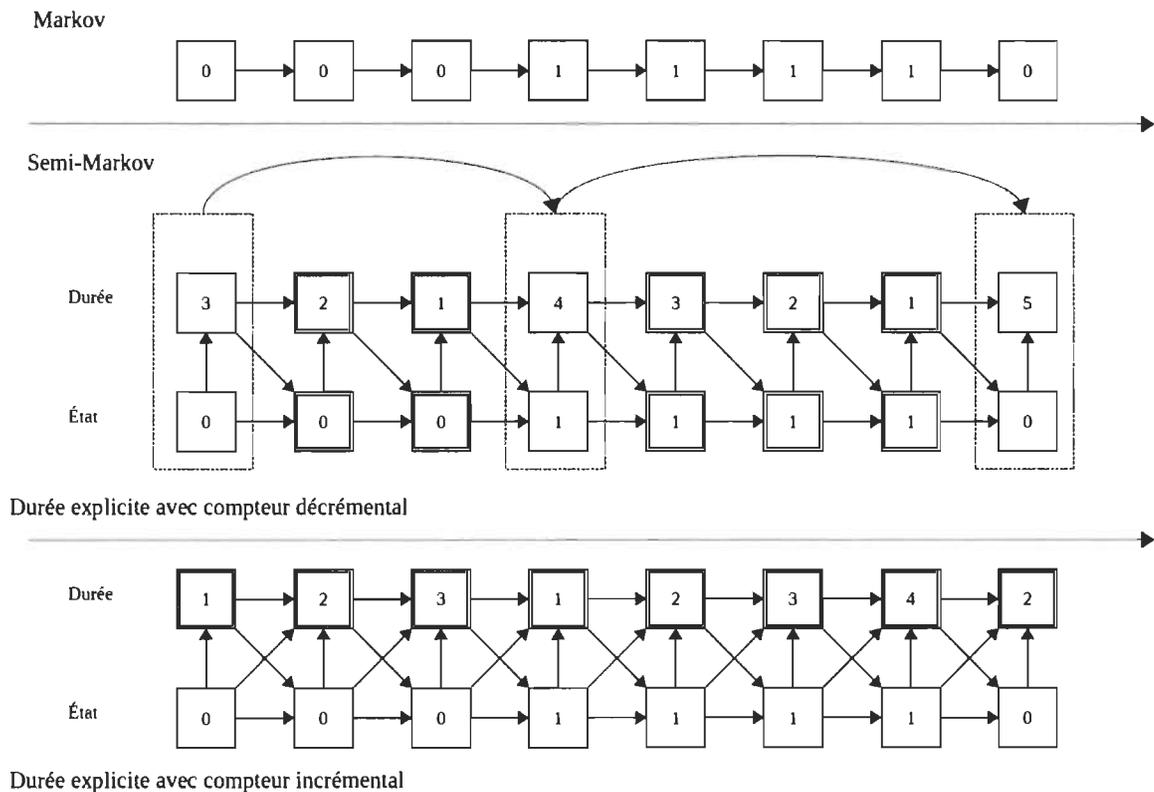


FIGURE 3-4 Réseaux bayésiens des processus markovien, semi-markovien et à durée explicite d'une séquence discrète à deux états.

Dans ce travail de thèse, l'approche [79] est priorisée avec quelques adaptations au problème de charges à états finis. Dans un premier temps, les probabilités de transition pour un modèle à deux états seront associées à la fonction de hasard discret (appelé aussi

risque instantané). Dans un deuxième temps, le comportement déterministe de la durée d_k est considéré à partir des incréments évoluant selon l'équation (3-8). Ce travail de modélisation permettra d'établir des lois de transition plus adaptées à la modélisation de charges thermostatiques pour le cas markovien ainsi que semi-markovien.

Pour mieux détailler l'analyse de transition, les différents aspects du modèle à durée explicite seront détaillés dans la section suivante (3.2.3). Nous montrerons que pour le cas des machines à deux états, le processus markovien est un cas particulier des modèles à durée explicite. Il faut souligner que, cette étude n'a pas pour objectif de faire une révision ou une étude approfondie sur les modèles semi-markoviens. Toutefois, nous partons du théorème de Bayes et d'autres notions de probabilité de base appliquées spécifiquement aux processus stochastiques discrets à deux états incluant les processus markoviens et à durée explicite.

3.2.3 *Modèle de transition incluant la durée du séjour*

Considérons d_k le temps discret écoulé à partir du dernier changement d'état. Ceci peut être vu comme un compteur avec une incrémentation déterministe et une réinitialisation aléatoire. L'évolution de d_k est décrite par l'équation suivante :

$$d_k = \delta(z_k, z_{k-1})d_{k-1} + 1 \quad (3-8)$$

Où $\delta(z_k, z_{k-1})$ configure l'événement de réinitialisation du compteur selon les situations observées aux temps k et $k - 1$. Un changement d'état indique une discontinuité sur la variable d'état de la forme $z_k \neq z_{k-1}$. La génération des événements de réinitialisation peut s'associer à la fonction delta de Kronecker qui correspond à la forme discrète de la fonction delta de Dirac. Par définition :

$$\delta(a, b) = \begin{cases} 1 & \text{si } a = b \\ 0 & \text{si } a \neq b \end{cases} \quad (3-9)$$

Étant donné qu'une transition entre deux états consécutifs différents génère un événement sur la variable observée, la fonction delta de Kronecker permet d'introduire le concept d'événement autant que l'évidence d'un changement d'état [80]. Cette fonction permet alors de réinitialiser à "un" la variable d_k si un changement d'état est survenu.

Étant donné le caractère aléatoire de la durée de chaque état, il est approprié d'exprimer les hypothèses sur les transitions en fonction des distributions ou des probabilités. Nous avons donc deux expressions pour la loi jointe de la variable d'états au présent et au passé ainsi que la durée de l'état. Dans ce sens, pour les événements $A = \{z_k \neq s\}$, $B = \{z_{k-1} = s\}$ et $C = \{d_{k-1} = d'\}$, les factorisations suivantes sont applicables selon la règle de conditionnement en chaîne [74] :

$$P(A,B,C) = P(A|B,C)P(C|B)P(B) \quad (3-10)$$

$$= P(C|A,B)P(A|B)P(B) \quad (3-11)$$

La loi *a posteriori* du changement d'état peut s'exprimer à partir de l'égalité des équations (3-10) et (3-11) selon l'expression conditionnelle suivante :

$$\begin{aligned} P(z_k \neq s | z_{k-1} = s, d_{k-1} = d') &= \frac{P(d_{k-1} = d' | z_k \neq s, z_{k-1} = s)P(z_k \neq s | z_{k-1} = s)}{P(d_{k-1} = d' | z_{k-1} = s)} \\ &= h(d' | \phi_s) \end{aligned} \quad (3-12)$$

Où $h(d' | \phi_s)$ est une fonction à déterminer et ϕ_s ses paramètres.

L'équation (3-12) décrit le processus de transition d'une machine à états discrets en considérant le concept de durée du séjour dans l'état. Toute la démarche probabiliste pour expliquer la dynamique des états et le processus de transition sera dérivée de l'analyse de ce modèle. Afin de mettre en place des algorithmes pour le décodage optimal des états cachés de machines, des lois de probabilité adaptées pour chaque terme de cette équation seront proposées. Ces lois seront définies selon les caractéristiques comportementales de la charge d'intérêt. Dans ce sens, deux types de situations peuvent avoir lieu à chaque instant de temps pour évaluer les hypothèses sur les transitions : la première situation concerne le changement de l'état et la deuxième indique la persistance de l'état. La relation entre les deux situations s'excluant mutuellement est déterminée par :

$$P(z_k = s | z_{k-1} = s, d_{k-1} = d') = 1 - P(z_k \neq s | z_{k-1} = s, d_{k-1} = d') \quad (3-13)$$

Les situations décrites préalablement configurent une épreuve de Bernoulli. Toutefois, les

probabilités associées aux deux possibilités ne sont pas constantes dans le temps. Celles-ci varient en fonction de la valeur du temps écoulé dans un état spécifique. Pour illustrer ce comportement, le pseudo-code de l'algorithme (3.1) schématise la dynamique d'une machine à deux états dont l'état bascule à partir d'un signal $\bar{\delta}_{k+1}$ non-déterministe. Ce signal est responsable du redémarrage du compteur du temps écoulé dans chaque état. Le calcul des probabilités de transition est réalisé selon l'équation (3-12) qui définit la fonction du hasard discret. En effet, au lieu d'échantillonner une durée en utilisant la distribution $f(d'|\phi_s)$, un signal de transition $\bar{\delta}_k \in \{0, 1\}$ est généré à chaque pas de simulation en échantillonnant une loi de Bernoulli avec paramètre $\alpha_k = h(d'|\phi_s)$ (ligne 5 de l'algorithme (3.1)). La mise à jour de l'état est réalisée en utilisant la fonction XOR entre l'état passé et l'événement de transition δ_k généré.

Algorithm 3.1: Algorithme pour la génération d'un profil stochastique à deux états.

Input: Paramètres : $\{\phi_0, \phi_1\}$

Output: Séquence des états discrets : z_1, z_2, \dots, z_K

```

1 begin
2   échantillonner  $z_1 \sim [\pi_0, \pi_1]$ 
3   initialiser  $d_k = 1$ 
4   for  $k = 2, 3, \dots, K$  do
5     échantillonner  $\bar{\delta}_k \sim \text{Bern}(h(d_{k-1}|\phi_{z_{k-1}}))$ 
6     mettre à jour  $z_k = z_{k-1} \oplus \bar{\delta}_k$ 
7     mettre à jour  $d_k = (1 - \bar{\delta}_k)d_{k-1} + 1$ 

```

3.2.3.1 Le temps de séjour et l'analyse de survie conditionnelle

Dans la méthodologie adoptée dans la modélisation du comportement du cycle thermostatique, l'analyse de la dynamique est spécifiquement liée à la caractérisation statistique des durées du séjour dans des états spécifiques. L'utilisation des lois de transition adaptées permettra d'inférer de façon optimale les séquences d'états thermostatiques. Selon la littérature, les probabilités $P(d_{k-1} = d' | z_k \neq s, z_{k-1} = s)$ et $P(d_{k-1} = d' | z_{k-1} = s)$ de l'équation (3-12) sont définies en termes de la distribution sur la durée et la fonction de survie de l'état respectivement [80], [81]. Justement, la probabilité de la durée à l'état s est exprimée par la

distribution sur la durée appelée aussi distribution du temps de séjour définie comme :

$$P(d_{k-1} = d' | z_k \neq s, z_{k-1} = s) = f(d' | \phi_s) \quad (3-14)$$

Où ϕ_s est un ensemble de paramètres associés à la distribution sur la durée de l'état s .

À partir de l'analyse de la loi sur le temps de séjour de l'état, il est possible d'estimer la probabilité de la durée sachant l'état comme une quantité proportionnelle à la fonction de survie associée à l'état s , c'est à dire, $P(d_{k-1} = d' | z_{k-1} = s) \propto S(d' | \phi_s)$ [79]. À cet égard, la distribution de probabilité sur les valeurs du compteur de temps est calculée en utilisant la fonction de survie laquelle est définie par l'expression suivante :

$$P(d_{k-1} = d' | z_{k-1} = s) = \frac{S(d' | \phi_s)}{\sum_{d'=1}^{\infty} S(d' | \theta_d^s)} \quad (3-15)$$

Dans l'analyse des modèles de durée, la fonction de survie est par définition le complément à la fonction de répartition du temps de séjour dans l'état s . Dans le temps discret, cette fonction est exprimée par le modèle suivant :

$$S(d' | \phi_s) = 1 - \sum_{\tau=1}^{d'} f(\tau | \phi_s) \quad (3-16)$$

3.2.3.2 Risque instantané et probabilité de transition

Afin de compléter le modèle de transition (3-12), une formulation pour le terme $P(z_k \neq s | z_{k-1} = s)$ en fonction des paramètres utilisés dans l'analyse de la durée de l'état est nécessaire. À cet égard, nous considérons le profil thermostatique comme un processus stationnaire homogène à deux états. Pour ce processus, la probabilité $P(z_k \neq s | z_{k-1} = s)$ est une quantité associée au rapport entre la probabilité d'un événement $P(z_k \neq s, z_{k-1} = s)$ et la probabilité stationnaire de l'état $P(z_{k-1} = s)$. Une analyse du processus à long terme permet de constater que ce rapport converge à l'inverse du temps discret moyen :

$$P(z_k \neq s | z_{k-1} = s) = \frac{P(z_k \neq s, z_{k-1} = s)}{P(z_{k-1} = s)} = \frac{1}{\mathbb{E}[d_{k-1} | z_k \neq s, z_{k-1} = s]} \quad (3-17)$$

Cette équation est explicable à partir du théorème élémentaire du renouvellement ALSMEYER [82]. En effet, en considérant des processus de comptage pour les évènements associés au numérateur et au dénominateur de l'équation (3-17), on peut constater que le nombre d'évènements associés au numérateur augmente de façon linéaire en fonction du nombre de cycles (ou nombre de visites à un état donné), tandis que le nombre d'évènements associés au dénominateur est lié au temps de séjour accumulé dans l'état s . En dénotant N_s le nombre de visites à l'état s pendant une longue période de durée K , le terme du numérateur peut s'écrire comme :

$$P(z_k \neq s, z_{k-1} = s) = \frac{N_s}{K} \quad (3-18)$$

De la même façon, le terme du dénominateur exprime la probabilité stationnaire de l'état s laquelle est définie comme :

$$P(z_{k-1} = s) = \frac{\mathbb{E}[d_{k-1} | z_k \neq s, z_{k-1} = s] N_s}{K} \quad (3-19)$$

Par conséquent, la probabilité de transition est définie par l'inverse de la durée moyenne de l'état, laquelle dans le cas limite correspond à l'espérance mathématique de la distribution du temps de séjour associé à l'état s .

Par ailleurs, dans l'équation (3-15) le dénominateur joue un rôle de normalisation. Sa valeur correspond à l'espérance mathématique calculée à partir de la fonction de survie [83], c'est à dire : $\sum_{d'=1}^{\infty} S(d' | \theta_d^s) = \mathbb{E}[d_{k-1} | z_k \neq s, z_{k-1} = s]$. À partir de ce constat, et en combinant les équations (3-12), (3-14), (3-17) et (3-15), nous avons une probabilité postérieure pour le changement d'état définie selon l'expression ci-après :

$$P(z_k \neq s | z_{k-1} = s, d_{k-1} = d') = \frac{f(d' | \phi_s)}{S(d' | \phi_s)} = h(d' | \phi_s) \quad (3-20)$$

Ce modèle de transition est aussi connu comme fonction de hasard discret (*discret hazard rate*, en anglais). Dorénavant, la difficulté dans le processus de modélisation de l'aspect dynamique des charges thermostatiques est liée à deux problèmes spécifiques. Premièrement, il faut proposer une loi $f(d' | \phi_s)$ adaptée à la dynamique des états des charges d'intérêt et calculer sa fonction de survie $S(d' | \phi_s)$. Deuxièmement, il sera nécessaire d'avoir une méthode

appropriée pour l'estimation des paramètres ϕ_s à partir des informations *a priori* et données expérimentales disponibles sur les charges.

Il faut noter que les méthodes markoviennes classiques évitent ce type d'analyse au détriment d'un modèle plus descriptif par rapport à la dynamique réelle des durées des états. Justement, le modèle (3-20) permet l'utilisation d'une distribution plus adaptée à la forme de la distribution empirique des durées thermostatiques. Nous expérimenterons ultérieurement avec quelques distributions paramétriques bien connues dans l'analyse de survie.

À partir de l'analyse de la fonction $h(d'|\phi_s)$ quelques propriétés peuvent être établies. Si cette équation aboutit à une expression indépendante de la durée, nous parlerons d'un processus "sans mémoire" ou markovien. Particulièrement, la propriété "sans mémoire" est introduite explicitement par une distribution géométrique sur la durée des états. En effet, la propriété de Markov est assumée dans la plupart des modèles de charges à états finis [37], [73]. Par contre, si une distribution autre que la loi géométrique est mieux explicative du comportement des états, un processus de comptage doit être ajoutés afin de faciliter le calcul des probabilités de transition non markoviennes. De plus, les distributions sur la durée discrètes peuvent être aussi calculées à partir de la relation $f(d'|\phi_s) = S(d'|\phi_s) - S(d' + 1|\phi_s)$. Ceci permet de passer directement à l'analyse des fonctions de survie lesquelles sont normalement plus simples que les distributions. Sur ce point, la probabilité de transition d'état peut s'exprimer de façon alternative par [84] :

$$h(d'|\phi_s) = 1 - \frac{S(d' + 1|\phi_s)}{S(d'|\phi_s)} \quad (3-21)$$

Un premier constat déterminant la nature des systèmes markoviens peut être fait à partir de l'équation (3-21). Étant donné qu'ils sont une probabilité de transition indépendante de la durée, leur fonction de hasard doit être constant. Effectivement, en résolvant l'équation (3-21) pour une fonction de survie géométrique $S(d'|\alpha_s) = (1 - \alpha_s)^{d'}$, la fonction de hasard résultant sera $h(d'|\phi_s) = \alpha_s$. Ce constat permet de concevoir un algorithme qui peut facilement basculer entre l'approche markovienne et semi-markovienne en modifiant convenablement la fonction de hasard discret.

Dans le travail de modélisation réalisé dans ce projet, la loi Weibull a été considérée pour le cas des modèles à deux états. Cette loi a été utilisée dans une approche NILM pour la reconstruction des profils à deux états présentés dans un article de journal publié en 2015

[41]. Afin de visualiser quelques distributions sur les durées et les modèles de transition d'une machine réelle, nous avons calculé les paramètres de trois lois différentes : 1) la loi Géométrique, 2) la loi de Weibull, et 3) la loi Log-logistique. Ces lois sont listées avec les fonctions de survie respectives dans le tableau 3-1. Dans cet exercice, le profil d'un réfrigérateur de la base ECO [85] a été traité pour déterminer les durées des états "ON" et "OFF" du thermostat. Ce profil est affiché à la Figure 3-5 .

Nom	Survie
Géométrique	$S(d' \alpha) = (1 - \alpha)^{d'}$
Weibull	$S(d' \lambda, \gamma, \varphi) = \exp\left(-\left(\frac{d'-\varphi}{\lambda}\right)^\gamma\right)$
Log-logistique	$S(d' \lambda, \gamma, \varphi) = \left(1 + \left(\frac{d'-\varphi}{\lambda}\right)^\gamma\right)^{-1}$

TABLE 3-1 Distributions utilisées pour modéliser les durées des états.

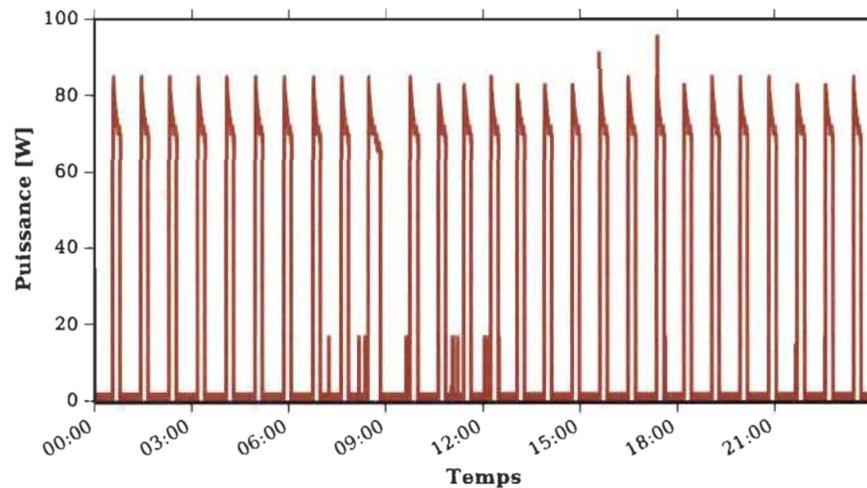


FIGURE 3-5 Profil de puissance et séquence d'états ON-OFF d'un réfrigérateur enregistrés à la seconde. Données provenant de la base ECO [85].

En plus du calcul et de l'affichage des distributions proposées, nous avons déterminé la distribution empirique ou histogramme des durées de l'état "OFF" du réfrigérateur enregistré pendant plus de 90 jours. Toutes les distributions sont tracées à la Figure 3-6a. Il est constaté visuellement que les lois de Weibull et Log-logistique ont une allure plus proche de celle de la distribution empirique que celle de la loi géométrique.

La Figure 3-6b montre les probabilités de transition de l'état "OFF" vers l'état "ON" du profil thermostatique du réfrigérateur évalué à une période d'échantillonnage $\Delta t = 1s$. Nous

observons que pour le cas de la loi géométrique, cette probabilité est constante en tout temps. Pour les deux autres distributions, cette probabilité est presque nulle au commencement, mais elle devient par la suite considérable quand la durée s'approche de la valeur moyenne de la distribution.

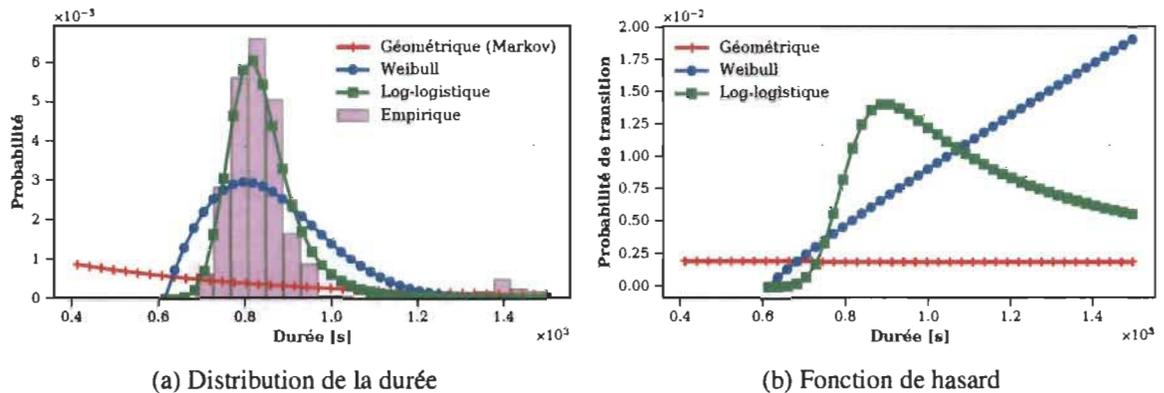


FIGURE 3-6 Distributions sur la durée et la fonction de hasard de l'état "ON" du profil thermostatique de la Figure 3-5.

Le fait de considérer de probabilités de transition non constantes améliore la précision du mécanisme d'inférence des états. Tout d'abord, les cycles brusques ou de courte durée sont naturellement rejetés dû à leur faible probabilité. En effet, un des défauts des modèles markoviens peut être constaté avec la courbe du hasard de la Figure 3-6b, où la probabilité est constante tout au long de la séquence. Cette situation implique que des transitions à n'importe quel moment seront acceptées avec la même probabilité. Les améliorations en terme de performance dans d'identification des profils seront mises en évidence avec une comparaison entre les résultats des cas markovien et semi-markovien. Le cadre comparatif sera précisé plus loin dans le Chapitre 4.

3.2.4 Reconnaissance probabiliste des profils à deux états

La reconnaissance d'un profil à états finis est bien encadrée dans les approches probabilistes de reconnaissance de séquences aléatoires. À partir des analyses préalables, il est constaté qu'une estimation adéquate des états permet une estimation précise de la consommation de la charge. Toutefois, cette estimation est conditionnée par une connaissance appropriée des paramètres du modèle statistique de la puissance et aussi des durées associées aux états [22],

[75]. Le réseau bayésien utilisé pour la modélisation de la charge à états finis correspond au modèle HMM à durée explicite illustrée à la Figure 3-7. Pour ce réseau spécifique, la probabilité jointe des séquences d'états cachés et de l'observation en considérant la durée est calculée avec le produit suivant :

$$\begin{aligned} P(\bar{x}, \bar{z}) &= P(\bar{z})P(\bar{x}|\bar{z}) \\ &= P(z_1)P(d_1|z_1)P(x_1|z_1) \prod_{k=2}^K P(z_k|z_{k-1}, d_{k-1})P(x_k|z_k) \end{aligned} \quad (3-22)$$

Où les facteurs sont définis selon les lois définies dans les sections préalables. Ils sont résumés dans la liste suivante :

- Probabilité initiale de l'état : $P(z_1) = \pi_{z_1}$
- Selon l'équation (3-15) : $P(d_1|z_1) = S(d_1|\phi_{z_1})/\mu_D^{z_1}$, où $\mu_D^{z_1}$ correspond à la durée moyenne de l'état initial z_1 .
- Vraisemblance de l'émission : $P(x_k|z_k) = f(x_k|\theta_{z_k})$
- Selon les équations (3-12) et (3-13) :

$$\begin{aligned} P(z_k|z_{k-1}, d_{k-1}) &= \begin{cases} 1 - h(d_{k-1}|\phi_{z_{k-1}}) & \text{si } z_k = z_{k-1} \\ h(d_{k-1}|\phi_{z_{k-1}}) & \text{si } z_k \neq z_{k-1} \end{cases} \\ &= (1 - \delta(z_k, z_{k-1}))h(d_{k-1}|\phi_{z_{k-1}}) + \delta(z_k, z_{k-1})(1 - h(d_{k-1}|\phi_{z_{k-1}})) \end{aligned} \quad (3-23)$$

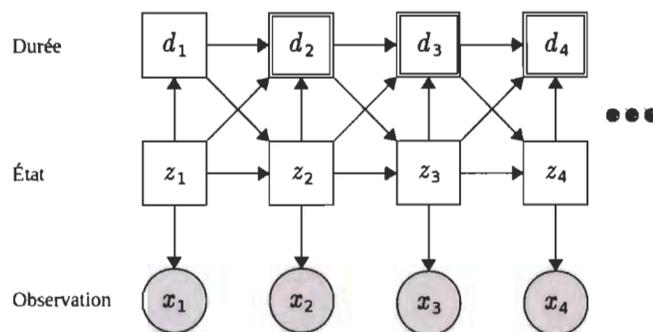


FIGURE 3-7 Réseau bayésien proposé pour le HMM à durée explicite.

Selon le modèle de probabilité (3-22), à chaque pas de temps k , le processus évolue selon des probabilités variantes en fonction de la durée de l'état. Dans la procédure d'estimation des états cachés, des compteurs du temps discret sont ajoutés pour chaque état afin d'évaluer les transitions hypothétiques des états entre deux instants consécutifs du temps discret. Il faut rappeler que dans le cadre général des modèles semi-markoviens la durée de l'état n'est pas une variable qui change en fonction des décisions sur les transitions, au contraire, elle est considérée comme une information non observée de l'état [80]. Afin de l'estimer, elle est normalement traitée comme un état caché, ce qui augmente de façon drastique le nombre de possibilités à évaluer [79]. Afin de gérer la complexité des modèles semi-markoviens, deux types d'approches approximatives sont proposées dans la littérature. La première est basée sur des techniques d'échantillonnage Monte-Carlo (MC) [38], [65], la deuxième est basée sur l'utilisation des estimateurs du temps de séjour en fonction de la variable de récursion [79], [86]. Cette dernière stratégie est plus cohérente avec le modèle de transition basé sur le compteur du temps d_{k-1} et la fonction de hasard $h(d_{k-1}|\phi_s)$. En effet, cette approximation considère que le temps moyen de séjour d'un état s peut être approché avec $\hat{d}_k^s = \beta_k^s \hat{d}_{k-1}^s + 1$, où β_k^s détermine la probabilité de l'état s à l'instant k sachant l'historique des données. Dans d'autres termes, on fait évoluer le compteur du temps de séjour d'un état spécifique selon la certitude de sa présence dans la séquence optimale à l'instant k .

Cependant, le temps de séjour n'a pas besoin d'être approximé. En fait, selon le modèle (3-8), le temps évolue de façon linéaire et déterministe par rapport aux situations présentées avec la variable d'état à l'instant k et $k - 1$. De surcroît, la composant aléatoire dans la dynamique du compteur correspond à l'événement de réinitialisation défini par le delta de Kronecker $\delta(z_k, z_{k-1})$ selon l'équation (3-9). Or, cette expression est calculable selon les possibilités de transition dans l'ensemble d'états possibles. Autrement dit, à l'instant k le temps de séjour de la charge dans l'état $s \in \{0, 1\}$ peut être récursivement calculé selon l'équation suivante :

$$d_k^s = \delta(s, q)d_{k-1}^s + 1 \quad (3-24)$$

Où q correspond au meilleur état précédent (à l'instant $k - 1$) connectant avec l'état s à l'instant k .

Dans la proposition [79], une estimation approximée des durées est nécessaire à cause d'une erreur d'interprétation des probabilités de transition homogènes ou invariantes dans le temps. En effet, dans la phase de modélisation, la probabilité de transition est définie par une loi postérieure prédictive de la forme $P(z_{k+1} = s | z_k = q, d_k^q)$, ce qui est avéré juste. Cependant, dans l'algorithme de programmation dynamique utilisé, cette probabilité a été calculée comme $P(z_k = s | z_{k-1} = q, d_k^q)$, d'où d_k^q est inconnu. Le problème se trouve dans le fait que d_k^q doit évoluer en fonction des valeurs de z_k et z_{k-1} , mais la valeur de z_k dépend aussi de d_k^q , ce qui résulte dans une boucle algébrique. Cette difficulté est contournée par une estimation de \hat{d}_k^q . Toutefois, en considérant l'invariance par translation dans le temps, cette probabilité doit garder la forme $P(z_k = s | z_{k-1} = q, d_{k-1}^q)$. Cette dernière est calculable sans approximation, car elle suit un chemin causal.

3.2.4.1 Maximisation du logarithme de la vraisemblance

Théoriquement, la solution optimale pour la séquence des états cachés consiste à trouver parmi toutes les solutions possibles la trajectoire qui maximise la probabilité (3-22) telle que définie par l'équation (3-2). En principe, il est possible d'utiliser la programmation dynamique pour maximiser ce produit. Cependant, cette analyse porte sur une longue multiplication successive de probabilités ce qui entraîne un problème au niveau de calcul numérique. En effet, les limites de la représentation en virgule flottante peuvent être atteintes très rapidement (dépassement de capacité par valeurs inférieures ou souppassement arithmétique) dans une séquence de quelques dizaines d'observations. Dans la pratique, il est préférable d'opérer le logarithme de la vraisemblance ou log-vraisemblance [63]. Puisque la fonction logarithme est strictement croissante, maximiser une vraisemblance $P(\vec{x}, \vec{z})$ est équivalent à maximiser son logarithme $\mathcal{L}(\vec{x}, \vec{z}) = \ln(P(\vec{x}, \vec{z}))$. Donc, à partir de la séquence d'émissions \vec{x} et ayant un modèle \mathcal{M} comprenant les distributions et ses paramètres, la meilleure séquence d'états cachés \vec{z}^* peut être estimée via le problème d'optimisation suivant :

$$\vec{z}^* = \arg \max_{\vec{z}} \mathcal{L}(\vec{x}, \vec{z} | \mathcal{M}) \quad (3-25)$$

Où $\mathcal{L}(\vec{x}, \vec{z}|\mathcal{M})$ est calculée à partir de l'équation (3-7) de la façon suivante :

$$\begin{aligned} \mathcal{L}(\vec{x}, \vec{z}|\mathcal{M}) = & \ln(\pi_{z_1}) + \ln(f(x_1|\theta_{z_1})) + \ln\left(\frac{S(d_1|\phi_{z_1})}{\mu_D^{z_1}}\right) \\ & + \sum_{k=2}^K (1 - \delta(z_k, z_{k-1})) \ln(h(d_{k-1}|\phi_{z_{k-1}})) \\ & + \sum_{k=2}^K \delta(z_k, z_{k-1}) \ln(1 - h(d_{k-1}|\phi_{z_{k-1}})) \\ & + \sum_{k=2}^K \ln(f(x_k|\theta_{z_k})) \end{aligned} \quad (3-26)$$

3.2.4.2 Programmation dynamique pour l'inférence des états

La programmation dynamique permet de déterminer la trajectoire d'états cachés qui explique le mieux la séquence d'observations à partir d'un modèle de tests. Par principe, la solution optimale d'un problème peut s'obtenir à partir des solutions optimales de sous problèmes. Ceci est connu comme le principe d'optimalité de Bellman « *toute politique optimale est composée de sous-politiques optimales* ». Cette technique appliquée au décodage de processus stochastiques à états finis est appelée algorithme de Viterbi. Ceci est normalement dérivé de la procédure de calcul progressif des probabilités (*en avant*).

Nous présentons par la suite l'algorithme de programmation dynamique en incluant le processus de comptage du temps de séjour et le calcul dynamique des probabilités de transitions. Cet algorithme suit la procédure classique définie par l'algorithme de Viterbi. Cependant, les coûts de transitions sont calculés à partir de la fonction de hasard discret $h(d_{k-1}|\phi_s)$. Il faut remarquer que si cette fonction est constante, l'algorithme est donc similaire à la méthode de Viterbi classique pour les modèles de Markov cachés.

Afin de simplifier la notation dans la suite du développement mathématique, les probabilités de transitions définies par le terme $P(z_k|z_{k-1}, d_{k-1})$ seront représentées par la forme matricielle suivante :

$$P(z_k|z_{k-1}, d_{k-1}) = A_{z_k, z_{k-1}} \quad (3-27)$$

Pour un modèle à deux états, celle-ci peut s'écrire sous la forme de matrice où les composantes sont définies à partir des modèles des durées comme suit :

$$A = \begin{bmatrix} 1 - h(d_{k-1}|\phi_0) & h(d_{k-1}|\phi_0) \\ h(d_{k-1}|\phi_1) & 1 - h(d_{k-1}|\phi_1) \end{bmatrix} \quad (3-28)$$

Étape récursive : La procédure avant, appliquée au décodage de séquences cachées est un jeu d'hypothèses dans lequel on se demande à chaque instant de temps k : si un état $s \in S$ fait partie de la séquence optimale quel serait son meilleur état précédent. En fait, nous choisissons toujours la transition d'état qui maximise la probabilité de la séquence pour tous les états à tous les temps. Dans ce cas, si la probabilité de la séquence passant par chacun des états précédents est dénotée par $\psi_{k-1,s}$, et les probabilités de transitions d'un état s vers un autre $q \in S$ sont données par la matrice $A_{q,s}$, le meilleur état précédent pour l'état s est estimé en résolvant le sous-problème d'optimisation suivant :

$$Q_{k,s} = \arg \max_{q \in S} (\ln(A_{q,s}) + \psi_{k-1,q}) \quad (3-29)$$

Pour démarrer cette procédure, il est nécessaire de calculer les valeurs des probabilités initiales $\psi_{1,s}$ pour tout $s \in S$. Celles-ci sont estimées en utilisant la distribution de l'état initial π_s et la vraisemblance de la première observation x_1 pour chacun des états :

$$\psi_{1,s} = \ln(\pi_s) + \ln(f(x_1|\theta_s)) \quad (3-30)$$

En connaissant les meilleures connexions entre les états, nous pouvons alors mettre à jour la probabilité de la séquence pour chaque état en incluant la probabilité sur l'observation à l'instant k :

$$\psi_{k,s} = \ln(f(x_k|\theta_s)) + \ln(A_{s,q^*}) + \psi_{k-1,q^*} \quad (3-31)$$

Où $q^* = Q_{k,s}$.

Par ailleurs, le calcul de la matrice de transition demande l'estimation de la durée hypothétique de chaque état à l'instant k tel que défini dans l'équation (3-8). Dans ce modèle de durée, le compteur du temps discret est réinitialisé quand un changement d'état est détecté.

Pour le cas d'estimation des états cachés, la dynamique de ce compteur est reliée à la variation entre le meilleur état précédent et l'état présent. Autrement dit, si $Q_{k,s}$ n'est pas égale à s , cela signifie que la meilleure décision consiste à changer d'état. Par conséquent la durée doit être réinitialisée. Le compteur incrémental associé à la durée de chaque état évolue de la façon suivante :

$$d_s = \delta(q^*, s) d_{q^*} + 1 \quad (3-32)$$

Détermination du meilleur chemin : La procédure avant permet de calculer de façon optimale la probabilité des séquences passant par chacun des états. Cependant, il faut un deuxième passage pour choisir la séquence optimale. Cette procédure permet de déterminer la séquence à partir de la connaissance de l'état final estimé à partir des probabilités $\psi_{K,s}$ [79]. Pour calculer l'état final optimal, il faut choisir celui, que maximise la probabilité finale de la séquence :

$$z_K = \arg \max_{s \in S} (\psi_{K,s}) \quad (3-33)$$

Une fois que l'état final optimal est calculé, le reste des éléments de la séquence est choisi en prenant le meilleur état précédent :

$$z_k = Q_{k,z_{k+1}} \quad k = K - 1, K - 2, \dots, 2, 1 \quad (3-34)$$

Le pseudo-code de l'algorithme (3.2) résume toute la procédure de décodage optimal d'un profil à deux états. Contrairement à l'algorithme de Viterbi classique, ceci permet de décoder des séquences discrètes considérant les durées sur les états comme information pertinente pour le calcul de probabilités de transition. Il faut remarquer que cet algorithme n'ajoute pas de la complexité de calcul par rapport à l'algorithme de Viterbi des HMMs conventionnels, laquelle est de l'ordre $O(|S|^2 K)$, où $|S|$ est le nombre d'états, $|S| = 2$ pour le cas des modèles à deux états. Il faut noter que, le modèle général pour analyser les processus stochastiques à durée d'état explicite ayant un nombre arbitraire d'états est formulé dans le cadre de modèles semi-markoviens. Ceux-ci ont une complexité de l'ordre $O(|S|^2 K^2 D_{max})$, où D_{max} correspondent à la durée maximale d'un état [87], ce qui donne un temps de calcul très élevé pour un ordinateur classique.

Algorithm 3.2: Algorithme de décodage pour un HMM à Durée Explicite (EDHMM) à deux états

Input:

Séquence d'observations : x_1, x_2, \dots, x_K

Paramètres des modèles sur les durées et l'émissions : $\{\phi_s, \theta_s, \tau_s\}_{s=0,1}$

Output: Séquence d'états cachés : z_1, z_2, \dots, z_K

```

1 begin
2    $S = \{0, 1\}$ 
3   Définir les matrices  $\psi, Q$  avec dimension  $K \times |S|$  et un vecteur  $d$  de dimension
    $|S| \times |S|$ 
4   for  $s \in S$  do ▷ Initialisation du treillis
5      $\pi_s = \frac{\tau_s}{\sum_q \tau_q}$ 
6      $\psi_1[s] = \log(\pi_s) + \log(f(x_1 | \theta_s))$ 
7      $d_s = \tau_s$ 
8   for  $k = 2, 3, \dots, K$  do ▷ Étape récursive
9     Estimer la matrice  $A$  selon l'équation 3-28
10    for  $s \in S$  do
11       $q^* = \arg \max_{q \in S} (A_{q,s} + \psi_{k-1,q})$ 
12       $\psi_{k,s} = \log(f(x_k | \theta_s)) + A_{q^*,s} + \psi_{k-1,q^*}$ 
13       $d_s = \delta(q^*, s) d_{q^*} + 1$ 
14       $Q_{k,s} = q^*$ 
15     $z_K = \arg \max_{s \in S} (\psi_{K,s})$ 
16    for  $k = K, K-1, K-2, \dots, 1$  do ▷ Détermination du meilleur chemin
17       $z_k = Q_{k,z_{k-1}}$ 

```

3.2.4.3 Étiquetage de séquences

La méthode statistique utilisée offre un mécanisme simple pour la classification et l'étiquetage de profils non agrégés. En considérant les modèles et paramètres associés à chaque profil individuel comme un ensemble d'information \mathcal{M}_m , le processus de classification, dans le sens du maximum de vraisemblance (MLE), est défini avec le problème suivant :

$$\hat{m} = \arg \max_m [P(\vec{x}, \vec{z} | \mathcal{M}_m)] \quad (3-35)$$

Ce problème permet de choisir de façon simple l'index de l'ensemble des modèles candidats. Cependant, des analyses supplémentaires sur la probabilité jointe de la séquence des états et

des émissions doivent être considérées afin de rejeter ou accepter les classifications obtenues dans la solution de (3-35). Plus précisément, des seuils de rejet doivent être définis par rapport à une vraisemblance minimale dans l'évaluation des séquences. C'est à dire, on accepte l'étiquette \hat{m} , si $P(\vec{x}, \vec{z} | \mathcal{M}_{\hat{m}}) > \eta_m$, où η_m est un seuil associé au modèle du profil m . Les valeurs pour ces seuils seront discutées plus loin.

En outre, pour un profil composé de plusieurs cycles, la vraisemblance décroît exponentiellement tandis que la log-vraisemblance le fait de façon linéaire. Par conséquent, une division par rapport à la longueur doit stabiliser la log-vraisemblance d'une longue séquence. À cet effet, la moyenne de log-vraisemblance d'une séquence de longueur K peut s'exprimer selon $\frac{1}{K} \mathcal{L}(\vec{x}, \vec{z} | \mathcal{M}_m)$. La classification des profils peut être redéfinie à partir du modèle suivant :

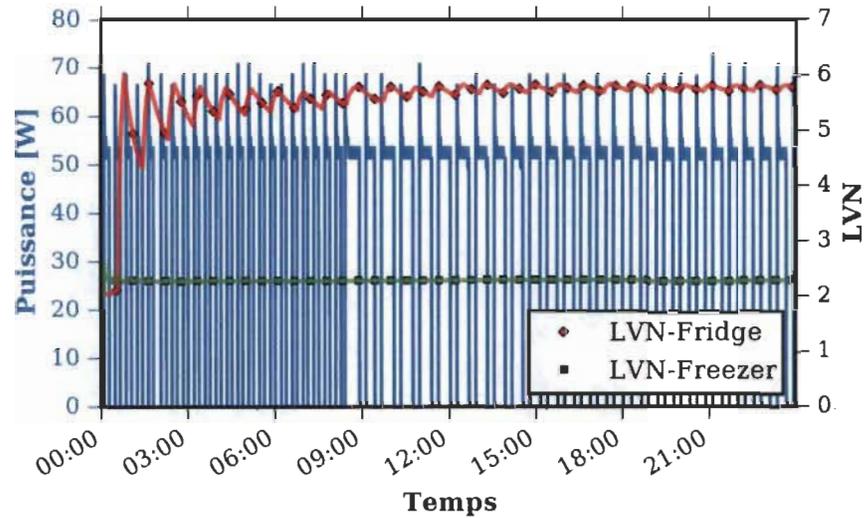
$$\hat{m} = \arg \min_m \left[-\frac{1}{K} \mathcal{L}(\vec{x}, \vec{z} | \mathcal{M}_m) \right] \quad (3-36)$$

La Figure 3-8 montre des résultats d'analyse croisée entre deux profils de charges thermostatiques. En effet, un profil thermostatique de réfrigérateur et un autre de congélateur pris de la base [85] ont été utilisés. Dans cette expérience, il est constaté que la moyenne de log-vraisemblance (LVN) reste stable tout au long de la séquence optimale pour chacun des modèles de test. On peut observer que la log-vraisemblance est stabilisée à partir de la division par le temps k . Cette procédure permettra de définir des seuils de rejet indépendants de la longueur de la séquence.

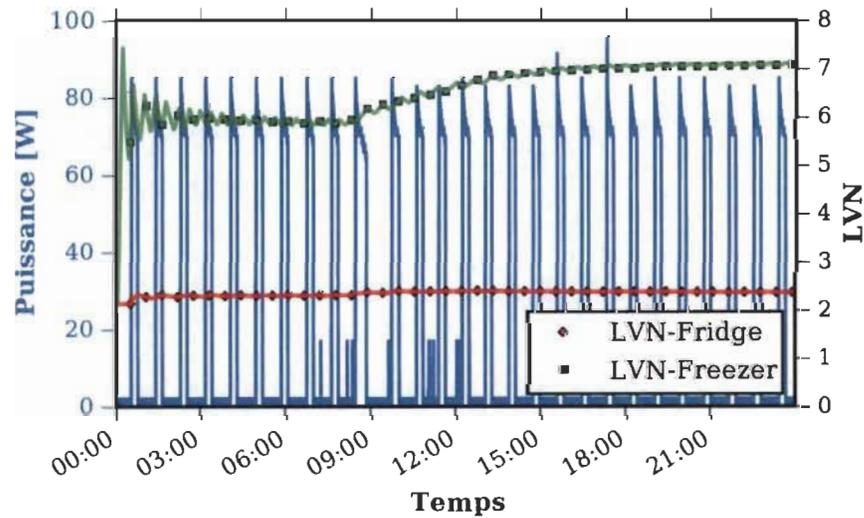
3.2.4.4 Probabilités stationnaires des états thermostatiques

Afin d'introduire l'information du type fréquentiel dans l'analyse du comportement des charges ayant un comportement caractérisé par des périodicités, nous allons décrire les paramètres des lois de transition en fonction de la période ou bien de la fréquence du cycle. La période espérée en temps discret d'un cycle du profil thermostatique peut s'exprimer comme la somme de valeurs espérées des états individuels du profil :

$$\tau = \tau_0 + \tau_1 \quad (3-37)$$



(a) Profil du congélateur



(b) Profil du réfrigérateur

FIGURE 3-8 Log-vraisemblance des profils individuels de la puissance de deux charges thermostatiques.

Où τ_0 et τ_1 correspondent aux moyennes des durées discrètes des états "OFF" et "ON" respectivement. Ces paramètres permettent de calculer la probabilité stationnaire des états. En effet, ces probabilités sont définies comme les proportions du temps que la charge demeure dans un état spécifique par rapport au temps d'analyse. En considérant la période du cycle τ comme temps d'analyse, la probabilité stationnaire est déterminée par l'équation suivante :

$$\pi_s = \frac{\tau_s}{\tau} \quad (3-38)$$

Pour les charges thermostatiques, cette quantité peut être associée au rapport cyclique des thermostats. Pour illustrer expérimentalement le calcul des probabilités stationnaires, nous avons traité le profil de la Figure 3-5. Ceci correspond au profil thermostatique de puissance active d'un réfrigérateur mesuré pendant une journée complète. Dans ce profil, les durées moyennes des états $\tau_0\Delta t$ et $\tau_1\Delta t$ correspondent à 2285 et 830 secondes respectivement (Δt est la période d'échantillonnage). Le calcul des probabilités de l'équation (3-38) donne comme résultats les valeurs $\pi_0 = 0.73$ et $\pi_1 = 0.27$. Ceci indique que cet appareil demeure 27% du temps dans l'état "ON" et le 73% restant dans l'état "OFF".

À partir des analyses supplémentaires, il est possible d'étudier le lien entre les paramètres temporaires et l'analyse fréquentielle. Ceci nous permettra de faire le chemin inverse, c'est-à-dire, faire certaines interférences sur les paramètres des distributions à partir des résultats des analyses fréquentielles. Ce travail est important pour pouvoir détecter l'existence des charges thermostatiques ayant des signatures spectrales assez particulières et pour pouvoir créer postérieurement un modèle probabiliste approprié dédié au décodage des états cachés des charges FSM. Dans un premier temps les thermostats électroniques à fréquence de commutation fixe sont étudiés. Toutefois, les approches proposées resteront suffisamment générales pour des charges thermostatiques pouvant avoir une fréquence de commutation soumise à des fluctuations considérables. De façon générale, pour un profil de type I, la fréquence moyenne f_c est estimable à partir des paramètres statistiques des durées :

$$f_c = \frac{1}{\tau} \quad (3-39)$$

3.2.4.5 Puissance et énergie moyenne d'un profil thermostatique

La consommation moyenne de la charge pendant un état précis peut être estimée en calculant l'espérance de la distribution de la puissance pour l'état en question :

$$\mu_s = \mathbb{E}(x_k | z_k = s) = \sum_x x f(x | \theta_s) \quad (3-40)$$

Par ailleurs, la puissance moyenne d'un cycle d'opération d'une machine à deux états peut être calculée en combinant linéairement les probabilités de la distribution stationnaire des

états (équation (3-40)) et les valeurs espérées pour la consommation de chaque état (équation (3-38)). En considérant $p = [\pi_0 \ \pi_1]^T$ la distribution stationnaire du cycle, et $u = [\mu_0 \ \mu_1]$ le vecteur de consommation pour les états du cycle, l'énergie consommée par une machine pendant une période de temps donnée t est :

$$E_t = t u p = t \sum_s \pi_s \mu_s \quad (3-41)$$

Afin d'évaluer le processus d'estimation du profil thermostatique, nous avons calculé l'erreur de reconstruction pour le profil du réfrigérateur étudié dans cette section. Les valeurs correspondant aux consommations des états "ON" et "OFF" μ_0 et μ_1 sont 0.48 et 76.07 watts respectivement. Pour cet appareil, la période moyenne du cycle thermostatique est $\tau \Delta t = 3111$ secondes, les temps moyens des états "ON" et "OFF" sont $\tau_0 \Delta t = 2271$ et $\tau_1 \Delta t = 840$ secondes respectivement. Ce qui permet de calculer la distribution stationnaire des états comme $p = [0.73, 0.27]$. À partir de ces informations, l'énergie consommée dans une journée en est $\hat{E}_{24h} = 20.62 \times 24 = 490$ wattheures (Wh). En comparant avec l'énergie réelle mesurée pour la journée d'analyse ($E_{24h} = 495$ Wh), l'erreur relative absolue en pourcentage ($100\% |\hat{E}_{24h} - E_{24h}| / E_{24h}$) est égale à 0.82%.

3.3 Agrégation de modèles et détection parallèle des états

Dans cette section, le concept de charge à multiples états comme une composition de profils élémentaires de type I est développé. À partir des éléments théoriques présentés préalablement, un modèle composé est construit. Ceci comprend les éléments nécessaires pour accomplir le processus de reconnaissance et reconstruction de profils créés par la combinaison de multiples thermostats. Ce développement cherche à caractériser la signature macroscopique des charges de type II. En effet, dans l'agrégation probabiliste de la puissance de ces machines, l'observation est une conséquence de la combinaison des états d'opération individuels. L'hypothèse principale adoptée dans l'analyse séquentielle consiste à considérer une indépendance fonctionnelle entre les différents éléments du mélange.

Étant donné un ensemble d'éléments thermostatiques fonctionnant simultanément, l'agrégation de la puissance est une conséquence des combinaisons des états ON-OFF des thermostats [88]. Pour représenter ces combinaisons en tant qu'états composés ou super-états d'une charge FSM, considérons $\mathbf{z}_k = [z_k^1 \ z_k^2 \ \dots \ z_k^M]$ un vecteur de dimension M contenant les valeurs instantanées des états de M profils de type I. Dénotons par y_k la puissance demandée par l'ensemble de charges à l'instant k . Conséquemment, la probabilité jointe d'une séquence d'observations agrégées $\vec{y} = \{y_k\}_{k=1,\dots,K}$ et une séquence multidimensionnelle d'états composés $\vec{z} = \{\mathbf{z}_k\}_{k=1,\dots,K}$ est définie par $P(\vec{z}, \vec{y}) = P(\vec{z})P(\vec{y}|\vec{z})$. Cette probabilité peut se factoriser selon l'équation suivante :

$$P(\vec{y}, \vec{z}) = P(\mathbf{z}_1)P(y_1|\mathbf{z}_1)P(\mathbf{d}_1|\mathbf{z}_1) \prod_{k=2}^K P(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{d}_{k-1})P(y_k|\mathbf{z}_k) \quad (3-42)$$

Où $\mathbf{d}_{k-1} = [d_{k-1}^1 \ d_{k-1}^2 \ \dots \ d_{k-1}^M]^T$ est un vecteur colonne associé aux durées des états de chacun des sous-profils.

L'objectif de la reconnaissance parallèle est encadré dans un problème d'optimisation où on cherche à maximiser le logarithme de la probabilité jointe des séquences \vec{y} et \vec{z} . En suivant le même développement méthodologique utilisé pour les charges du type I, on définit le problème d'optimisation suivante :

$$\vec{z}^* = \arg \max_{\vec{z}} \mathcal{L}(\vec{x}, \vec{z}) \quad (3-43)$$

Où $\mathcal{L}(\vec{x}, \vec{z})$ désigne la fonction à maximiser. Celle-ci adopte la forme suivante :

$$\begin{aligned} \mathcal{L}(\vec{x}, \vec{z}) &= \ln(P(\vec{y}, \vec{z})) \\ &= \ln(P(\mathbf{z}_1)) + \ln(P(y_1|\mathbf{z}_1)) + \ln(P(\mathbf{d}_1|\mathbf{z}_1)) \\ &\quad + \sum_{k=2}^K \ln(P(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{d}_{k-1})) + \ln(P(y_k|\mathbf{z}_k)) \end{aligned} \quad (3-44)$$

Le vecteur des durées évolue selon la récursion suivante :

$$\mathbf{d}_k = \delta(\mathbf{z}_k, \mathbf{z}_{k-1})\mathbf{d}_{k-1} + \mathbf{1}$$

Où $\delta(\mathbf{z}_k, \mathbf{z}_{k-1})$ est un vecteur ligne contenant les deltas de Kronecker pour chacun des sous-profiles à détecter : $\delta(\mathbf{z}_k, \mathbf{z}_{k-1}) = [\delta(z_k^1, z_{k-1}^1) \delta(z_k^2, z_{k-1}^2) \dots \delta(z_k^M, z_{k-1}^M)]$.

Afin de fournir les modèles et les expressions pour les facteurs de l'équation (3-42), nous étudierons d'abord la probabilité jointe des états $P(\mathbf{z}_k | \mathbf{z}_{k-1}, \mathbf{d}_{k-1})$. Postérieurement, une loi $P(y_k | \mathbf{z}_k)$ pour les émissions agrégées est développée. La séquence d'états cachés est estimée à partir de procédures de programmation dynamique "en avant" comme il a été proposé pour le cas d'un profil à deux états.

3.3.1 Transitions entre super-états

Pour la représentation de la combinaison des états, nous utiliserons le concept de super-état. Ceci est un élément d'un ensemble généré à partir de la combinaison des petits ensembles binaires $\{0, 1\}$ définissant le support de la variable d'état des profils individuels. Pour ce faire, l'ensemble de super-états peut être construit à partir de la codification binaire de l'ensemble de nombres entiers $S = \{0, 1, \dots, 2^M - 1\}$ représentant chaque une des combinaisons engendrées par l'opération parallèle des éléments à deux états. L'ensemble des super-état binaires est défini par $\mathbf{S} = \{[\text{BCD}(s)]\}_{s=1 \dots 2^M - 1}$, où $\text{BCD}(s)$ désigne le processus de conversion binaire du nombre entier s . Par exemple, deux éléments thermostatiques formeront l'ensemble de super-états $\mathbf{S} = \{[00], [01], [10], [11]\}$.

D'autre part, à partir de l'hypothèse d'indépendance entre les processus individuels, la probabilité de transition entre super-états peut se calculer avec le produit : $\prod_{m=1}^M P(z_k^m | z_{k-1}^m, d_{k-1}^m)$. Évidemment, chacun des facteurs correspond à la probabilité de transition individuelle calculée à partir de la fonction de hasard discret $h(d_{k-1}^m | \phi_{z_{k-1}}^m)$ tel que défini dans l'équation (3-23). À cet égard, si les transitions individuelles suivent une loi de Bernoulli, l'ensemble des transitions entre super-états suivent une loi de Bernoulli multidimensionnelle où la probabilité de transition entre les super-états $\mathbf{q} \in \mathbf{S}$ et $\mathbf{s} \in \mathbf{S}$ est donnée par :

$$P(\mathbf{z}_k = \mathbf{s} | \mathbf{z}_{k-1} = \mathbf{q}, \mathbf{d}_{k-1}) = \prod_{m=1}^M P(z_k^m = s_m | z_{k-1}^m = q_m, d_{k-1}^m) \quad (3-45)$$

Où chaque facteur est estimé en fonction des éléments individuels.

$$P(z_k^m = s_m | z_{k-1}^m = q_m, d_{k-1}^m) = (1 - \delta(s_m, q_m))h(d_{k-1}^m | \phi_{q_m}) + \delta(s_m, q_m)(1 - h(d_{k-1}^m | \phi_{q_m})) \quad (3-46)$$

3.3.2 Émissions agrégées

L'analyse du problème d'agrégation instantanée de la variable y_k est basée sur le modèle de la vraisemblance $P(y_k | \mathbf{z}_k)$. Notamment, on s'intéresse à évaluer la façon dont cette distribution est affectée par la combinaison d'états \mathbf{z}_k . Il est à noter que dans les HMMs conventionnels une combinaison d'états devient un état. Il convient par conséquent d'observer la puissance associée à chaque combinaison. Afin d'éviter une longue et incertaine phase d'entraînement pour chaque état, l'analyse factorielle permet de prédire le comportement et la puissance demandée d'un super-état sans la nécessité d'une observation expérimentale. Cette propriété représente un avantage considérable par rapport aux autres techniques de l'apprentissage de machine dans NILM. En fait, la désagrégation est naturellement un problème combinatoire, et la plupart de classificateurs sont spécialisés dans la reconnaissance du profil individuel et n'ont pas la capacité de prédire de façon explicite les émissions des combinaisons quand elles n'ont pas été utilisées dans le processus d'entraînement. En effet, la puissance agrégée correspond à la combinaison linéaire des puissances individuelles comme définie par l'équation (2-1), la valeur instantanée de cette puissance est $y_k = x_k^1 + x_k^2 + \dots + x_k^M$. Afin d'illustrer cette agrégation, un profil de puissance agrégée est tracé à la Figure 3-10. Ceci correspond à l'addition des quatre profils thermostatiques montrés à la Figure 3-9. Ces données ont été acquises au laboratoire à une fréquence d'échantillonnage de 4 Hertz.

Par ailleurs, nous avons montré que les probabilités des super-états $P(\mathbf{z}_k)$ peuvent être factorisées sous l'hypothèse d'indépendance. Cependant, cette hypothèse n'est pas applicable pour la probabilité conditionnelle $P(y_k | \mathbf{z}_k)$. Ceci à cause la somme des puissances y_k est une conséquence directe des combinaisons des états. Sur ce point, étant donné y_k en tant qu'une somme des variables indépendantes, il est donc possible d'avoir une estimation de sa distribution conditionnelle (sachant le super-état). En théorie, cette distribution est calculée par le produit de convolution entre les distributions individuelles. En effet, si l'ensemble de

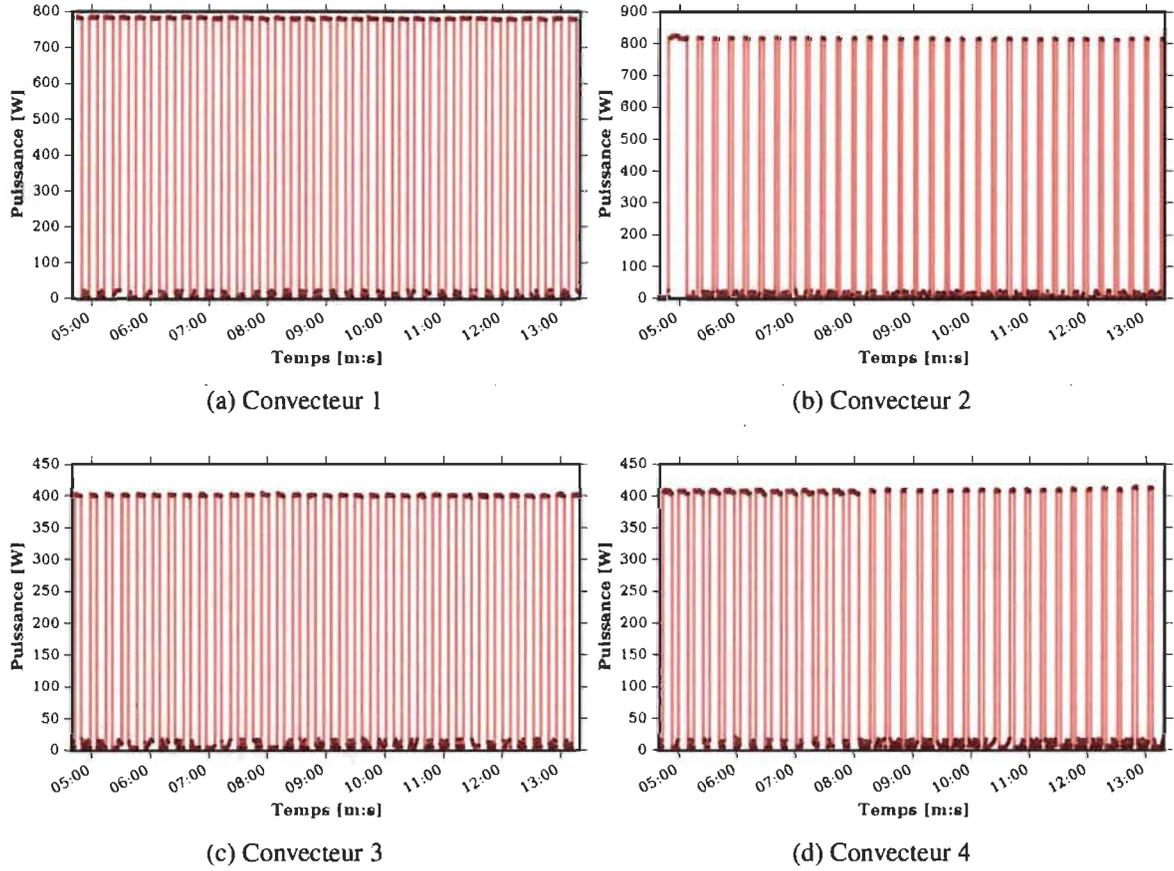


FIGURE 3-9 Profils individuels de la puissance de quatre convecteurs électriques contrôlés par de thermostats électroniques.

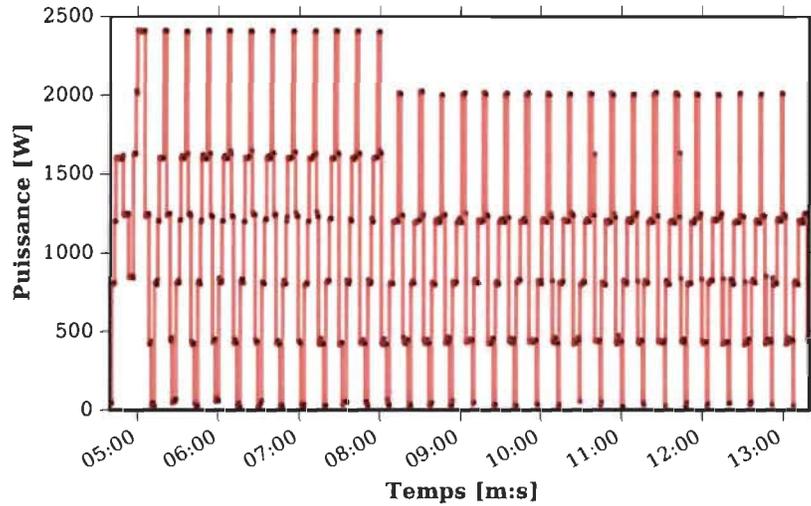


FIGURE 3-10 Profil FSM résultant de l'agrégation des quatre profils thermostatiques affichés à la Figure 3-9.

charges à deux états se trouve dans une combinaison d'états $\mathbf{s} = [s_1 \ s_2 \ \dots \ s_M]$, où $s_m \in \{0, 1\}$, la distribution conditionnelle pour la somme de puissances individuelles est définie par le produit de convolution suivant :

$$f(y|\theta_{\mathbf{s}}) = f(x|\theta_{s_1}^1) \otimes f(x|\theta_{s_2}^2) \otimes \dots \otimes f(x|\theta_{s_M}^M) \quad (3-47)$$

Avec $\theta_{\mathbf{s}}$, l'ensemble de paramètres de la distribution résultante.

L'équation (3-47) est générale et valide pour le cas des distributions discrètes (fonction de masse) et continues (fonction de densité). Elle est aussi applicable aux lois paramétriques ou non si elles existent. Si une loi paramétrique est adaptée pour décrire la distribution de la puissance d'une charge, et s'il existe une solution analytique de l'équation (3-47) pour les lois en question, il sera possible alors de calculer les paramètres globaux du mélange à partir d'une combinaison des paramètres individuels. Dans le domaine NILM, la modélisation paramétrique utilise fréquemment la loi normale comme approximation des densités conditionnelles. Conséquemment, à partir des analyses présentées préalablement, la loi composée résultante d'une somme de variables aléatoires indépendantes suivant une distribution normale est aussi gaussienne avec paramètres définis selon l'équation (3-49) et (3-50) :

$$f(y|\theta_{\mathbf{s}}) = \mathcal{N}(x|\mu_{s_1}^1, v_{s_1}^1) \otimes \mathcal{N}(x|\mu_{s_2}^2, v_{s_2}^2) \otimes \dots \otimes \mathcal{N}(x|\mu_{s_M}^M, v_{s_M}^M) = \mathcal{N}(y|\mu_{\mathbf{s}}, v_{\mathbf{s}}) \quad (3-48)$$

Où $\theta_{\mathbf{s}} = \{\mu_{\mathbf{s}}, v_{\mathbf{s}}\}$. De façon générale, pour une somme des variables aléatoires indépendantes, les valeurs moyennes et les variances des distributions résultantes correspondent aux combinaisons linéaires des moyennes et variances individuelles :

$$\mu_{\mathbf{s}} = \mathbb{E}[y_k | \mathbf{z}_k = \mathbf{s}] = \sum_{m=1}^M \mu_{s_m}^m \quad (3-49)$$

$$v_{\mathbf{s}} = \mathbb{E} \left[((y_k | \mathbf{z}_k = \mathbf{s}) - \mu_{\mathbf{s}})^2 \right] = \sum_{m=1}^M v_{s_m}^m \quad (3-50)$$

La Figure 3-11 montre trois distributions agrégées relatives aux états combinés des profils de puissance illustrés à la Figure 3-9. Notons que pour chaque combinaison d'états \mathbf{s} , une nouvelle distribution est engendrée selon l'équation (3-47). Dans cet exemple, un total de $2^4 =$

16 différentes distributions sont obtenues à partir du produit de convolution des distributions non paramétriques et aussi par l'approximation à une loi normale. Les distributions résultantes sont comparées avec la distribution empirique des données observées pour les combinaisons évaluées. Dû au nombre de combinaisons, nous illustrons seulement trois de ces distributions correspondant aux super-états $\mathbf{s} = [0\ 0\ 0\ 0]$, $\mathbf{s} = [0\ 1\ 1\ 1]$ et $\mathbf{s} = [1\ 1\ 1\ 1]$ respectivement.

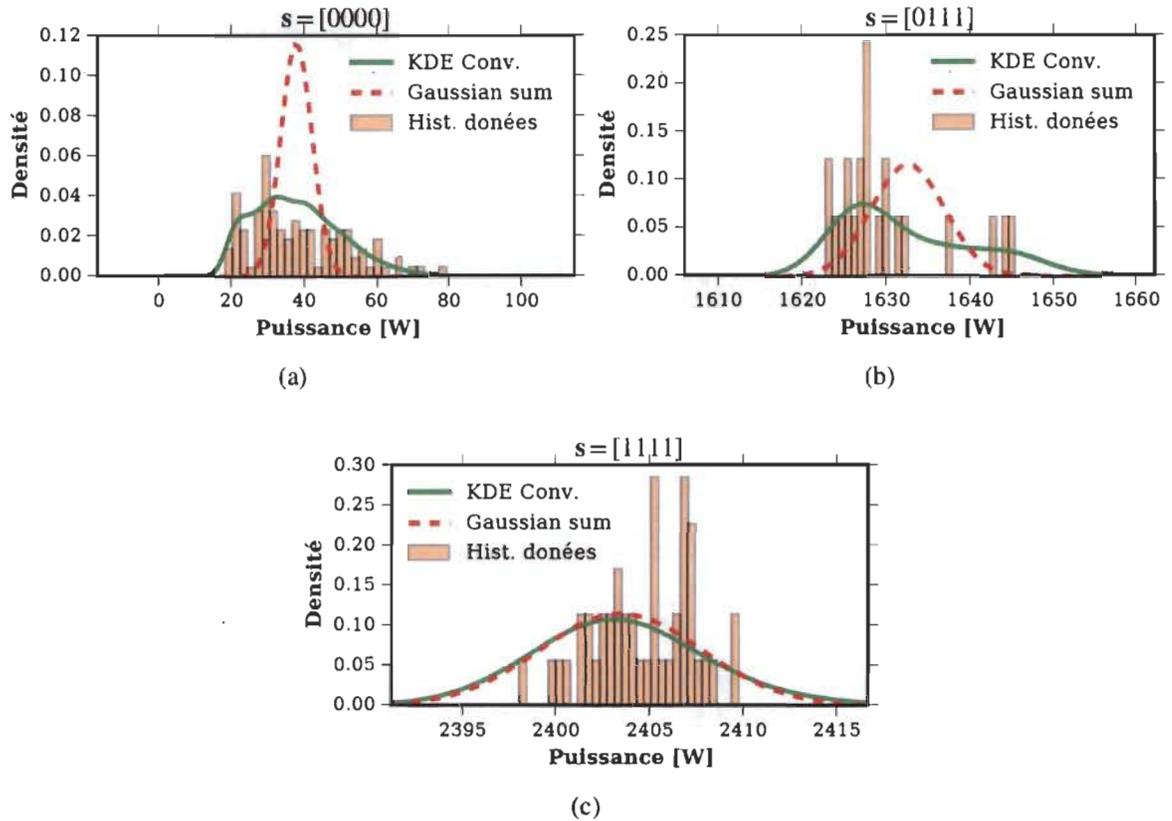


FIGURE 3-11 Distributions conditionnelles de la puissance agrégée de quatre éléments thermostatiques.

3.3.3 Pré-allocation des probabilités de transition avec les opérations de Kronecker

Des études sur les modèles markoviens factoriels proposent l'utilisation du produit de Kronecker. En tant que cas particulier du produit tensoriel, ce produit abouti à une technique compacte pour calculer les distributions résultantes de plusieurs processus stochastiques évoluant en parallèle [89], [90]. À partir de ces opérations, on propose une approche qui permet la construction d'un modèle à multiples états à partir de la combinaison des matrices de

transitions individuelles. De façon générale, si la probabilité de transition entre les super-états est calculée à partir de la loi de Bernoulli multidimensionnelle, une représentation avec le produit de Kronecker permet de calculer de façon matricielle les combinaisons possibles dans une transition entre super-états [91]. À ce propos, on définit une matrice de transition $\mathbf{A} = \{a_{q,s}\}$ dont chaque élément représente la probabilité de transition $a_{q,s} = P(\mathbf{z}_k = \mathbf{s} | \mathbf{z}_{k-1} = \mathbf{q}, \mathbf{d}_{k-1}^q)$. Les lignes de cette matrice sont construites à partir d'un produit de Kronecker (symbolisé par l'opérateur \otimes) entre les probabilités de transition individuelles :

$$\mathbf{A}[q, :] = \bigotimes_{m=1}^M \mathbf{A}_q^m[\mathbf{q}[m], :], \quad \forall q \in \mathcal{S} \quad (3-51)$$

Où \mathbf{A}_q^m corresponde à la matrice de transition du profil m associé au super-état avec indice q . Cette matrice est construite comme il suit :

$$\mathbf{A}_q^m = \begin{bmatrix} 1 - h(\mathbf{d}_{k-1}^q[m] | \phi_0^m) & h(\mathbf{d}_{k-1}^q[m] | \phi_0^m) \\ h(\mathbf{d}_{k-1}^q[m] | \phi_1^m) & 1 - h(\mathbf{d}_{k-1}^q[m] | \phi_1^m) \end{bmatrix} \quad (3-52)$$

Il faut noter qu'un total de M hypothétiques durées doit être assigné à chaque super-état afin de calculer la probabilité de transition selon l'équation (3-45). Sur ce point, la gestion des durées est faite à partir d'un passage des valeurs entre les super-états précédents et les super-états présents. C'est à dire, si à l'instant k le meilleur super-état précédent pour \mathbf{s} est \mathbf{q} , le vecteur des durées assigné au super-état \mathbf{s} doit évoluer selon l'équation suivante :

$$\mathbf{d}_k^s = \delta(\mathbf{s}, \mathbf{q}) \mathbf{d}_{k-1}^q + \mathbf{1} \quad (3-53)$$

Effectivement, le produit de Kronecker facilite l'évaluation de combinaisons probabilistes. Ceci a été étudié dans la construction des modèles markoviens factoriels composés de multiples modèles subjacents indépendants. Le produit de Kronecker est une façon alternative et compacte pour représenter les distributions résultantes de plusieurs processus stochastiques évoluant en parallèle [89], [90]. Afin d'illustrer la relation entre le produit de Kronecker et l'agrégation des modèles de probabilité, prenons comme exemple le cas des distributions avec probabilités $p_1 = [p_0^1, p_1^1]$ et $p_2 = [p_0^2, p_1^2]$, le produit de Kronecker entre ces deux vecteurs

détermine les probabilités des combinaisons $\{00, 01, 10, 11\}$. Celles-ci sont contenues dans la distribution résultante $\mathbf{p} = p_1 \otimes p_1 = [p_0^1 p_0^2, p_0^1 p_1^2, p_1^1 p_0^2, p_1^1 p_1^2]$. Une exemple de calcul de ce produit est illustré à la Figure 3-12. Dans celle-ci, la distribution stationnaire des super-états générés par 4 profils thermostatiques (Figure 3-12b) a été calculée à partir du produit de Kronecker entre les distributions individuelles (Figure 3-12a).

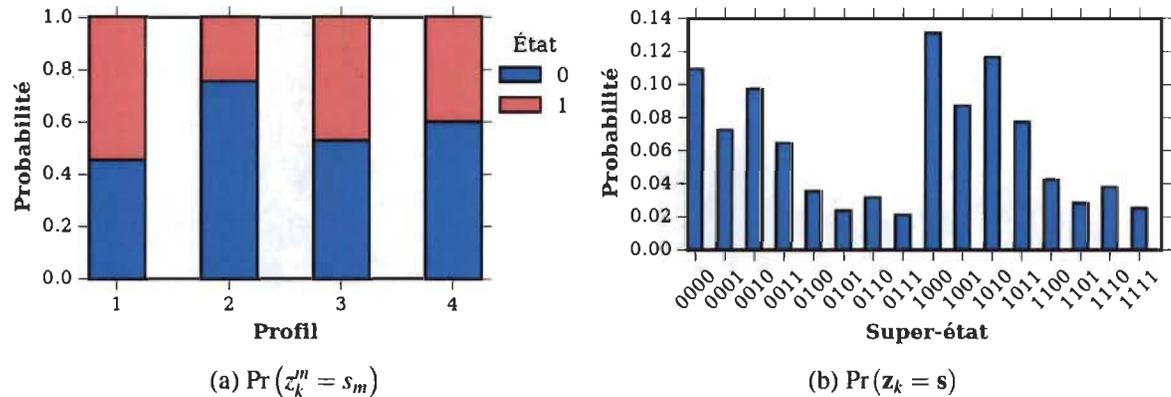


FIGURE 3-12 Résultats du produit de Kronecker entre quatre distributions stationnaires. a) Distributions stationnaires de quatre charges thermostatiques. b) Distribution stationnaire des super-états thermostatiques.

Dans cette partie de la proposition, nous avons introduit l'expression de l'équation (3-51) comme une variante de la notation conventionnelle du produit de Kronecker pour la construction de matrices de transition de modèles de Markov factoriels. En fait, si les modèles individuels sont markoviens, la matrice de transitions entre super-états est calculée par le produit de Kronecker entre les matrices individuelles : $\mathbf{A} = \otimes_{m=1}^M A_m$. Cependant, ceci n'est pas possible avec des probabilités de transition dynamiques, car chaque super-état a un vecteur de durées \mathbf{d}_{k-1}^s différent. Toutefois, dans le cas markovien, les probabilités de transition restent invariants par rapport à la durée. Il faut remarquer que chaque ligne d'une matrice de transition représente une distribution conditionnelle. C'est à dire, la probabilité de l'état présent sachant l'état passé. Conséquemment, les distributions pour les super-états doivent être calculées avec l'information apportée par le super-état précédent. Dans le modèle semi-markovien, cette information est associée aux durées et chaque super-état engendre une durée hypothétiquement différente pour les éléments individuels ; ce qui empêche l'utilisation de l'opérateur de Kronecker dans la construction de colonnes. En revanche, le constructeur

défini par (3-51) est général, permettant aussi la manipulation des probabilités de transition markoviennes et non markoviennes.

Il faut souligner que les produits de Kronecker ne diminuent pas la complexité de calcul dans les procédures de programmation dynamique. Par contre, celles-ci facilitent la pré-allocation des matrices de transition en évitant un calcul redondant des fonctions de hasard. En effet, en utilisant directement l'équation (3-45) pour calculer toutes les probabilités de transition, à chaque instant du temps, un total de $M \times 2^M \times 2^M$ évaluations des fonctions $h(d_{k-1}^m | \phi_{s[m]}^m)$ seront nécessaires. Par contre, le produit de Kronecker fournit ces probabilités en utilisant $2^M \times 2^M$ opérations. De plus, les opérations de Kronecker peuvent être facilement utilisées dans le calcul de la log-vraisemblance. Évidemment, appliquer le logarithme sur le résultat d'un produit de Kronecker est équivalent à appliquer une addition de Kronecker sur ces logarithmes. Encore plus, les combinaisons des paramètres pour les émissions gaussiennes de super-états peuvent être aussi calculées par des additions de Kronecker. Dans cette perspective, les vecteurs des moyennes et des variances paramétrant toutes les combinaisons (sommées de gaussiennes) sont calculés à partir des additions de Kronecker. À cet égard, étant donné les vecteurs paramètres individuels $\mathbf{u}_m = [\mu_0^m, \mu_1^m]$, $\mathbf{v}_m = [v_0^m, v_1^m]$ comprenant les moyennes et les variances respectivement, les vecteurs des paramètres regrésés sont calculés comme il suit :

$$\mathbf{u} = \bigoplus_{m=1}^M \mathbf{u}_m, \quad \mathbf{v} = \bigoplus_{m=1}^M \mathbf{v}_m \quad (3-54)$$

À partir des vecteurs \mathbf{u} et \mathbf{v} on peut déterminer la vraisemblance de l'observation agrégée conditionnée au super-état \mathbf{s} comme : $p(y_k | \mathbf{z}_k = \mathbf{s}) = \mathcal{N}(y_k | \mathbf{u}[\mathbf{s}], \mathbf{v}[\mathbf{s}])$, où \mathbf{s} correspond à la représentation entière de \mathbf{s} .

En outre, les paramètres pour les émissions des états combinés peuvent être calculées à partir des sommes de Kronecker sur les vecteurs \mathbf{u} et \mathbf{v} . En effet, si au lieu de considérer la puissance agrégée comme observation on considère ses différences consécutives $\Delta y_k = y_k - y_{k-1}$, la vraisemblance des ces émissions : $p(\Delta y_k | \mathbf{z}_k = \mathbf{s}, \mathbf{z}_{k-1} = \mathbf{q})$, est calculée par une fonction gaussienne $\mathcal{N}(y_k | \mathbf{U}[\mathbf{s}, \mathbf{q}], \mathbf{V}[\mathbf{s}, \mathbf{q}])$ [73]. Dans celle-ci, les matrices \mathbf{U} et \mathbf{V} comprennent respectivement les moyennes et les variances des transitions en puissance générées par le

changement d'état d'un ensemble de thermostats.

$$\mathbf{U} = \mathbf{u} \oplus -\mathbf{u}^T \quad (3-55)$$

$$\mathbf{V} = \mathbf{v} \oplus \mathbf{v}^T \quad (3-56)$$

Pour illustrer cette opération, prenons l'exemple des deux charges thermostatiques ayant des consommations moyennes pour chaque état thermostatique $\{0, 1\}$ représentées par les vecteurs $\mathbf{u}_1 = [\mu_0^1, \mu_1^1]$ et $\mathbf{u}_2 = [\mu_0^2, \mu_1^2]$ respectivement. Si l'on considère la vraisemblance de la puissance agrégée, on aura un vecteur de moyennes agrégées calculé selon l'addition de Kronecker (3-54) comme $\mathbf{u} = [\mu_0^1 + \mu_0^2, \mu_0^1 + \mu_1^2, \mu_1^1 + \mu_0^2, \mu_1^1 + \mu_1^2]$. Ce qui représente les possibles consommations générées par l'opération simultanée des deux éléments. Par contre, si l'on s'intéresse aux transitions consécutives Δy_k , les changements en puissance observés au moment de changement des états thermostatiques sont prédits par la matrice (3-55). Dans celle-ci, chaque élément représente la variation moyenne dans la courbe de puissance produite pour le changement simultané des super-états thermostatiques :

$$\mathbf{U} = \begin{bmatrix} 0 & \mu_1^2 - \mu_0^2 & \mu_1^1 - \mu_0^1 & \mu_1^2 - \mu_0^2 + \mu_1^1 - \mu_0^1 \\ \mu_0^2 - \mu_1^2 & 0 & \mu_1^1 - \mu_0^1 + \mu_0^2 - \mu_1^2 & \mu_1^1 - \mu_0^1 \\ \mu_0^1 - \mu_1^1 & \mu_0^1 - \mu_1^1 + \mu_1^2 - \mu_0^2 & 0 & \mu_1^2 - \mu_0^2 \\ \mu_0^1 - \mu_1^1 + \mu_0^2 - \mu_1^2 & \mu_0^1 - \mu_1^1 & \mu_0^2 - \mu_1^2 & 0 \end{bmatrix}$$

Cette matrice a une diagonale nulle, car les transitions vers le même super-état ne produisent pas une variation dans la puissance moyenne du super-état.

Par ailleurs, la même stratégie de combinaison avec les opérateurs de Kronecker peut être utilisée pour la construction des distributions initiales des super-états. Ces opérations seront présentées dans la section prochaine qui résume l'algorithme de Viterbi en tant que procédé pour le décodage parallèle de multiples éléments thermostatiques. La description du modèle est faite à partir de la puissance agrégée. Cependant, il est facilement adaptable à l'analyse événementielle des thermostats en changeant la vraisemblance de la puissance totale $p(y_k | \mathbf{z}_k = \mathbf{s})$ par celle de ses variations $p(\Delta y_k | \mathbf{z}_k = \mathbf{s}, \mathbf{z}_{k-1} = \mathbf{q})$.

3.3.4 *Algorithme de Viterbi*

La stratégie d'agrégation des modèles de transitions probabilistes à deux états permettra d'évaluer de façon parallèle les multiples instances thermostatiques afin de faire un décodage optimal des états cachés des machines ciblées. De plus, à la différence de méthodes semi-markoviennes proposées dans la littérature, la stratégie d'analyse utilisée n'introduit pas de la complexité de calcul. L'approche de reconnaissance des super-états cachés est résumée dans l'Algorithme (3.3). Il est constaté que la structure du processus est quasiment la même que celle de l'algorithme de Viterbi classique. Toutefois, la différence principale de cette méthode se trouve dans les mécanismes de calcul de la matrice de la transition. En effet, le modèle considère de façon explicite les durées des états des processus subjacentes afin d'estimer les transitions entre super-états. En plus, la méthode profite du caractère compact et simple des produits de Kronecker dans le calcul des probabilités entre les super-états.

Algorithm 3.3: Algorithme de Viterbi parallèle pour multiple machines probabiliste à deux états avec émission agrégée.

Input: Séquence d'observations agrégées : y_1, y_2, \dots, y_K
 Paramètres des modèles : $\{\{\phi_m\}, \{\theta_m\}, \{[\tau_0^m, \tau_1^m]\}\}_{m=1..M}$
Output: Séquence de super-états cachés : z_1, z_2, \dots, z_K

```

1 begin
2    $S = \{0, 1, 2, \dots, 2^M\}$ 
3   Définir  $\psi$  et  $Q$  de dimension  $K \times |S|$ , et  $\mathbf{d}$  de dimension  $|S| \times M$ 
4   Estimer la distribution initiale des super-états :  $\pi = \bigoplus_{m=1}^M \ln(\pi_m)$ 
5   Combiner les paramètres  $\{\theta_m\}$  avec Kronecker (équation (3-54)) pour avoir
     l'ensemble de paramètres  $\{\theta_s\}$  associés aux super-états  $s \in S$ 
6   for  $s \in S$  do ▷ Initialisation
7      $\psi_1^s = \pi_s + \ln(f(y_1 | \theta_s))$ 
8      $\mathbf{d}_s = [\tau_0^m, \tau_1^m] \quad \forall m \in \{1, \dots, M\}$ 
9   for  $k = 2, 3, \dots, K$  do ▷ Procédure récursive
10    for  $s \in S$  do
11      Binariser l'indice  $s$  :  $\mathbf{s} = \text{BCD}(s)$ 
12      Définir une matrice  $A$  de dimension  $M \times 2$ 
13      for  $m \in \{1, 2, \dots, M\}$  do
14        Estimer la probabilité de transition individuelle :  $h(\mathbf{d}_s^m | \phi_{\mathbf{s}[m]}^m)$ 
15        Calculer la ligne  $m$  de  $A$  selon l'équation (3-51)
16      Estimer la ligne de la matrice de transition :  $\mathbf{A}[s, :] = \bigoplus_{m=1}^M \ln(A[m, :])$ 
17      for  $s \in S$  do
18        Estimer le meilleur précédent pour  $s$  :  $q^* = \arg \max_{q \in S} (\mathbf{A}_{s,q} + \psi_{q-1}^q)$ 
19        Mettre à jour la probabilité de la séquence :  $\psi_{k,s} = \ln(f(y_k | \theta_s)) + \mathbf{A}_{s,q^*} + \psi_{k-1}^{q^*}$ 
20        Transférer les durées de  $q^*$  vers  $s$  :  $\mathbf{d}_s = \delta(\text{BCD}(q^*), \text{BCD}(s)) \mathbf{d}_{q^*} + 1$ 
21        Stocker le meilleur précédent de  $s$  :  $Q_k^s = q^*$ 
22     $z_K = \arg \max_{s \in S} (\psi_{K,s})$ 
23    for  $k = K, K-1, K-2, \dots, 1$  do ▷ Meilleur chemin
24       $z_k = Q_k^{z_{k+1}}$ 

```

3.4 Synthèse du chapitre

Dans ce chapitre, nous avons présenté les développements fondamentaux de l'approche de désagrégation de multiples charges thermostatiques. Dans la première section, un modèle probabiliste d'une machine stochastique à deux états est développé. Ce modèle est fondé sur une variante des modèles de Markov cachés appelée modèle de Markov à durée explicite

ou semi-markovienne. En effet, cette variante considère de façon explicite le temps écoulé dans un état afin d'introduire des probabilités de transition mieux adaptées à la dynamique réelle des charges thermostatiques. Cette stratégie est déjà proposée dans le domaine NILM, cependant, les approches existantes priorisent la distribution de la durée dans le calcul de probabilités de transition. Ceci débouche dans un processus discret déterministe par morceaux avec des points de rupture dans lesquels les machines font leurs transitions (renouvellement markovien). Cette modélisation implique une inférence sur les durées possibles à chaque point de rupture, ce qui rend la durée une variable aléatoire à inférer, normalement, avec un grand nombre de possibilités [21]. Nous avons réussi à transférer toute la composante aléatoire vers les états discrets des machines. Ceci à partir de l'utilisation de la fonction de hasard (complément de la probabilité de survie conditionnelle). Dans ce nouveau processus les états ne seront plus déterministes par morceaux, en revanche, on évite l'inférence sur la durée laquelle est rendue déterministe par rapport à l'état. Dans la deuxième section, nous avons présenté une méthode capable de détecter de façon parallèle les états des machines modélisée dans la première section. La stratégie de calcul de probabilité de transition proposée a une conséquence importante favorisant l'analyse de multiples charges à états finis. Cette conséquence est reliée à la facilité dans la parallélisation de processus indépendants (la variable d'état d'une charge) avec une émission commune et non indépendante (la puissance agrégée). En effet, en utilisant le modèle de transition basé sur la fonction de hasard, la synchronisation des probabilités de transition des processus parallèles est plus simple car le hasard discret de l'ensemble de processus est égale au produit des valeurs individuelles. Ceci permet de changer l'état de l'ensemble de charges surveillées quand l'observation est favorable à ce changement. Dans l'approche semi-markovienne conventionnelle, les opérations conjointes ont besoin de l'analyse des probabilités postérieures sur les durées. Encore une fois, trop de possibilités à évaluer.

Chapitre 4 - Évaluation de la méthode

Ce chapitre présente une étude comparative de la performance des méthodes proposées et décrites dans le chapitre antérieur. À cet effet, différents scénarios d'agrégation et d'incertitude sont étudiés afin de démontrer en simulation et en expérimentation l'efficacité des différentes méthodes implémentées. De plus, des comparaisons sur la performance de détection des états thermostatiques ainsi que la reconstruction des profils sont effectuées en utilisant des techniques avancées dérivées de l'état de l'art. À ce propos, les résultats de détection des charges thermostatiques conventionnelles sont d'abord présentés. Cette première phase d'évaluation cherche à déterminer la précision des algorithmes dans la détection des charges stationnaires bien connues dans le domaine. Ensuite, l'efficacité de l'approche dans la manipulation de multiples éléments de chauffage est validée.

4.1 Cadre comparatif et scénarios d'évaluation

Cette validation cherche à démontrer le fonctionnement de l'algorithme dans l'analyse de charges stationnaires à deux états. Pour ce faire, nous avons étudié et programmé deux approches dont leurs performances sont connues dans la littérature.

La première approche utilisée dans cette comparaison correspond aux HMM Factoriels (FHMM). Ces derniers font partie de la famille des modèles génératifs. Ils permettent une description appropriée des signaux agrégés sous des effets aléatoires et incertains. En effet, comme discuté pendant le développement de la proposition, les FHMM sont considérés comme un cas particulier des modèles à durée explicite dont la distribution sur la durée est définie par une loi géométrique. De plus, le modèle factoriel est facilement construit à partir des opérateurs de Kronecker [89]. En conséquence, le développement algorithmique exposé dans le chapitre antérieur est compatible et ainsi suffisant pour la programmation des FHMM.

La deuxième approche employée fait partie de la famille discriminante. Plus précisément, la méthode des Réseaux de Neurones Récurents (RNN) permettant l'analyse des structures séquentielles a été utilisée. Ces méthodes sont à la base des modèles d'apprentissage profond. Étant donné que les réseaux de neurones conventionnels ne sont pas adaptés à la modélisation des systèmes ayant des dépendances temporelles à court et moyen terme. Les RNN peuvent introduire des unités avec capacité de traitement de séquences d'états des automates probabilistes. En effet, un grand intérêt pour les techniques d'apprentissage profond a été suscité récemment dans le domaine NILM. Dans cette perspective, quelques méthodes ont été mises au point afin de reconnaître les états de fonctionnement des charges résidentielles [92], [93].

Pour examiner les aspects de la méthode employée, un ensemble de scénarios a été proposé comme suit :

1. Identification et désagrégation des charges conventionnelles à deux états en mesurant la puissance agrégée nette sans perturbation. Dans ce cas, tous les éléments de la courbe de charge sont connus et modélisés.
2. Identification et désagrégation de charges conventionnelles en présence de charges inconnues. Dans ce scénario seulement deux charges sont connues, trois autres charges font partie du profil agrégé, mais celles-ci sont inconnues pour les algorithmes de reconnaissance.
3. Détection de charges thermostatiques ayant des puissances similaires.
4. Identification des paramètres des charges thermostatiques.

4.2 Apprentissage supervisé des approches NILM étudiées

Le processus d'entraînement d'un système NILM supervisé requiert un ensemble de données d'entrées-sorties afin de déterminer les paramètres des modèles utilisés pour la détection. Dans cette phase d'évaluation préliminaire, des données individuelles provenant de bases de données publiques et aussi des mesures recueillies au laboratoire sont utilisées.

4.2.1 Estimation des états à partir des profils individuels

Les états thermostatique de chaque profil peuvent être estimés selon deux procédures différentes. La première consiste à implémenter un processus d'étiquetage des observations du profil individuel à partir d'un algorithme de classification automatique tel que k-moyens, ou bien, un modèle de mélange gaussien (GMM, pour Gaussian Mixture Model en anglais). Ces types de modèles sont appropriés pour *clustériser* des données de puissance bruitées. Une deuxième technique consiste à définir un seuil de puissance comme il est proposé dans [65]. En effet, avec une méthode de seuillage les états seront étiquetés selon la position des données par rapport au seuil. Normalement, la valeur du seuil est définie à partir des connaissances empiriques sur le profil ou bien en prenant la médiane statistique. D'une façon ou d'une autre, étant donné le seuil en puissance \bar{x}_m , la variable d'état z_k peut s'estimer comme il suit :

$$z_k = \begin{cases} 0, & \text{si } x_k \leq \bar{x}_m \\ 1, & \text{si } x_k > \bar{x}_m \end{cases} \quad (4-1)$$

Où x_k représente la puissance du profil individuel à l'instant k . Pour la plupart de charges thermostatiques, un seuil entre 30 et 50 watts fonctionne correctement.

Après avoir déterminé les états thermostatiques z_k pour chaque profil de puissance, un module d'apprentissage est implémenté pour chacune des méthodes de détection à évaluer. Les méthodes et les outils utilisés pour construire les modèles permettant la désagrégation supervisée seront décrits plus bas.

4.2.2 Entraînement des modèles à durée explicite

Les modèles individuels à durée explicite sont la base de la proposition de cette thèse. Chaque paramètre des modèles de charges thermostatiques est utilisé pour construire un modèle combiné capable de poursuivre parallèlement le comportement de multiples instances thermostatiques. À ce propos, pour chaque état s les paramètres sur les émissions en puissance θ_s ainsi que les paramètres ϕ_s pour les lois sur les durées des états sont déterminés à partir des estimateurs proposés dans la littérature. De ce fait, l'ensemble des données relatives aux émissions des états est défini comme $\bar{\mathbf{x}}_s = \{x_k | z_k = s\}_{k=1..K}$.

D'abord, pour le cas des paramètres sur les émissions, il est assumé que la puissance de chaque état suit une distribution gaussienne avec des paramètres $\theta_s = \{\mu_s, \nu_s\}$. En conséquence, les estimateurs empiriques pour ces paramètres sont définis selon les équations suivantes :

$$\hat{\mu}_s = \frac{1}{|\bar{\mathbf{x}}_s|} \sum_{x \in \bar{\mathbf{x}}_s} x, \quad \hat{\nu}_s = \frac{1}{|\bar{\mathbf{x}}_s|} \sum_{x \in \bar{\mathbf{x}}_s} (x - \hat{\mu}_s)^2 \quad (4-2)$$

Où $|\bar{\mathbf{x}}_s|$ correspond à la cardinalité ou nombre d'éléments de l'ensemble $\bar{\mathbf{x}}_s$.

Ensuite, les paramètres ϕ_s de la loi de transition $h(d'|\phi_s)$ sont estimés à partir des estimateurs Maximum a posteriori (MAP). Sur ce point, nous avons présenté trois fonctions paramétriques dans le tableau 3-1. Parmi ces trois, la loi de Weibull discrète s'adapte bien à la dynamique des profils thermostatiques de charges résidentielles conventionnelles. Cette loi est définie par l'équation suivante :

$$f(d'|\lambda, \gamma, \varphi) = \exp\left(-\left(\frac{d' - \varphi}{\lambda}\right)^\gamma\right) - \exp\left(-\left(\frac{d' - \varphi + 1}{\lambda}\right)^\gamma\right) \quad (4-3)$$

Cette loi est importante dans le processus d'apprentissage, car elle représente le modèle de vraisemblance pour lequel le mécanisme d'estimation de paramètres est implémenté. Par ailleurs, la fonction de hasard associée à la loi de Weibull permettant le calcul de la probabilité de transition selon l'équation (3-20) adopte la forme suivante :

$$h(d'|\lambda, \gamma, \varphi) = 1 - \exp\left(-\left(\frac{d' - \varphi}{\lambda}\right)^\gamma + \left(\frac{d' - \varphi + 1}{\lambda}\right)^\gamma\right) \quad (4-4)$$

Malgré la popularité de la loi de Weibull dans l'analyse des durées, il n'y a pas une solution analytique pour ses paramètres à partir des méthodes Maximum of likelihood (MLE) ou le MAP. Toutefois, des outils de calcul permettant des solutions approximatives ont été développés dans les dernières années. Dans cette perspective, plusieurs techniques sont proposées dans le cadre des méthodes MCMC. L'algorithme de Metropolis-Hastings et l'échantillonnage de Gibbs se trouvent parmi les plus connus. Néanmoins, la description mathématique des méthodes MCMC ne sera pas approfondie dans ce manuscrit. La méthodologie suivie et les outils utilisés pour cette tâche sont décrits de façon générale.

Afin d'estimer les paramètres des distributions sur les durées, le module Python PyMC¹ est utilisé. Dans ce module les algorithmes d'apprentissage sont déjà implémentés et testés. Ce outil présente plusieurs types de modèles statistiques bayésiens et les algorithmes d'estimation du type MAP y compris les méthodes MCMC. Sa flexibilité et son extensibilité le rendent applicable au grand nombre de problèmes d'estimation de variables latentes.

Afin de fournir les données nécessaires au processus d'estimation, les durées des états sont enregistrées selon les changements d'état détectés sur la variable z_k . Pour cela, le constructeur d'ensemble est défini : $\vec{d}_s = \{d_{k-1} | z_k \neq s, z_{k-1} = s\}_{k=1 \dots K}$. Ensuite, l'ensemble de paramètres ϕ_s de la distribution (4-4) est échantillonné et enregistré afin d'être utilisé postérieurement dans l'algorithme de désagrégation (3.3). Dans ce modèle, les distributions initiales $P(\phi_s)$ sur les paramètres sont considérées comme uniformes.

4.2.3 Entraînement des modèles FHMM

La construction et l'apprentissage des modèles markoviens à deux états est une tâche relativement plus simple comparée au modèle semi-markovien. En effet, les composantes de la matrice de transition d'un HMM peuvent être définies à partir des paramètres statistiques selon leurs durées. Plus particulièrement, il suffit d'estimer les durées moyennes de chaque état afin de compléter la matrice de transition suivante :

$$A = \begin{bmatrix} 1 - \frac{1}{\tau_0} & \frac{1}{\tau_0} \\ \frac{1}{\tau_1} & 1 - \frac{1}{\tau_1} \end{bmatrix} \quad (4-5)$$

Les durées moyennes τ_s peuvent être calculées à partir d'un estimateur empirique :

$$\tau_s = \frac{1}{|\vec{d}_s|} \sum_{d \in \vec{d}_s} d \quad (4-6)$$

Où $|\vec{d}_s|$ correspond à la cardinalité ou le nombre d'éléments de l'ensemble \vec{d}_s .

Quant aux paramètres des distributions sur les émissions, ils sont déterminés tels que définis pour les modèles à durée explicite. Dans ce cas, les paramètres définis dans l'équation (4-2) permettront de calculer la vraisemblance gaussienne sur les observations de puissance.

1. <https://pymc-devs.github.io/pymc/>

Notons aussi que l'approche d'agrégation des modèles à partir des opérateurs Kronecker est facilement applicable pour les matrices de transition et pour les distributions gaussiennes agrégées.

4.2.4 Entraînement des modèles RNN

Afin d'effectuer une étude comparative en employant des approches de la famille discriminante, nous avons réalisé des analyses de détection et désagrégation des charges thermostatiques à l'aide d'une technique d'apprentissage profond. En ce qui concerne la programmation dans les phases d'apprentissages et reconnaissances, le module de haut niveau Keras² programmé dans le langage Python³ a été utilisé.

En fait, le module Keras permet de manipuler facilement les structures des réseaux des neurones récurrents Réseau de Neurones Récurrents (RNN). Il facilite aussi la validation des problèmes d'apprentissage supervisé. Le choix de cette approche est lié à la capacité d'analyse des séries temporelles et des systèmes ayant un effet mémoire. Nous nous concentrons ensuite sur le développement de l'application plutôt que la programmation de l'approche. Il faut aussi remarquer que le module Keras intègre le moteur d'apprentissage automatique de bas niveau TensorFlow⁴ [94]. C'est un module développé par Google et mis en *open source* en 2015. TensorFlow permet une instanciation rapide et structurée des modèles d'apprentissage de machine pour les recherches et développement d'applications pratiques.

En ce qui concerne le problème de détection des états cachés des charges thermostatiques, nous avons entraîné le RNN en utilisant le profil agrégé comme une entrée du système. Ainsi, les états individuels des charges sont établis pour les sorties du système. En effet, nous voulons que le RNN soit capable de prédire les états à partir des observations de la puissance agrégée. Le schéma d'entraînement est illustré à la Figure 4-1. Cette stratégie est aussi utilisée pour les approches markoviennes et semi-markoviennes utilisées dans cette thèse.

La structure du RNN utilisée comporte deux couches récurrentes ayant chacune 14 cellules avec retour d'état. Une fonction d'activation du type Relu (Unité de rectification linéaire) a été utilisée dans la couche d'entrée et celles cachées. Par contre, une fonction d'activation

2. <https://keras.io/>

3. <https://www.python.org/>

4. <https://www.tensorflow.org/>

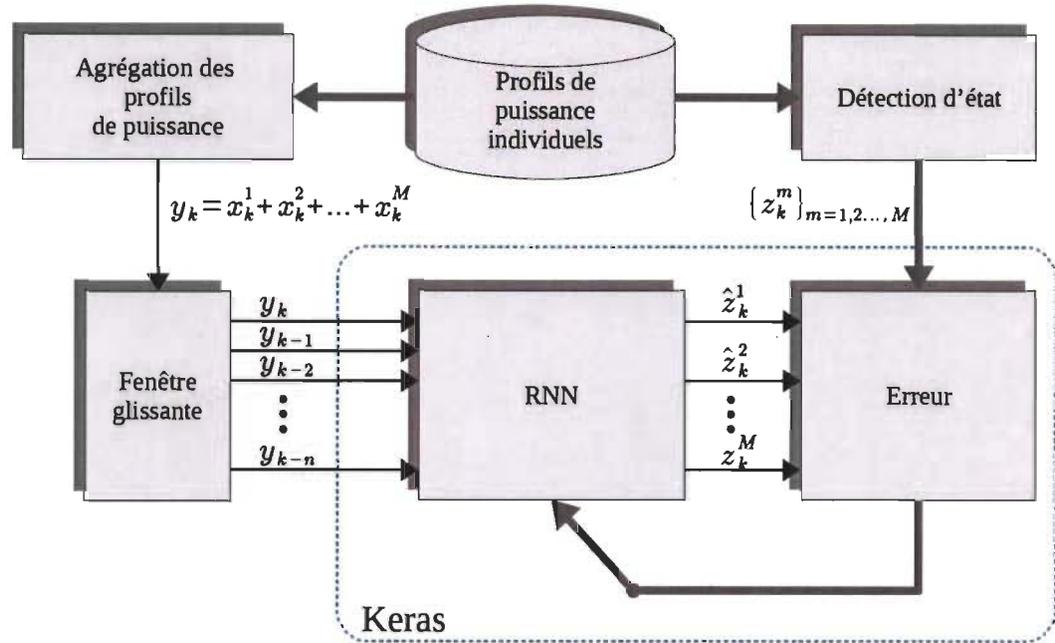


FIGURE 4-1 Schéma du processus d'entraînement du RNN appliqué à la désagrégation supervisée des charges thermostatiques.

sigmoïde est implémentée sur la couche de sortie. Cette fonction permet d'avoir des valeurs entre zéro et un pour les états estimés. Postérieurement, les sorties seront arrondies à l'entier plus proche afin d'avoir ces valeurs binaires. Par ailleurs, l'entraînement du réseau est réalisé en utilisant le module d'optimisation *Adam* implémenté sur TensorFlow. Ce dernier est une variante de l'algorithme du gradient de descente stochastique [95].

4.3 Métriques et indicateurs de performance

Afin d'évaluer et comparer les résultats sur la détection et la désagrégation des charges. La littérature NILM a développé plusieurs métriques et indicateurs de performance. D'une façon générale, deux types de métriques sont utilisées en fonction de la nature de la variable ciblée [96]. Dans le premier cas, nous trouvons les métriques orientées à l'analyse des variables discrètes ou de classification. Dans le deuxième cas, nous repérons les métriques évaluant la capacité de reconstruction des variables continues ou métriques de régression.

4.3.1 Métriques pour la détection de l'état

Dans le cas des charges à états finis, les métriques utilisées pour les variables discrètes permettent d'évaluer la capacité de détection des états de charges représentés par la variable discrète z_k . En outre, la détection d'événements, la classification binaire ainsi que la détection des états d'opération de charges sont effectuées par un indicateur permettant de quantifier la proportion de prédictions produites correctement [21].

$$\text{Accuracy}(\hat{\mathbf{z}}, \mathbf{z}) = \frac{1}{K} \sum_{k=1}^K [\hat{z}_k = z_k] \quad (4-7)$$

Où $[\hat{z}_k = z_k]$ correspond au crochet d'Iverson, celui-ci est une notation retournant 1 si la condition est vraie et 0 sinon.

Bien que la métrique de l'équation (4-7) fournisse une idée du pourcentage d'échantillons qui ont été bien détectés, elle demeure fiable seulement dans l'analyse des machines ayant une distribution d'états uniforme. En d'autres termes, si un des états possède une forte probabilité, un détecteur trivial qui déclare en tout temps cet état comme prédiction aura une performance égale à la probabilité de la classe. Ceci peut mener à une mauvaise interprétation de la qualité du classificateur [97]. Afin d'éviter toute quantité biaisée dans l'analyse, les métriques de performance dérivées des matrices de confusion sont considérées comme plus fiables [26]. En effet, les indicateurs F-score pondérés et le Coefficient de Corrélation de Matthews (MCC) répondent mieux au problème de détection des états avec une distribution stationnaire non uniforme [98]. Dans cette perspective, le MCC est calculé à partir des éléments de la matrice de confusion : TP (nombre de vrais positifs), TN (nombre des vrais négatifs), FP (nombre de faux positifs), FN (nombre de faux négatifs). Le MCC a été récemment introduit dans l'analyse de performances NILM par [98] et [97]. Il est défini selon l'équation suivante :

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4-8)$$

Le MCC renvoie 1 quand la classification est parfaite, 0 quand le classificateur a le même taux de réussite que d'échec et -1 quand il fait complétement le contraire, c'est à dire, quand tous les échantillons sont classifiés incorrectement.

4.3.2 Métriques pour la reconstruction du profil

Les métriques pour la détection de l'état peuvent être utilisées pour évaluer les résultats des classifications des profils. Elles peuvent être aussi étendues au problème de reconnaissance des profils individuels du type "accepté" ou "rejeté". Cependant, elles n'évaluent pas la performance de la reconstruction du profil de puissance. Cela veut dire qu'un système NILM peut avoir une très bonne performance dans la détection ou la classification des charges, mais sa capacité d'estimer la puissance consommée par la charge peut s'avérer d'être de faible précision. C'est la raison pour laquelle les indicateurs comparant le signal réel avec celui estimé sont aussi proposés dans NILM. Une des métriques les plus utilisées dans le domaine correspond à la précision de la reconstruction du profil de puissance (Précision de la Reconstruction en Puissance (PRA)) définie par l'équation suivante [99] :

$$\text{PRA} = 1 - \frac{\sum_{k=1}^K |\hat{x}_k - x_k|}{2 \sum_{k=1}^K x_k} \quad (4-9)$$

Où \hat{x}_k dénote le profil individuel de puissance estimée.

4.4 Résultats préliminaires

Dans cette section les performances de désagrégation de l'approche proposée sont évaluées et comparées avec les méthodes FHMM et RNN. L'objectif de cette première phase d'évaluation consiste à démontrer la capacité de l'approche proposée à reconnaître et reconstruire des profils de charges résidentielles classiques. Pour ce faire, quelques appareils de la base de données *Electricity Consumption and Occupancy (ECO)* [85] ont été sélectionnés et analysés. Nous comparons les performances de désagrégation en soumettant les différents algorithmes aux mêmes hypothèses et informations de départ. D'abord, tous les algorithmes ont l'accès aux profils non agrégés afin de faciliter l'apprentissage supervisé. Nous avons entraîné les modèles en utilisant les profils mesurés pendant quelques journées et nous avons exécuté le processus de désagrégation en utilisant des jours différents à ceux de l'entraînement. Deuxièmement, la puissance agrégée est échantillonnée à la même fréquence et aucun filtrage ou prétraitement n'a été utilisé pendant la phase de désagrégation.

Afin de constater la performance des modèles étudiés, une version simple de la validation croisée a été utilisée. À cet effet, le jeu de données a été divisé en deux sous-ensembles : un ensemble d'entraînement et un ensemble de test. L'ensemble d'entraînement comprend au moins 30% du total des données. Les indicateurs de performance ont été appliqués sur l'ensemble de données de validation. Dans cette étude, les analyses de validation n'ont pas été orientées vers l'optimisation de la structure des modèles RNN ou HMM. En effet, on cherche à valider la capacité de modèles de Markov factoriels afin de prédire les caractéristiques des profils agrégés à partir de l'apprentissage sur des profils individuels relativement courts. Cependant, les profils d'apprentissage contiennent un grand nombre de cycles d'opération des charges. Par exemple, un profil correspondant à une journée de la base ECO comprend environ 20 cycles arrêt-démarrage d'un réfrigérateur. Cette situation garantit une certaine confiance dans la paramétrisation des modèles de Markov factoriels. Par ailleurs, étant donné que les modèles RNN sont entraînés en utilisant des données du profil agrégé, il est nécessaire d'avoir des périodes plus longues afin d'observer toutes les combinaisons des charges possibles. Toutefois, nous avons exposé les trois modèles à la même quantité de données pour l'apprentissage ainsi que pour la validation.

4.4.1 Désagrégation des charges conventionnelles à deux états

Pour cette analyse, nous avons créé un profil agrégé simple à partir de l'addition des données individuelles des cinq charges conventionnelles de la base de données ECO. Les charges en question sont : un réfrigérateur, un congélateur, un HTPC⁵, un lave-vaisselle et une bouilloire électrique. Le Tableau 4-1 montre les paramètres associés à chacune des charges après la modélisation. La colonne du paramètre τ correspond à la période statistique de la charge. La colonne relative à l'énergie indique le pourcentage de l'énergie totale demandée par chaque charge pendant la période de test. Le reste du tableau montre les paramètres de l'état "ON" associés aux modèles des charges utilisées. La colonne π_m correspond à la probabilité stationnaire de l'état "ON" où celle-ci indique la proportion du temps d'utilisation de l'appareil. Pour les cas du lave-vaisselle et de la bouilloire électrique, les probabilités stationnaires sont

5. HTPC : Home theater personal computer

très faibles comparées au reste des machines. Le profil résultant de l'agrégation est tracé à la Figure 4-2.

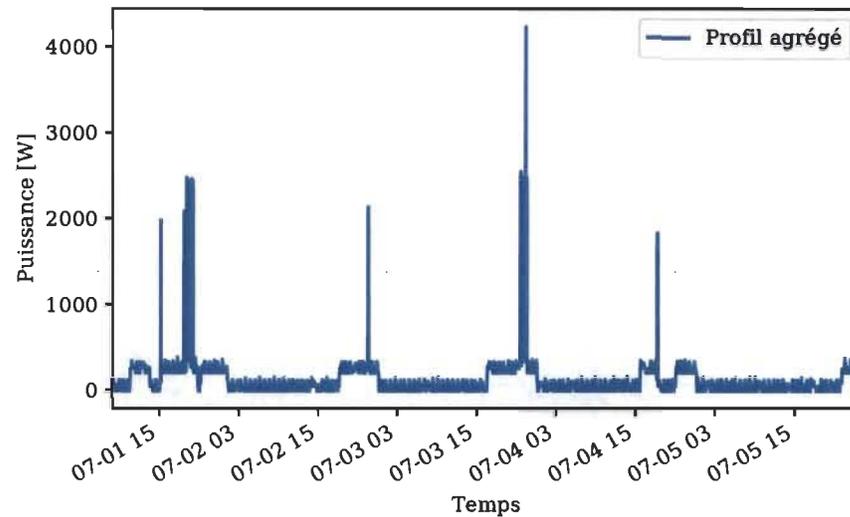


FIGURE 4-2 Profil agrégé des cinq charges thermostatiques.

TABLE 4-1 Paramètres associés à l'état "ON" des charges de la base ECO.

Charge	μ_1 [W]	σ_1 [W]	π_1	τ [min]	Énergie [%]
Réfrigérateur	73	4.9	0.3	46.4	20
Congélateur	52	2.1	0.55	30.44	27
HTPC	146	75	0.32	1013	37
Lave-vaisselle	2002	458	0.01	1500	10
Bouilloire	1776	207	0.01	699	6

Dans cette expérience, la fréquence d'échantillonnage de tous les profils a été établie à une donnée par minute. L'entraînement des méthodes a été réalisé en utilisant une tranche de données comprenant 15 jours. Pour la validation, nous avons mis à profit des données mesurées les 30 jours postérieurs à ceux de l'entraînement. Nous avons exécuté les trois algorithmes en gardant toujours les mêmes conditions d'opération et les mêmes quantités des données d'entraînement. À la fin de chaque processus de détection, la métrique de performance MCC est calculée. Il faut souligner que dans cette phase préliminaire les algorithmes traitent un mélange simple où les composants sont connus. C'est-à-dire, le profil agrégé est composé des cinq charges à deux états dont leurs profils individuels ont été utilisés pour construire le profil agrégé. Bien que ceci soit un scénario peu réaliste due à l'absence des perturbations, il

nous permettra de déterminer la capacité de la méthode proposée et de décomposer des profils thermostatiques simples à faible fréquence d'échantillonnage.

La Figure 4-3 illustre les résultats de désagrégation des trois approches pour les charges thermostatiques évaluées. Bien que la méthode RNN soit capable de détecter les états thermostatiques avec une bonne précision des quatre appareils, certaines difficultés sont ressorties dans la discrimination des charges ayant des puissances similaires. En effet, la distribution de puissance de la bouilloire chevauche celle du lave-vaisselle. Comme conséquence, dans toute l'expérience le RNN n'a pas détecté la bouilloire électrique. De plus, la méthode RNN performe moins bien lorsque la complexité de la courbe agrégée augmente, soit quand plus de charges fonctionnent au même moment. Cette situation est mise en évidence à la Figure 4-4, où le profil reconstruit du congélateur perd sa périodicité quand la puissance agrégée devient plus élevée.

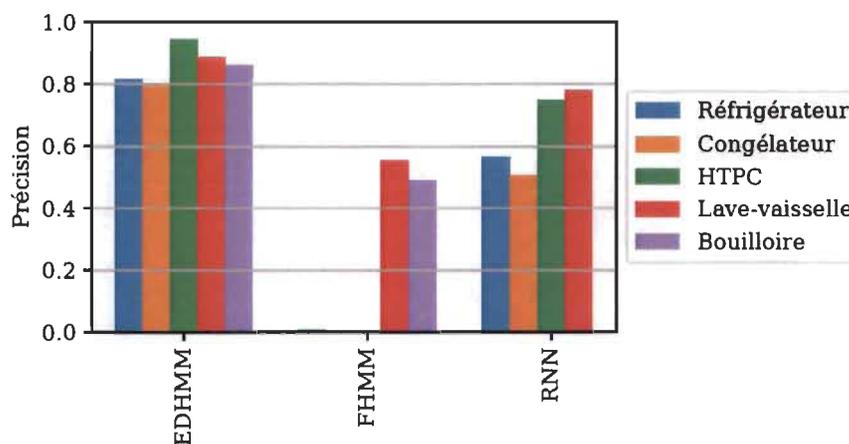


FIGURE 4-3 Indicateur MCC pour la détection des cinq charges thermostatiques de la base de données ECO.

Par ailleurs, étant donné que l'approche FHMM a une très courte mémoire, elle peut permettre des séjours de longue durée dans les états thermostatiques. Conséquemment, cette méthode est incapable de commuter les états des charges de façon appropriée. En outre, la supériorité de l'approche EDHMM repose dans sa capacité explicite de rejeter des transitions dont les distributions des durées ne sont pas respectées. La méthode a été capable de détecter toutes les charges du mélange avec une précision supérieure à 80%. En effet, dans les scénarios où des similarités en puissance peuvent apparaître, la transition d'état est coordonnée avec les

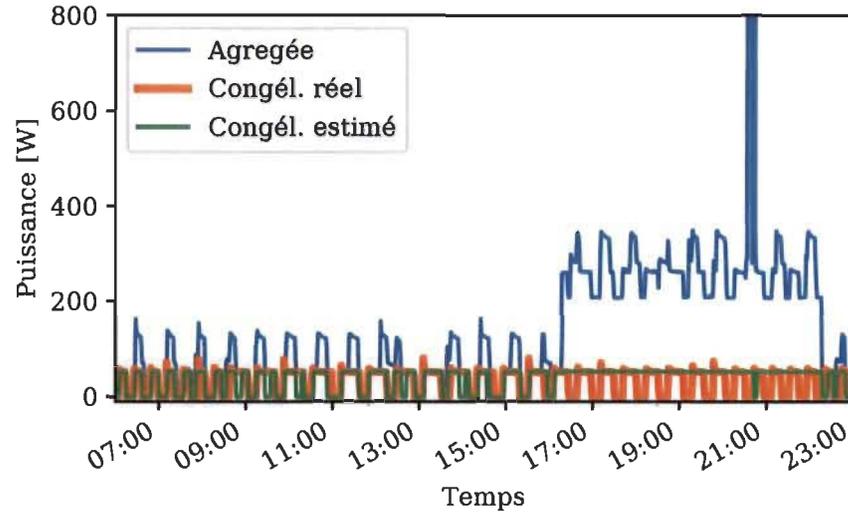


FIGURE 4-4 Puissance du congélateur désagrégée par la méthode RNN.

calculs de probabilité basée sur la durée explicite de l'état de la machine. Donc, les transitions de faible vraisemblance ne sont pas acceptées par le processus de décodage de Viterbi. La Figure 4-5 montre une partie du profil agrégé, ainsi que le profil réel du congélateur et son estimé à partir de la méthode EDHMM. Différemment du RNN, le modèle à durée explicite fournit une reconstruction plus ou moins fidèle quand la courbe agrégée devient complexe.

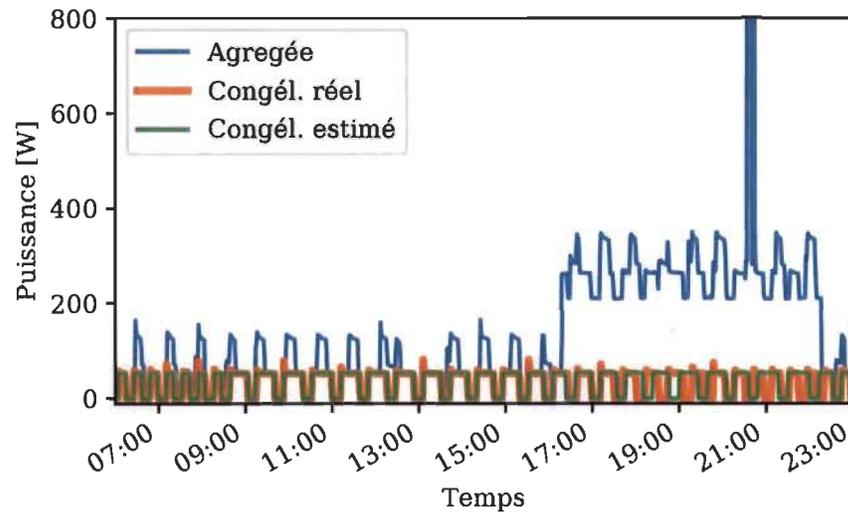


FIGURE 4-5 Puissance du congélateur désagrégée par la méthode EDHMM.

4.4.2 Effet des charges inconnues

Dans cette expérience, la désagrégation de seulement deux profils est ciblée. Ces profils correspondent à la puissance active du réfrigérateur et du congélateur. Le reste des charges listées à la Figure 4-1 font partie de la puissance agrégée mais leur modèles restent inconnus pour tous les algorithmes de désagrégation. À ce propos, nous cherchons à détecter les états du profil du réfrigérateur \bar{x}_{re} et du congélateur \bar{x}_{fr} additionnés dans un profil agrégé de la forme suivante :

$$\bar{y} = \bar{x}_{re} + \bar{x}_{fr} + \bar{x}_{o1} \quad (4-10)$$

Où \bar{x}_{o1} est un composant inconnu ou non considéré dans la phase d'apprentissage.

Les résultats de la détection des états des deux charges cibles sont affichés à la Figure 4-6. On constate que la présence des composants inconnus dans le profil agrégé a un impact considérable sur les trois algorithmes. Néanmoins, l'approche EDHMM garde une performance supérieure à 75%, celle-ci étant significativement meilleure que les performances des méthodes FHMM et RNN dans les mêmes conditions. Étant donné que les méthodes employées se basent sur l'analyse des niveaux de la puissance agrégée, les résultats obtenus peuvent être considérés comme normaux. En théorie, les techniques utilisées sont capables de prédire et gérer les combinaisons d'amplitudes résultantes de l'addition lorsque les composants sont suffisamment connus. Toutefois, la variation de la puissance introduite par un profil non modélisé sera hors de portée de la capacité prédictive des modèles. Cette situation incertaine a stimulé le développement des méthodes orientées vers l'analyse des transitions ou événements plutôt que l'analyse des niveaux.

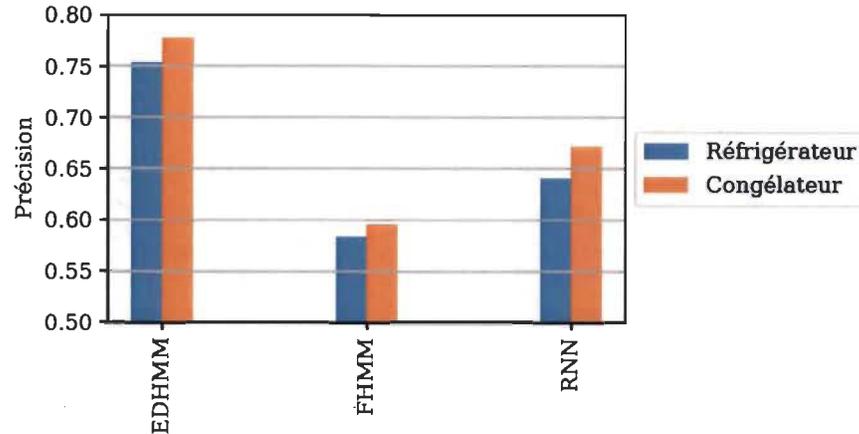


FIGURE 4-6 Indicateur MCC pour la détection des deux charges thermostatiques en présence de charges inconnues.

4.5 Désagrégation de charges thermostatiques actives à des puissances similaires

La dernière analyse conçoit des objectifs principaux de l'approche proposée. En effet, l'incapacité d'une gestion de multiples éléments résistifs ayant une puissance d'opération semblable reste toujours une des principales faiblesses des algorithmes NILM. Sur ce point, l'objectif de cette étude consiste à évaluer l'efficacité des méthodes EDHMM des charges thermostatique dont les puissances individuelles peuvent avoir des distributions égales ou très semblables. De plus, dans l'analyse agrégée de charges thermostatiques, les combinaisons de puissances agrégées peuvent aussi se chevaucher en amplitude. Ceci peut entraîner des difficultés majeures sur la détection de super-états.

C'est dans cette optique que nous avons analysé quatre charges thermostatiques utilisées pour le chauffage de l'espace. Ce sont des convecteurs électriques contrôlés par des thermostats électroniques et connectés selon le schéma de la Figure 4-7. Les puissances nominales de ces quatre éléments sont : 400, 400, 800 et 800 Watts. Dans ce groupe de convecteurs, il y a deux paires ayant la même puissance d'opération. De plus, l'addition des puissances des éléments plus petits se chevauche avec la puissance individuelle des éléments plus grands. En outre, l'agrégation entre les éléments de 800 et 400 Watts est aussi entrelacée. En effet, quand une puissance agrégée de 1200 Watts est observée, le système de détection n'assure pas avec certitude les combinaisons d'éléments formées de la puissance. L'unique combinaison dont nous garantissons avec certitude est l'état de tous les éléments correspondant à 2400 Watts.

Pour celle-ci, tous les thermostats devront être dans l'état "ON".

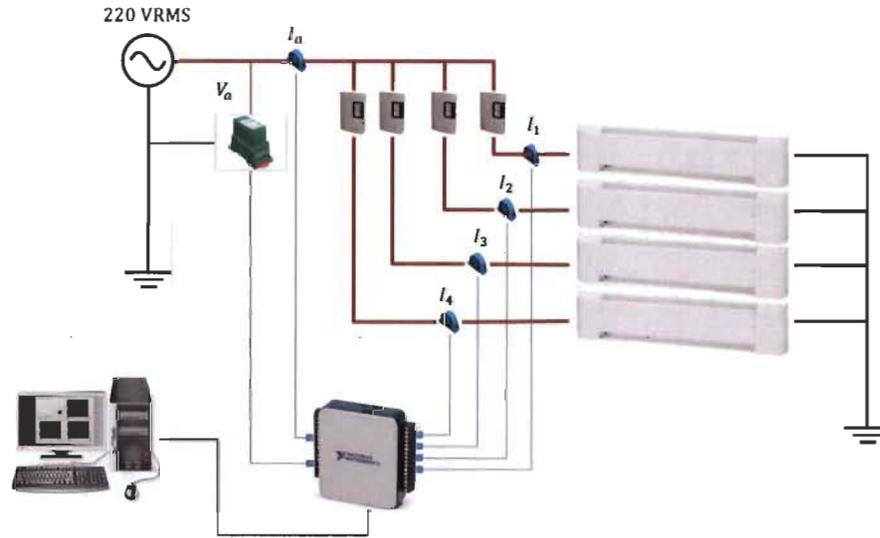


FIGURE 4-7 Système de mesure de la puissance pour des charges thermostatiques résistives.

Étant donné les similarités en puissance, le modèle EDHMM proposé exploite le maximum des différences temporelles ou dynamiques des charges thermostatiques afin de dissocier leurs profils temporels. Précisément, ce sont les distributions sur les durées qui permettent une modélisation dynamique mieux adaptée aux comportements thermostatiques des charges à deux états. Pour les quatre charges étudiées dans cette expérience, nous avons tracé la loi de Weibull pour la durée de chaque état. Celles-ci sont illustrées à la Figure 4-8. Sur le plan pratique, les différences les plus ou les moins remarquables entre ces distributions jouent un rôle important dans l'inférence optimale des états thermostatiques cachés.

En outre, le manque d'une gestion explicite des durées des états pénalise significativement les approches de courte mémoire telles que FHMM et les méthodes orientées vers l'amplitude comme les RNN. Cet avantage important est reflété dans les indices de performance montrés dans le Tableau 4-2 et 4-3.

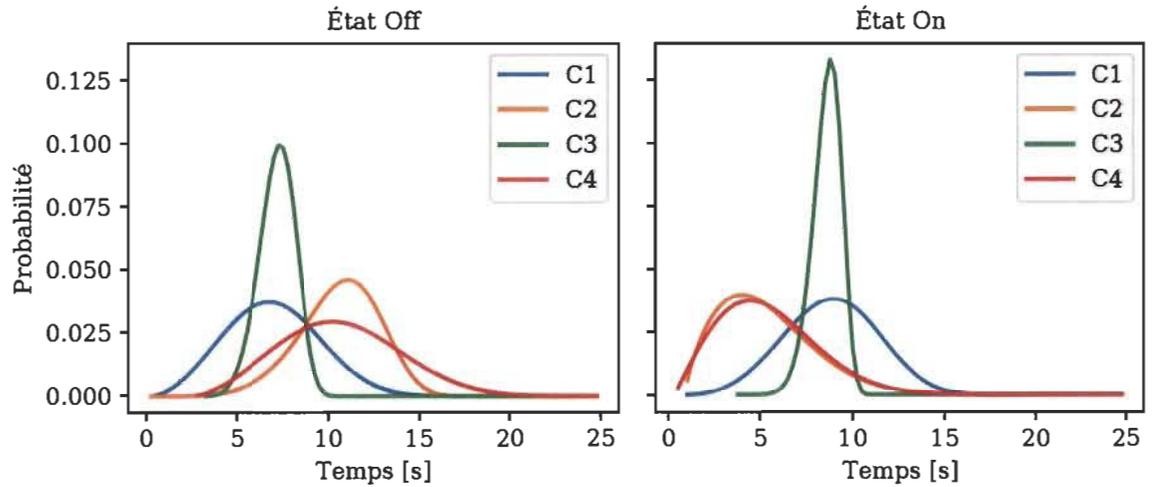


FIGURE 4-8 Distribution de Weibull pour la durée des états thermostatiques des convecteurs électriques.

TABLE 4-2 Indicateur MCC pour la désagrégation charges thermostatiques actives à des puissances similaires.

	EDHMM	FHMM	RNN
Convecteur			
1	0.999	0.979	0.524
2	0.997	0.971	0.549
3	0.980	0.714	0.157
4	0.978	0.679	0.105

TABLE 4-3 Indicateur PRA pour la désagrégation charges thermostatiques actives à des puissances similaires.

	EDHMM	FHMM	RNN
Convecteur			
1	0.996	0.987	0.793
2	0.991	0.975	0.716
3	0.984	0.869	0.637
4	0.973	0.796	0.512
Globale	0.986	0.907	0.665

4.6 Cas des charges de chauffage contrôlées par des thermostats électroniques

Les analyses faites préalablement valident le caractère approprié de la stratégie basée sur la technique EDHMM pour la désagrégation des charges thermostatiques ayant des similarités

en puissance. Ces validations montrent aussi l'avantage en termes de performance de cette méthode sur d'autres méthodes innovantes connues dans la littérature. Effectivement, nous avons constaté que le modèle proposé prédit des résultats qui dépassent significativement les performances des approches FHMM et RNN, et ceci, dans les mêmes conditions d'apprentissage et d'incertitude.

Par ailleurs, une des difficultés de la méthode EDHMM se trouve dans la dépendance de l'apprentissage des modèles de durées. Ces derniers peuvent être facilement biaisés par les données (ceci peut être aussi un problème pour les techniques RNN et FHMM). En effet, les distributions de Weibull illustrées à la Figure 4-8 ont des paramètres que nous estimons à partir des données mesurées. Cependant, les caractéristiques dynamiques de ces charges peuvent varier dans le temps. Pour le cas de charges de chauffage, les distributions des durées sont très influencées par la demande d'énergie de chauffage qui est en même temps très dépendant des facteurs aléatoires tels que la température extérieure et les habitudes des occupants.

4.6.1 *Modèle MLI avec rapport cyclique stochastique*

L'affectation dans le comportement des thermostats électroniques est reflétée directement sur une variation du rapport cyclique thermostatique. Étant donné que ce rapport cyclique reste une variable non observée, son estimation à partir des mesures agrégées devient fondamentale dans le processus de modélisation de l'approche EDHMM. C'est pour cela que nous avons introduit une méthode permettant le calcul des paramètres des distributions des durées des états thermostatiques. La procédure utilisée dans la phase d'apprentissage ne demande pas l'observation des signatures de puissance des charges à modéliser. Elle a besoin seulement des paramètres de haut niveau qui peuvent être fournis par les fabricants des thermostats et des convecteurs. Ils pourront, dans une certaine mesure, être estimés en utilisant des données agrégées. Cependant, dans cette analyse, nous les considérons en tant qu'informations connues.

Les thermostats électroniques effectuent la régulation de la température en utilisant un signal de sortie à la modulation de largeur d'impulsion (PWM) tel qu'illustré à la Figure 4-9. Afin de modéliser le cycle ON-OFF, nous introduisons les variables et les paramètres associés au fonctionnement, qui pourraient être détectables dans le profil de la puissance électrique. Ces informations sont résumées dans la liste suivante :

- τ : Période de commutation.
- τ_0 : Durée de l'état "ON" pendant un cycle thermostatique.
- τ_1 : Durée de l'état "OFF" du cycle thermostatique.
- DC : Rapport cyclique, ceci définit la proportion du temps ON par rapport à la période du cycle τ .

En considérant le rapport cyclique DC un nombre réel défini entre 0 et 1, les relations entre les différents paramètres de temps sont exprimées par les équations :

$$\tau_1 = \tau DC \quad \tau_0 = \tau(1 - DC) \quad (4-11)$$

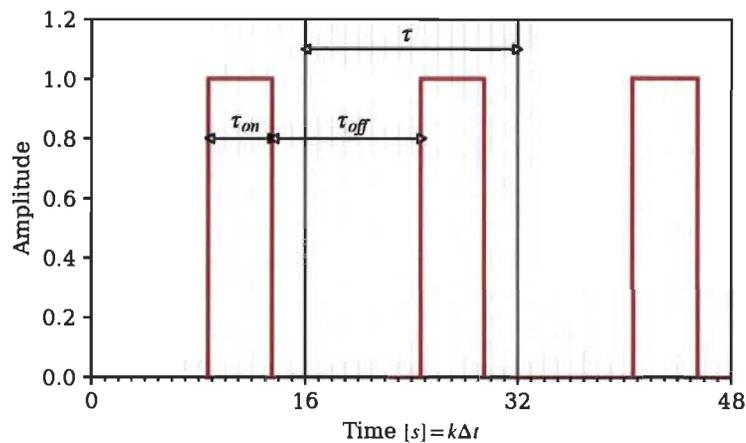


FIGURE 4-9 Profil d'état ON-OFF d'un thermostat électronique.

Pour le contrôle thermique des chambres, les fluctuations de la demande d'énergie sont transférées au rapport cyclique de thermostats ce qui fait varier automatiquement les distributions sur les durées des états "ON" et "OFF". En conséquence, les variations des valeurs espérées dans les durées des cycles modifient significativement les probabilités de transition. Par les approches markoviennes, ces dernières sont supposées être constantes. Afin de donner une continuité dans les approches proposées, nous présentons une procédure pratique permettant l'estimation des modèles de transitions dans un cadre semi-supervisé. C'est à dire, le système est entraîné à partir des données agrégées avec l'aide de quelques informations sur les éléments thermostatiques. Plus particulièrement, cette procédure permet d'estimer les paramètres des distributions de Weibull nécessaires dans le calcul de probabilités de transition.

4.6.2 Initialisation et mise à jour des paramètres

L'initialisation des paramètres pour les distributions sur les durées est faite à partir des connaissances générales sur le fonctionnement des thermostats électroniques. Dans un premier temps, nous considérons une valeur arbitraire pour le rapport cyclique initial. De cette valeur, nous pouvons calculer ensuite les paramètres initiaux λ , φ , γ de la loi de Weibull. Cette distribution doit avoir une moyenne égale au rapport cyclique multiplié par la période de commutation du thermostat. Cette condition permet d'introduire l'égalité suivante :

$$\tau_{DC} = \lambda \Gamma \left(1 + \frac{1}{\gamma} \right) \quad (4-12)$$

Où $\Gamma()$ correspond à la fonction gamma. Ceci est basé sur l'équation définissant la moyenne d'une loi de Weibull à partir de ses paramètres.

Pour le cas d'un thermostat, une hypothèse raisonnable consiste à établir la valeur initiale du rapport cyclique à 0.5, c'est à dire, l'état "ON" et "OFF" ont la même probabilité. Par ailleurs, nous pouvons fixer la valeur du paramètre γ en cherchant une allure de la distribution de Weibull proche de celle de la distribution normale à l'aide d'un $\gamma = 2$. Finalement, dû à l'absence d'information permettant de définir la valeur du paramètre de translation φ , ce dernier a été initialisé à zéro. Dans ces conditions, la valeur de λ peut se calculer avec l'équation (4-12). Cette configuration initiale permet d'avoir une distribution couvrant toute la période τ . De cette façon, on s'attend à ce que la plupart des durées thermostatiques soient dans cette plage du temps. Ce comportement est illustré avec la distribution tracée à la Figure 4-10. Celle-ci correspond à la loi de Weibull sur la durée normalisée par rapport à la période du thermostat. Cette distribution à une moyenne égale à 0.5 et plus du 97% des durées sont dans l'intervalle $[0, \tau]$.

À l'aide du mécanisme d'estimation utilisé les paramètres initiaux permettent d'obtenir une première désagrégation. Postérieurement, ils seront réestimés en utilisant les profils d'état individualisés et le mécanisme d'échantillonnage MCMC. De plus, ces procédures de désagrégation-estimation sont exécutées de façon itérative jusqu'à convergence. Une description graphique du processus de modélisation et désagrégation de multiples thermostats est présentée à la Figure 4-11. De façon générale, cette stratégie suit la procédure proposée

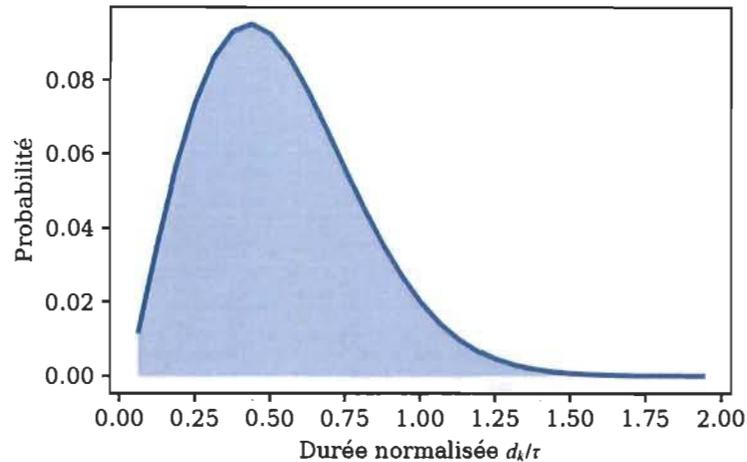


FIGURE 4-10 Distribution de Weibull initiale pour les durées des états thermostatiques.

dans [76]. Celle-ci est basée sur la technique d'inférence des paramètres appelée *entraînement de Viterbi* ou *extraction de Viterbi* (Viterbi extraction, or training). Contrairement à la méthode conventionnelle basée sur la maximisation de l'espérance (Expectation-maximization (EM)), l'entraînement de Viterbi fait une extraction de la meilleure séquence pour des paramètres donnés. Par la suite, les paramètres sont réestimés en utilisant la séquence d'états calculée précédemment. Toutefois, cette procédure n'a pas la précision des méthodes EM et peut fournir des estimations biaisées [100]. En revanche, elle gagne en rapidité et facilite le calcul. Il faut noter que la méthode EM nécessite des calculs prohibitifs pour les modèles semi-markoviens factoriels. Pour le cas spécifique de charges thermostatiques, la fiabilité et la propriété de l'extraction de Viterbi sont évaluées dans les résultats présentés ci-après.

4.6.3 Résultats

Afin de valider la procédure de modélisation et de désagrégation semi-supervisée des profils thermostatiques, nous avons testé l'algorithme avec des données agrégées réelles de quatre convecteurs. La période de la porteuse MLI correspond à $\tau = 16$ secondes. Les puissances de chaque convecteur sont les mêmes du montage de la Figure 4-7. La fréquence d'échantillonnage a été fixée à quatre échantillons par seconde. Tandis que les fenêtres des données de la puissance agrégée sont établies à 5 minutes pour une longueur de séquence de 1200 points. Par ailleurs les variances de la puissance pour les états "ON" sont calculées en

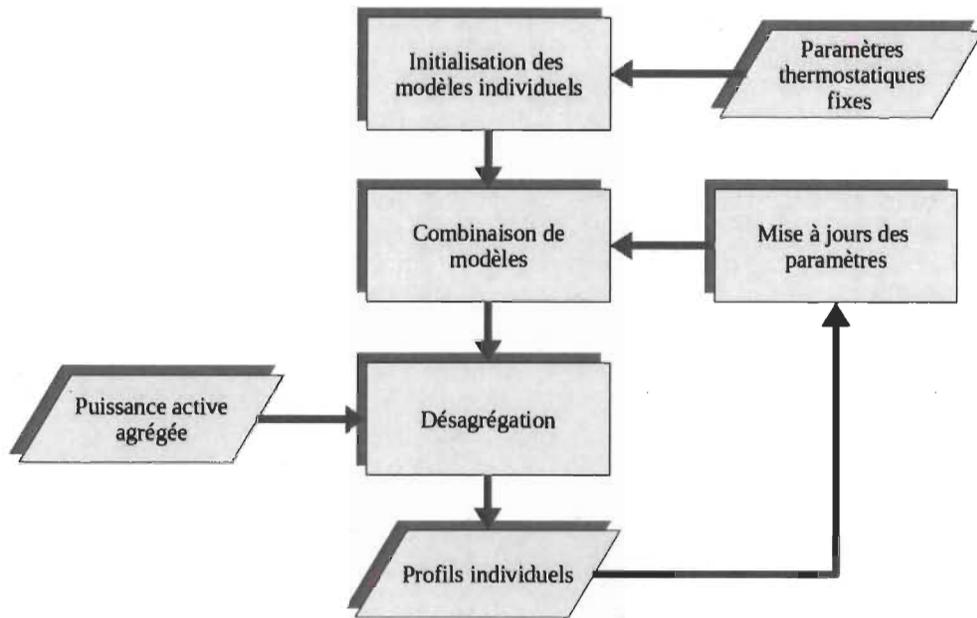


FIGURE 4-11 Structure du système de désagrégation non supervisée des profils thermostatiques.

fonction de la puissance comme $v_1^m = \xi \mu_1^m$ où ξ est une constante.

Les distributions de l'état "ON" et "OFF" d'un des thermostats utilisés dans la phase d'évaluation expérimentale sont tracées à la Figure 4-12. La façon dont le mécanisme de mise à jour des paramètres converge après quelques itérations est ainsi illustré. Il est constaté que la distribution résultante est mieux adaptée à la forme de la distribution empirique des données (courbe avec surface grise). Les distributions empiriques des données sont calculées à partir des données non agrégées. Les autres courbes tracées dans cette expérience correspondent aux distributions intermédiaires calculées pour chaque itération.

Par ailleurs, la Figure 4-13 montre le profil agrégé des quatre thermostats reconstruits à partir de la somme des profils individuels estimés à partir de la méthode de désagrégation EDHMM. Les quatre profils individuels reconstruits avec les propres profils réels sont tracés à la Figure 4-14. Ces résultats démontrent la capacité prédictive de la méthode dans le traitement des profils thermostatiques ayant des signatures identiques. Il est à noter que cette méthode est semi-supervisée, car elle part initialement avec la connaissance du paramètre τ , du nombre de profils et des puissances nominales des éléments chauffants connectés à chacun des thermostats.

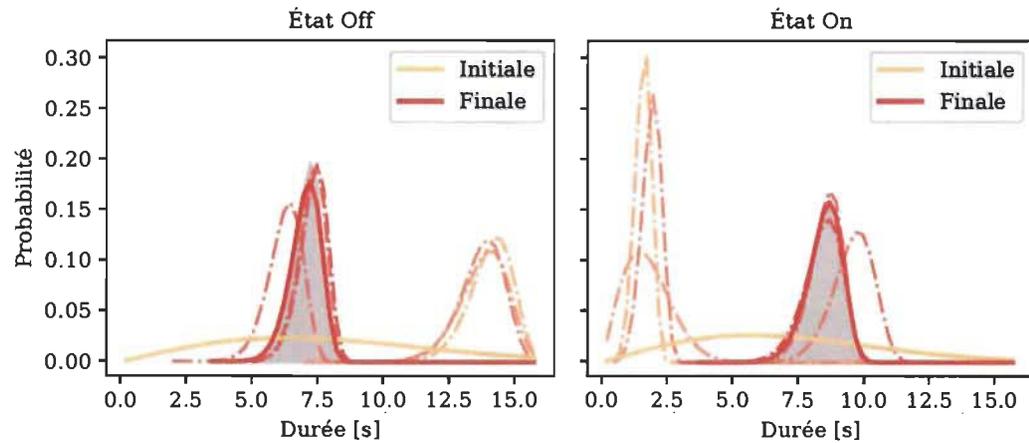


FIGURE 4-12 Estimation en ligne de la distribution sur la durée de l'état thermostatique .

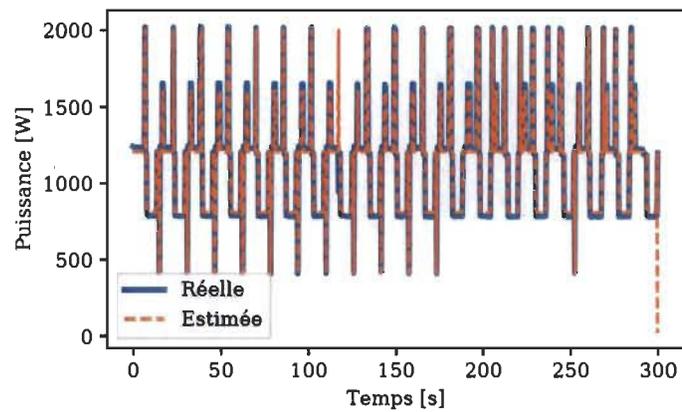


FIGURE 4-13 Profil reconstruit des quatre profils thermostatiques de la Figure 4-14.

Dans cette section, nous n'avons pas réalisé une étude comparative avec la méthode RNN à cause des difficultés associées à l'utilisation de ces techniques dans un cadre non supervisé (sans phase d'apprentissage). De plus, il n'y a pas une méthode appropriée pour introduire les informations a priori sur les éléments à désagréger dans la structure du réseau.

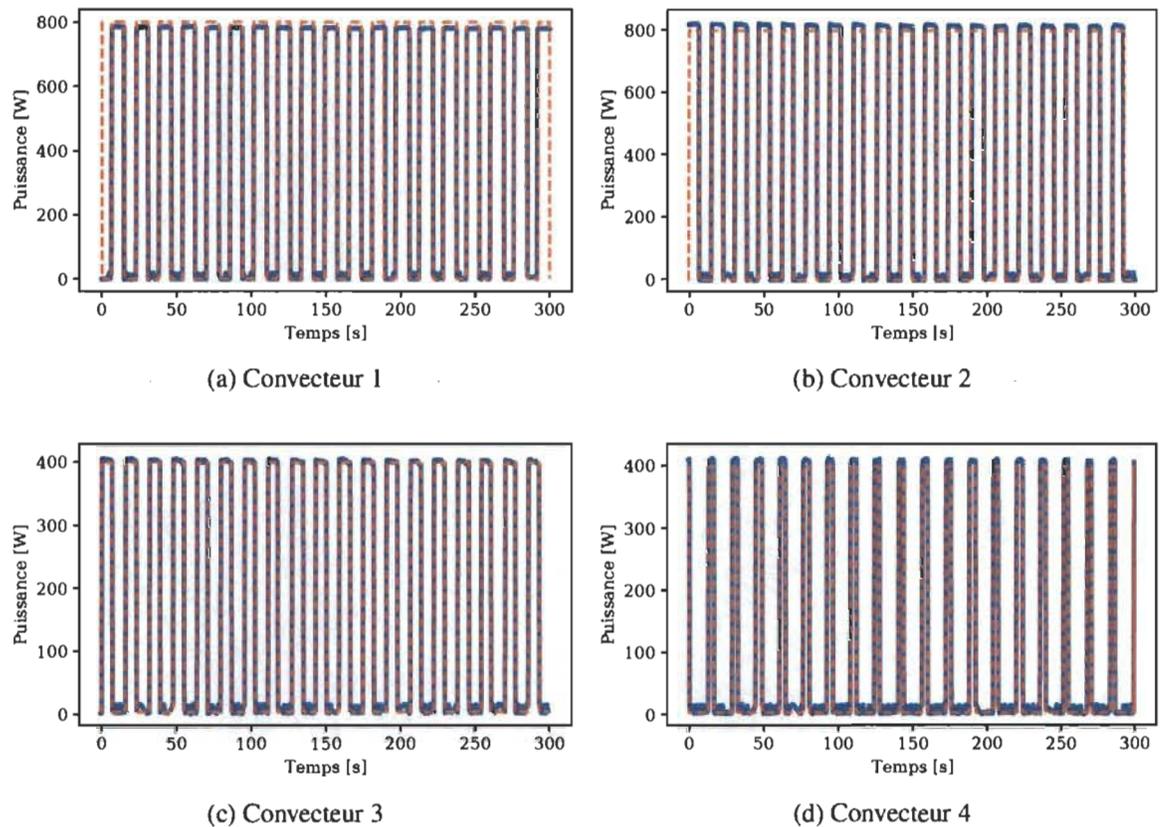


FIGURE 4-14 Profils individuels reconstruits de quatre convecteurs électriques contrôlés par de thermostats électroniques.

4.7 Discussion sur les résultats de la validation

Ce chapitre a été consacré à l'évaluation à partir de données expérimentales de la méthode EDHMM proposée. Nous avons constaté l'efficacité de l'algorithme dans l'analyse de charges actives thermostatiques. Les résultats obtenus démontrent l'applicabilité de cette technique dans la désagrégation de profils ayant des similitudes étroites. De plus, la technique utilisée a été soumise une comparaison avec d'autres méthodes avancées de l'état de l'art. Sur ce point, les résultats de comparaison sont favorables à la technique proposée.

Les tests de validation ont été orientés pour faire ressortir la capacité de l'approche à durée explicite à fin d'exploiter les caractéristiques temporelles des profils thermostatiques. De plus, contrairement aux méthodes EDHMM déjà publiées dans la littérature NILM, l'utilisation des fonctions de hasard discret calculant les probabilités de transition en fonction de la durée de l'état permet une manipulation plus simple de multiples processus simultanées. À cet égard,

les expériences réalisées ont permis de constater le fonctionnement de l'algorithme de Viterbi pour l'interférence d'états cachés. En effet, l'utilisation des méthodes de programmation dynamique exactes dans l'analyse semi-markovienne est peu attirante à cause de la complexité de calcul propre aux procédures de récursion. Toutefois, la méthode utilisée offre une solution optimale en demandant les mêmes ressources de calcul qu'un HMM classique.

Dans cette phase de validation, les trois modèles évalués ont été entraînés en utilisant toujours les mêmes données et conditions d'apprentissage. Bien que la méthode RNN prenne moins de temps pendant la désagrégation, son entraînement est plus exigeant en termes de données et du temps de convergence. En ce qui concerne la fiabilité des méthodes face aux perturbations, les EDHMM gardent une performance toujours supérieure par rapport aux deux autres techniques analysées. Par ailleurs, leur compatibilité avec l'apprentissage semi-supervisé rend les modèles markoviens plus appropriés pour le déploiement dans le secteur résidentiel.

Il faut souligner que les efforts de cette thèse ont été concentrés dans l'évaluation de l'applicabilité de la méthode dans le traitement de charges avec des signatures similaires. En effet, les techniques pour la détection des charges ayant des signatures différentes sont nombreuses dans la littérature NILM. Cependant, les hypothèses fréquemment adoptées sont peu compatibles avec le problème de détection des éléments de chauffage, dont la puissance nominale et la dynamique du contrôle peuvent être semblables. Sur ce point, des essais avec des méthodes conventionnelles ont été réalisés pour la désagrégation de convecteurs. Par exemple, une première étude publiée dans le cadre de cette thèse en 2014 [41] démontre la facilité de désagrégation de charges classiques avec une méthode simple d'analyse événementielle. Cependant, cette technique échoue en présence de charges thermostatiques de chauffage tel que spécifié dans un autre document publié en 2018 [101]. Justement, dans cette publication, nous avons abordé la désagrégation de multiples charges thermostatiques avec une variante de la méthode FHMM utilisée dans le cadre comparatif de cette thèse. Nonobstant, bien que cette technique permette de désagréger le profil de puissance d'un ensemble d'éléments, elle montre des résultats peu satisfaisants dans la désagrégation plus détaillée, soit dans la séparation complète des profils thermostatiques individuels.

En outre, l'évaluation réalisée dévoile l'importance et l'efficacité de l'analyse macroscopique semi-markovienne. Celle-ci se montre plus immune aux perturbations. De plus, elle offre une continuité dans les développements et améliorations futures à cause de sa capacité à intégrer des lois sur les durées mieux adaptées selon la dynamique macroscopique de la charge, contrairement aux méthodes HMM conventionnelles où le processus d'inférence des états cachés perd le contrôle sur la distribution des durées des états.

Chapitre 5 - Conclusions et recommandations

Nous avons présenté une approche pour la désagrégation de charges de type thermostatique capable de reconstruire des profils ayant plusieurs états. Les méthodes utilisées sont majoritairement basées sur des méthodes de l'analyse statistique bayésienne. L'approche de désagrégation opère sur la puissance active échantillonnée à faible fréquence et utilise très peu de paramètres de configuration. De façon générale, le développement est adressé à l'identification et à la reconstruction de profils de puissance des charges thermostatiques. Les efforts ont été concentrés principalement sur la conception de techniques pour le traitement des caractéristiques macroscopiques décrivant le comportement de la puissance active (à un horizon du cycle thermostatique). L'analyse aboutit à une proposition des lois adaptées pour la modélisation non déterministe des durées des états et leur intégration postérieure dans les modèles de transition du type markovien et à durée explicite.

Les travaux de développement théorique ainsi que ceux pour l'évaluation et la validation des approches proposées sont énoncés ensuite.

5.1 Synthèse des travaux

- Les analyses réalisées dans le cadre de ce projet de thèse ont permis le développement d'un modèle de détection des charges thermostatiques. Ces charges catégorisées comme machines à états finis peuvent avoir une dynamique de transition du type markovien ou à durée explicite. À cet égard, les techniques d'analyse conçues permettent l'utilisation de lois adaptées afin de déterminer les probabilités de transition entre les états. En cherchant le contrôle sur la durée de l'état, une analyse probabiliste du comportement stochastique des durées a permis d'établir un modèle de transition basé sur la fonction de risque discret. Étant donné la nature binaire de la représentation d'états, ce modèle

calcule les probabilités de transition à partir d'une loi de Bernoulli avec paramètre dynamique. Il est établi dans la littérature que ce paramètre (considéré maintenant comme une variable) correspond au risque instantané discret. En effet, les modèles markoviens ont une fonction de risque constante, au contraire des modèles semi-markoviens dont le risque instantané varie en fonction de la durée de l'état. De plus, le risque instantané est une expression mathématique facilement dérivable à partir des fonctions de densité et de survie.

- Une fois les modèles de transition individuels établis, une combinaison est faite afin de caractériser statistiquement l'agrégation de la puissance d'un nombre fini d'éléments. La structure résultante de cette combinaison permet de suivre de façon parallèle le comportement de l'ensemble d'éléments. Ceci configure un modèle factoriel à durée explicite où les probabilités de transition entre les éléments correspondent à une loi de Bernoulli multidimensionnelle, ce qui est cohérent avec l'analyse à deux états et permet de modéliser un type de machines à états finis composés. Les charges en question correspondent à celles munies d'un nombre fini d'éléments thermostatiques, par exemple, le chauffage à convecteurs, le chauffe-eau à multiples éléments, ainsi que les cuisinières électriques. Une des originalités de cette proposition se trouve dans la capacité de propager les durées des états individuels vers une structure d'analyse factorielle. Pour ce faire, les durées hypothétiques associées à un état de l'ensemble (ou super-état) sont transférées vers le super-état maximisant la probabilité de transition.
- La structure algorithmique proposée permet l'agrégation de multiples éléments à deux états. Elle permet en même temps l'ajout ou le retrait d'appareils en intégrant facilement les informations individuelles. Cette technique d'analyse adoptée est justement capable d'introduire de nouveaux modèles sans affecter les modèles existants et sans avoir besoin d'un nouvel entraînement du système. Cela s'explique par le fait que le modèle probabiliste factoriel prédit les nouvelles observations sur la variable agrégée quand un nouvel élément est introduit. En outre, l'agrégation de multiples instances est basée sur les opérations de Kronecker. Ce qui permet une représentation compacte de l'opération parallèle d'un nombre fini d'automates probabilistes. Les opérations de Kronecker sont déjà utilisées dans le cadre des modèles de Markov factoriels pour la détection

- parallèle de charges. Toutefois, dans cette thèse, ces opérations ont été étendues afin d'intégrer dans une seule approche les analyses markoviennes et semi-markoviennes.
- Le problème de reconnaissance a été formulé dans un cadre d'optimisation basé sur la programmation dynamique. En effet, une version adaptée à la gestion des processus de comptage de l'algorithme de Viterbi a été proposée. Cet algorithme cherche la maximisation du logarithme de la probabilité conjointe de la séquence d'observations et des états cachés. Dans la procédure de décodage, les compteurs du temps sont incrémentés de façon déterministe par rapport au comportement de la variable d'état. Cependant, ils sont réinitialisés en fonction des décisions optimales de transition. Cette stratégie évite une complexité élevée dans l'espace de décision lequel est un problème majeur des méthodes semi-markoviennes conventionnelles.
 - En utilisant un modèle spécifique, la méthode de reconnaissance fournit une estimation de la probabilité de la meilleure séquence. Cette estimation est utilisée dans le processus de classification et étiquetage automatique des profils individualisés. Profitant du fait que la méthode d'optimisation soit basée sur le calcul de la log-vraisemblance de la séquence, un seuil par rapport à cette valeur est fixé afin de déterminer si le profil appartient à une charge candidate. De plus, il est possible de classer les profils en choisissant parmi les modèles celui ayant la log-vraisemblance la plus élevée. Cette stratégie de reconnaissance est entièrement basée sur les modèles génératifs construits pour chacun des appareils. En conséquence très peu de paramètres et aucun entraînement supplémentaire ne sont requis dans la construction du classificateur.
 - Une évaluation des performances des approches proposées ainsi que des comparaisons avec des méthodes de l'état de l'art ont été réalisées. L'objectif principal de ces études consistait à déterminer la capacité des approches proposées dans :
 - La détection et la désagrégation de multiples éléments thermostatiques dans un cadre supervisé. C'est à dire, lorsqu'on construit les modèles à partir de signatures individuelles de chacune des charges d'intérêt. Plusieurs scénarios ont été évalués afin de déterminer l'efficacité de l'algorithme face aux incertitudes. De plus, le problème de la désagrégation de charges ayant des signatures similaires a été traité. Dans tous les cas évalués, la méthode probabiliste utilisée présentait les meilleures

performances. Dans ce cadre comparatif, les techniques de l'état de l'art utilisées correspondent aux Modèles de Markov Factoriels et aux Réseaux de Neurones récurrents dans tous les tests effectués.

- La désagrégation d'éléments thermostatiques dans un cadre semi-supervisé où quelques informations préalables sur le fonctionnement des charges sont transférées aux modèles de probabilité. Spécifiquement, le cas de détection des charges de chauffage contrôlées par de thermostatiques électroniques a été évalué. Les paramètres de haut niveau utilisés comprenaient la période de commutation de la Modulation à Largeur d'Impulsion (MLI) des thermostats et les puissances nominales des charges. À partir de ces paramètres de départ et à l'aide d'une méthode itérative basée sur l'apprentissage de Viterbi, les paramètres des lois de transition ont été mis à jour.
- En utilisant les approches probabilistes proposées, les connaissances *a priori* sur les éléments thermostatiques ont été intégrées. Ceci afin d'initialiser les procédures d'estimation des états thermostatiques. Cette stratégie aboutit à un entraînement semi-supervisé du système de détection de charge thermostatique pour le chauffage. Elle permet l'initialisation des modèles EDHMM à partir des paramètres connus des thermostats et des convecteurs.

5.2 Limitations et difficultés

5.2.1 Complexité de la méthode factorielle

Une des difficultés majeures des méthodes markoviennes factorielles se trouve dans la complexité de calcul de l'algorithme de Viterbi. En effet, pour analyser M profils à deux états ayant une longueur K , la complexité est de l'ordre $\mathcal{O}(K2^{2M})$ [102]. Pour le cas des modèles semi-markoviens, la complexité est considérablement plus élevée. Celle-ci atteint l'ordre $\mathcal{O}(K2^{2M}D_{max}^2)$, où D_{max} est la durée la plus longue considérée dans le modèle MURPHY [78]. Dans l'approche proposée dans cette thèse, nous avons évité la complexité semi-markovienne en ajoutant des compteurs déterministes du temps de séjour. Cependant, l'analyse parallèle garde le même nombre de pas de calcul de l'algorithme de Viterbi pour FHMM. Ceci est un

problème à calcul quasi prohibitif quand le nombre de profils est supérieur à 18 [103]. Pour un nombre considérable de charges, des méthodes approximatives de calcul sont nécessaires [102].

5.2.2 *Maximums locaux dans l'apprentissage de Viterbi*

L'apprentissage des paramètres des modèles markoviens et semi-markoviens est normalement fait à partir des algorithmes EM. Pour le cas des HMM, cet algorithme adopte le nom de Baum-Welch [104]. Pour les modèles semi-markoviens, l'analyse est plus compliquée et plusieurs techniques sont proposées dans la littérature [38], [86]. À cause de la complexité et de la convergence lente des techniques EM, nous avons opté pour l'apprentissage de Viterbi. Cette stratégie permet l'estimation des paramètres à partir de séquences d'états estimées par l'algorithme de Viterbi. Pour certains cas d'initialisation des paramètres, l'algorithme peut bloquer dans des maximums locaux. Ce problème peut entraîner des biais dans l'estimation des durées des états thermostatiques. Nonobstant, le signal agrégé est normalement bien expliqué par les modèles, et les profils reconstruits ont une précision qui dépasse celle des modèles markoviens.

5.2.3 *L'hypothèse d'indépendance entre les éléments*

Sous l'hypothèse de l'indépendance entre les différents profils subjacents, le terme correspondant à la probabilité d'une combinaison d'états $P(\vec{z})$ peut être factorisé. Donc le décodage des états individuels sera possible à l'aide des procédures de programmation dynamique. Bien que dans la réalité il soit difficile d'assurer une indépendance absolue entre les dynamiques des éléments qui composent une charge à multiples éléments, l'hypothèse d'indépendance apporte une simplification considérable dans les calculs des probabilités jointes. Cependant, cette hypothèse doit être évaluée selon le degré d'interaction entre les différents éléments. Dans certains cas, une dépendance forte peut se présenter quand l'état d'un élément est conditionné par l'opération d'un autre.

5.3 Recommandations

5.3.1 *Lois paramétriques alternatives pour la durée de l'état*

Bien que les modèles à durée explicite ajoutent de la complexité et demandent plus de mémoire dans les étapes de calcul, ils donnent un degré de liberté supplémentaire permettant le contrôle sur la durée de l'état. Techniquement, il existe un grand nombre de lois paramétriques pouvant être testées et étudiées dans la modélisation des durées. Les modèles souvent utilisés dans la littérature NILM sont basés sur la Loi binomiale négative [38] et la loi Gamma [37]. Dans le chapitre 3, nous avons cité et étudié les cas discrets des lois de Weibull et Log-Logistique. Cependant, d'autres lois discrètes, continues et même non paramétriques pourront être testées afin d'améliorer les modèles de transition. Il faut souligner que dans l'analyse de transition, on s'intéresse au risque instantané dérivé de la fonction de survie. Il est donc utile d'avoir des lois ayant des fonctions de survie analytiquement calculables.

Afin d'évaluer la capacité des lois paramétriques à décrire de façon adéquate le comportement des durées, des analyses comparatives entre les distributions empiriques (non paramétrique) et celles paramétriques adoptées doivent être réalisées. Une méthodologie appropriée consiste à calculer la divergence de Kullback-Leibler [73] (ou entropie relative) entre la distribution réelle ou empirique et la loi paramétrique étudiée. Avec cette mesure, il est possible de prioriser la loi ayant une divergence plus faible.

5.3.2 *Distributions priori sur les paramètres de la loi de Weibull*

Tel que proposé dans [73], il est possible de construire dans un cadre semi-supervisé, des distributions a priori sur les paramètres des modèles probabilistes utilisés dans la désagrégation. Introduire justement des informations préalables sous la forme d'une distribution est une manière consistante et pratique de donner une configuration initiale aux systèmes NILM non supervisés. Par exemple, une loi générale sur la durée des états thermostatiques d'une charge pourrait servir comme première approximation. Postérieurement, ses paramètres pourront être améliorés à partir des analyses sur le signal agrégé. Selon la méthodologie bayésienne hiérarchique, l'utilisation de lois a priori sur les paramètres dans la procédure d'inférence est une façon appropriée de combiner les connaissances générales sur le mode

de fonctionnement d'un système et les données observées. En effet, dans l'apprentissage bayésien, on cherche à maximiser une loi postérieure $P(\phi|\vec{d}) \propto P(\vec{d}|\phi)P(\phi)$, où \vec{d} sont des données observées, $P(\vec{d}|\phi)$ est la vraisemblance définie par la loi de Weibull avec paramètres $\phi = \{\gamma, \lambda, \varphi\}$, et $P(\phi)$ est une distribution qui a été supposée uniforme dans les expériences réalisées dans le cadre de cette thèse. Toutefois, cette distribution peut avoir une allure plus informative selon le type de charge à détecter c'est-à-dire qu'une étude approfondie sur le comportement général de ces paramètres selon une charge spécifique devrait aboutir à des lois sur ces paramètres. L'utilisation des lois à priori permettra l'implantation d'une meilleure stratégie d'initialisation des procédures d'estimation en ligne. Ceci pourrait aussi diminuer les possibilités des maximums locaux dans ces procédures d'apprentissage.

5.3.3 Estimation non supervisée des paramètres thermostatiques

Dans les approches de désagrégation proposées, quelques informations importantes associées à l'opération des thermostats doivent être fournies. En effet, pour les modèles à durées explicites des thermostats électroniques, la période τ de commutation du MLI doit être connue. Néanmoins, des analyses fréquentielles pourront aider à détecter cette période. Un exemple d'analyse est montré dans la Figure 5-1. Dans celle-ci, une transformation de Fourier rapide (FFT) sur la puissance agrégée a été réalisée pour une fenêtre de 5 minutes. L'analyse dévoile une fréquence fondamentale correspondant à celle de commutation des thermostats ($f_c = \frac{1}{\tau} = \frac{1}{16} \text{s}^{-1}$). Pour des améliorations futures, ces informations pourront être introduites dans des algorithmes d'initialisation de paramètres pour les modèles markoviens et semi-markoviens.

De plus, les puissances nominales des convecteurs pourront être aussi extraites à partir des analyses de transitions récurrentes. En effet, en suivant la méthodologie proposée dans [41], le traitement de la densité des variations de la puissance agrégée permet de détecter la puissance nominale des convecteurs. La densité d'une tranche de la puissance active analysée pendant une journée est illustrée dans la Figure 5-2. Dans celle-ci, les pics correspondent aux variations récurrentes, lesquelles sont normalement associées à la puissance nominale des charges thermostatiques. Toutefois, des analyses supplémentaires seront nécessaires afin de détecter les faux événements formés par les transitions chevauchées.

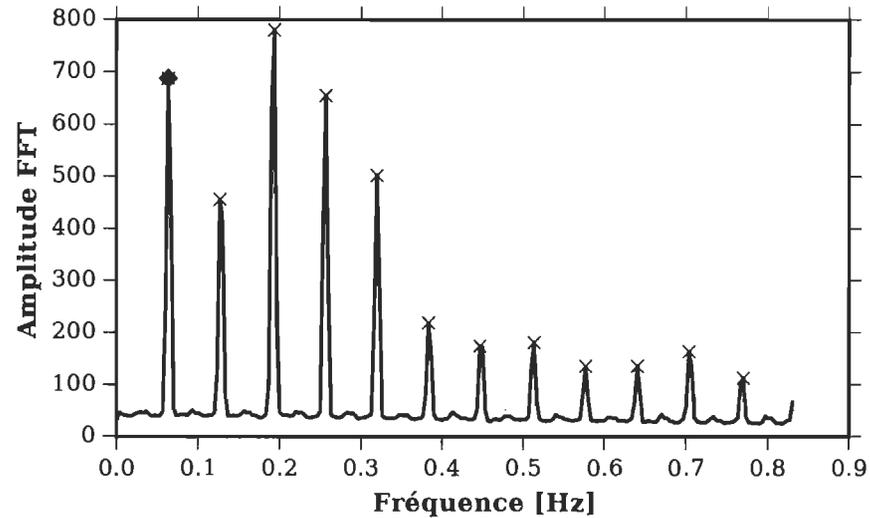


FIGURE 5-1 STFFT de la puissance agrégée contenant multiple charges thermostatiques.

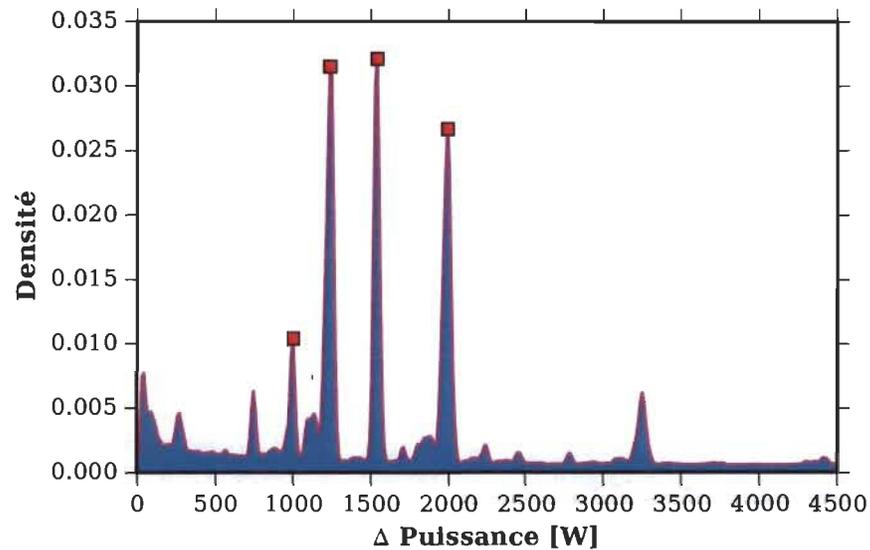


FIGURE 5-2 Analyse de la densité des transitions en puissance de charges thermostatiques.

5.3.4 Profit des ressources NILM disponibles

Dans cette thèse, nous avons travaillé avec le langage de programmation Python qui est un outil très développé pour l'analyse des données et le développement des algorithmes scientifiques. De plus, il y a une grande communauté consacrée au développement scientifique et technique des approches NILM sur ce langage. En effet, il est opportun de citer le module open source Python pour NILM appelé **NILMTK** (Toolkit for Non-intrusive Load Monitoring [96]). Ce module s'est établi comme un outil spécialisé par excellence dans le domaine. Ses

principaux avantages se trouvent dans sa capacité de gérer des bases de données publiques et les possibilités de créer des scénarios de simulation pour l'évaluation des nouvelles techniques.

Les algorithmes NILM implémentés sur NILMTK sont :

- FHMM
- Optimisation combinatoire
- Algorithme de George Hart's [17]
- Réseaux de neurones artificielles (Neuralnilm)
- Latent Bayesian melding (LBM)

De plus, NILMTK offre la possibilité de manipuler plusieurs bases de données publiques.

Parmi les plus connues dans la littérature, on cite :

- REDD¹ : Reference Energy Disaggregation Dataset.
- ECO² : Electricity Consumption and Occupancy data set.
- Dataport³
- PLAID⁴ : Plug Load Appliance Identification Dataset.
- DRED⁵ : Dutch Residential Energy Dataset.

Afin d'avoir une compatibilité avec NILMTK, les nouvelles bases de données doivent suivre le format NILM Metadata. Ceci fait partir d'un effort pour standardiser la structure des bases données orientées à l'analyse de l'énergie résidentielle. Le format définit l'architecture du mesurage et les étiquettes pour les usages et appareils surveillés.

Par ailleurs, un problème important pour l'étape de validation expérimentale et de comparaison des propositions sur le domaine est le manque de données et de scénarios de test standards. Il n'y a pas une base de données complète disponible pour une libre utilisation qui permet de faire une analyse générale des performances d'un système NILM. Les bases disponibles pour évaluer les méthodes NILM varient considérablement entre les projets de recherche qui les ont recueillies et la région géographique où les acquisitions ont été faites. Certains appareils d'intérêts en Amérique du Nord ne sont pas présents dans les bases plus

1. <http://redd.csail.mit.edu/>

2. <https://www.vs.inf.ethz.ch/res/show.html?what=eco-data>

3. <https://dataport.pecanstreet.org>

4. <http://portoalegre.andrew.cmu.edu/>

5. <http://www.st.ewi.tudelft.nl/~akshay/dred/>

reconnues. Dans ce cas il faut chercher des données ou faire des campagnes de mesures pour ces appareils pendant les différentes saisons et les ajouter aux bases existantes.

Bibliographie

- [1] S PACALA et R SOCOLOW. « Stabilization wedges : solving the climate problem for the next 50 years with current technologies. » In : *Science (New York, N.Y.)* 305.5686 (2004), p. 968–972. DOI : 10.1126/science.1100103.
- [2] OFFICE OF ENERGY EFFICIENCY. *Energy Efficiency Trends in Canada 1990 to 2009*. Rapp. tech. M141-1/2009. Ottawa, CA : Natural Resources Canada, Office of Energy Efficiency, 2011, p. 1–60.
- [3] INTERNATIONAL ENERGY AGENCY. *ELECTRICITY INFORMATION, 2012 with 2011 data*. Rapp. tech. OECD/IEA, 2012. Paris, France : OECD/IEA, 2012.
- [4] INTERNATIONAL ENERGY AGENCY et IEA. *25 Energy Efficiency Policy Recommendations*. Rapp. tech. International Energy Agency, 2011, p. 1–12.
- [5] K. CARRIE ARMEL, Abhay GUPTA, Gireesh SHRIMALI et al. « Is disaggregation the holy grail of energy efficiency ? The case of electricity ». In : *Energy Policy* 52 (2013), p. 213–234. DOI : 10.1016/j.enpol.2012.08.062.
- [6] Ahmad FARUQUI, Sanem SERGICI et Ahmed SHARIF. « The impact of informational feedback on energy consumption-A survey of the experimental evidence ». In : *Energy* 35.4 (2010), p. 1598–1608. DOI : 10.1016/j.energy.2009.07.042.
- [7] Mahmoud A. ALAHMAD, Patrick G. WHEELER, Avery SCHWER et al. « A comparative study of three feedback devices for residential real-time energy monitoring ».

- In : *IEEE Transactions on Industrial Electronics* 59.4 (2012), p. 2002–2013. DOI : 10.1109/TIE.2011.2165456.
- [8] Adeel Abbas ZAIDI, Friederich KUPZOG, Tehseen ZIA et al. « Load recognition for automated demand response in microgrids ». In : *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society* (2010), p. 2442–2447. DOI : 10.1109/IECON.2010.5675022.
- [9] Desley VINE, Laurie BUYS et Peter MORRIS. « The Effectiveness of Energy Feedback for Conservation and Peak Demand : A Literature Review ». In : *Open Journal of Energy Efficiency* 02.01 (2013), p. 7–15. DOI : 10.4236/ojee.2013.21002.
- [10] Zheng O'NEILL, Xiufeng PANG, Madhusudana SHASHANKA et al. « Model-based real-time whole building energy performance monitoring and diagnostics ». In : *Journal of Building Performance Simulation* 7.2 (2013), p. 83–99. DOI : 10.1080/19401493.2013.777118.
- [11] Jon FROELICH. « Promoting Energy Efficient Behaviors in the Home through Feedback : The Role of Human-Computer Interaction ». In : *HCIC 2009 Winter Workshop*. DUB Institute, University of Washington. Colorado, 2009.
- [12] A ZOHA, A GLUHAK, M NATI et al. « Low-power appliance monitoring using Factorial Hidden Markov Models ». In : *2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*. IEEE, 2013, p. 527–532. DOI : 10.1109/ISSNIP.2013.6529845.
- [13] OFFICE OF ENERGY EFFICIENCY. *Survey of Household Energy Use*. Rapp. tech. Natural Resources Canada, 2011, p. 1–226.
- [14] Dominik EGARTER, Manfred PÖCHACKER et Wilfried ELMENREICH. « Complexity of Power Draws for Load Disaggregation ». In : (2015), p. 1–10. arXiv : 1501.02954.
- [15] Ahmed ZOHA, Alexander GLUHAK, Muhammad Ali IMRAN et al. « Non-intrusive load monitoring approaches for disaggregated energy sensing : a survey. » In : *Sensors (Basel, Switzerland)* 12.12 (2012), p. 16838–66. DOI : 10.3390/s121216838.

- [16] Stephen MAKONIN, Fred POPOWICH et Bob GILL. « The cognitive power meter : Looking beyond the smart meter ». In : *2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (2013), p. 1–5. DOI : 10.1109/CCECE.2013.6567686.
- [17] George W. HART. « Nonintrusive appliance load monitoring ». In : *Proceedings of the IEEE* 80.12 (1992), p. 1870–1891. DOI : 10.1109/5.192069.
- [18] Nipun BATRA, Oliver PARSON, Mario BERGES et al. « A comparison of non-intrusive load monitoring methods for commercial and residential buildings ». In : *arXiv :1408.6595 [cs.SY]* (2014). arXiv : 1408.6595.
- [19] Taha HASSAN, Fahad JAVED et Naveed ARSHAD. « An Empirical Investigation of V-I Trajectory Based Load Signatures for Non-Intrusive Load Monitoring ». In : *IEEE Transactions on Smart Grid* 5.2 (2014), p. 870–878. DOI : 10.1109/TSG.2013.2271282.
- [20] Yao-Chung FAN, Xingjie LIU, Wang-Chien LEE et al. « Efficient Time Series Disaggregation for Non-intrusive Appliance Load Monitoring ». In : *2012 9th International Conference on Ubiquitous Intelligence and Computing and 9th International Conference on Autonomic and Trusted Computing*. IEEE, 2012, p. 248–255. DOI : 10.1109/UIC-ATC.2012.122.
- [21] Oliver PARSON. *Using Hidden Markov Model Variants for Non-intrusive Appliance Load Monitoring from Smart Meter Data*. Rapp. tech. University of Southampton, 2012.
- [22] Ahmed ZOHA, Alexander GLUHAK, Muhammad Ali IMRAN et al. « Non-intrusive load monitoring approaches for disaggregated energy sensing : a survey. » In : *Sensors (Basel, Switzerland)* 12.12 (2012), p. 16838–66. DOI : 10.3390/s121216838.
- [23] M BARANSKI et J VOSS. « Genetic algorithm for pattern detection in NIALM systems ». In : *IEEE International Conference on Systems, Man and Cybernetics*. T. 4. The Hague, Netherlands, 2004, p. 3462 –3468. DOI : 10.1109/ICSMC.2004.1400878.

- [24] Roman STREUBEL et Bin YANG. « Identification of electrical appliances via analysis of power consumption ». In : *2012 47th International Universities Power Engineering Conference (UPEC) (2012)*, p. 1–6. DOI : 10.1109/UPEC.2012.6398559.
- [25] Jian LIANG, Simon K. K. NG, Gail KENDALL et al. « Load Signature Study—Part I : Basic Concept, Structure, and Methodology ». In : *IEEE Transactions on Power Delivery* 25.2 (2010), p. 551–560. DOI : 10.1109/TPWRD.2009.2033799.
- [26] Marisa FIGUEIREDO. « Contributions to Electrical Energy Disaggregation in a Smart Home ». Thèse de doct. University Institute of Lisbon, 2013.
- [27] Zhenyu WANG et Guilin ZHENG. « Residential Appliances Identification and Monitoring by a Nonintrusive Method ». In : *IEEE Transactions on Smart Grid* 3.1 (2012), p. 80–92. DOI : 10.1109/TSG.2011.2163950.
- [28] Men-Shen TSAI et Yu-Hsiu LIN. « Modern development of an Adaptive Non-Intrusive Appliance Load Monitoring system in electricity energy conservation ». In : *Applied Energy* 96 (2012), p. 55–73. DOI : 10.1016/j.apenergy.2011.11.027.
- [29] M S TSAI et Y H LIN. « Development of a non-intrusive monitoring technique for appliance' identification in electricity energy management ». In : *2011 International Conference on Advanced Power System Automation and Protection (2011)*, p. 108–113. DOI : 10.1109/APAP.2011.6180394.
- [30] Taha HASSAN, Fahad JAVED et Naveed ARSHAD. « An Empirical Investigation of V-I Trajectory Based Load Signatures for Non-Intrusive Load Monitoring ». In : *IEEE Transactions on Smart Grid* (2013), p. 1–9. DOI : 10.1109/TSG.2013.2271282.
- [31] Michael ZEIFMAN, Craig AKERS et Kurt ROTH. « Nonintrusive appliance load monitoring : Review and outlook ». In : *IEEE Transactions on Consumer Electronics* 57.1 (2011), p. 76–84. DOI : 10.1109/TCE.2011.5735484.
- [32] Zhenyu GUO, Z Jane WANG et Ali KASHANI. « Home Appliance Load Modeling From Aggregated Smart Meter Data ». In : *IEEE Transactions on Power Systems* 30.1 (2015), p. 254–262. DOI : 10.1109/TPWRS.2014.2327041.

- [33] Yung Fei WONG, Y. AHMET SEKERCIOGLU, Tom DRUMMOND et al. « Recent approaches to non-intrusive load monitoring techniques in residential settings ». In : *2013 IEEE Computational Intelligence Applications in Smart Grid (CIASG) (2013)*, p. 73–79. DOI : 10.1109/CIASG.2013.6611501.
- [34] Leslie K. NORFORD et Steven B. LEEB. « Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms ». In : *Energy and Buildings* 24.1 (1996), p. 51–64. DOI : 10.1016/0378-7788(95)00958-2.
- [35] S.B. LEEB, S.R. SHAW et J.L. KIRTLEY. « Transient event detection in spectral envelope estimates for nonintrusive load monitoring ». In : *IEEE Transactions on Power Delivery* 10.3 (1995), p. 1200–1210. DOI : 10.1109/61.400897.
- [36] Hsueh-Hsien CHANG. « Non-Intrusive Demand Monitoring and Load Identification for Energy Management Systems Based on Transient Feature Analyses ». In : *Energies* 5.12 (2012), p. 4569–4589. DOI : 10.3390/en5114569.
- [37] Hyungsul KIM, M. MARWAH, M. F. ARLITT et al. « Unsupervised Disaggregation of Low Frequency Power Measurements ». In : *Proceedings of the 11th SIAM International Conference on Data Mining (2011)*, p. 747–758. DOI : 10.1137/1.9781611972818.64.
- [38] Matthew J JOHNSON et Alan S WILLSKY. « Bayesian Nonparametric Hidden Semi-Markov Models ». In : *Journal of Machine Learning Research* 14 (2012), p. 673–701. arXiv : 1203.1365.
- [39] Mahmood AKBAR et Zubair Ahmad KHAN. « Modified nonintrusive appliance load monitoring for nonlinear devices ». In : *INMIC2007 - 11th IEEE International Multi-topic Conference (2007)*. DOI : 10.1109/INMIC.2007.4557691.
- [40] Michael BARANSKI et Jürgen VOSS. « Detecting patterns of appliances from total load data using a dynamic programming approach ». In : *Proceedings - Fourth IEEE International Conference on Data Mining, ICDM 2004*. Ieee, 2004, p. 327–330. DOI : 10.1109/ICDM.2004.10003.

- [41] Nilson HENAO, Kodjo AGBOSSOU, Souso KELOUWANI et al. « Approach in Nonintrusive Type I Load Monitoring Using Subtractive Clustering ». In : *IEEE Transactions on Smart Grid* (2015), p. 1–1. DOI : 10.1109/TSG.2015.2462719.
- [42] Pietro DUCANGE, Francesco MARCELLONI et Michela ANTONELLI. « A Novel Approach Based on Finite-State Machines with Fuzzy Transitions for Nonintrusive Home Appliance Monitoring ». In : *IEEE Transactions on Industrial Informatics* 10.2 (2014), p. 1185–1197. DOI : 10.1109/TII.2014.2304781.
- [43] BU ISLAM. « Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems ». In : *International Journal of Computer Science Issues(IJCSI ...* 8.5 (2011), p. 504–513.
- [44] Zhilin ZHANG, Jae Hyun SON, Ying LI et al. « Training-Free Non-Intrusive Load Monitoring of Electric Vehicle Charging with Low Sampling Rate ». In : *Computing Research Repository* abs/1404.5 (2014), p. 1–7. DOI : 10.1109/IECON.2014.7049328. arXiv : 1404.5020.
- [45] Oliver PARSON, Siddhartha GHOSH, Mark WEAL et al. « Non-intrusive Load Monitoring using Prior Models of General Appliance Types ». In : *26th AAAI Conference on Artificial Intelligence*. Sous la dir. d'ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE (WWW.AAAI.ORG). Toronto, ON, Canada, 2012.
- [46] Oliver PARSON, Siddhartha GHOSH, Mark WEAL et al. « An unsupervised training method for non-intrusive appliance load monitoring ». In : *Artificial Intelligence* 217 (2014), p. 1–19. DOI : 10.1016/j.artint.2014.07.010.
- [47] Marisa FIGUEIREDO, Ana de ALMEIDA et Bernardete RIBEIRO. « Home electrical signal disaggregation for non-intrusive load monitoring (NILM) systems ». In : *Neurocomputing* 96 (2012), p. 66–73. DOI : 10.1016/j.neucom.2011.10.037.
- [48] Sidhant GUPTA, Matthew S REYNOLDS et Shwetak N PATEL. « ElectriSense ». In : *Proceedings of the 12th ACM international conference on Ubiquitous computing - Ubicomp '10*. 2010, p. 139. DOI : 10.1145/1864349.1864375.

- [49] Hsueh-Hsien CHANG, Ching-Lung LIN et Jin-Kwei LEE. « Load identification in non-intrusive load monitoring using steady-state and turn-on transient energy algorithms ». In : *The 2010 14th International Conference on Computer Supported Cooperative Work in Design* (2010), p. 27–32. DOI : 10.1109/CSCWD.2010.5472008.
- [50] Bingsheng WANG, Haili DONG et Chang-tien LU. « Using Step Variant Convolutional Neural Networks for Energy Disaggregation ». In : *NIPS workshop on Machine Learning for Sustainability*. Lake Tahoe, NV, USA, 2013, p. 1–8.
- [51] Aaron R. RABABAHAH et Eniye TEBEKAEMI. « Electric load monitoring of residential buildings using goodness of fit and multi-layer perceptron neural networks ». In : *2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE)*. T. 2. Aaron2012 : IEEE, 2012, p. 733–737. DOI : 10.1109/CSAE.2012.6272871.
- [52] Mario E. BERGES, Ethan GOLDMAN, H. Scott MATTHEWS et al. « Enhancing Electricity Audits in Residential Buildings with Nonintrusive Load Monitoring ». In : *Journal of Industrial Ecology* 14.5 (2010), p. 844–858. DOI : 10.1111/j.1530-9290.2010.00280.x.
- [53] Alan MARCHIORI, Douglas HAKKARINEN, Qi HAN et al. « Circuit-Level Load Monitoring for Household Energy Management ». In : *IEEE Pervasive Computing* 10.1 (2011), p. 40–48. DOI : 10.1109/MPRV.2010.72.
- [54] Krishnan SRINIVASARENGAN, Y.G. GOUTAM, M. Girish CHANDRA et al. « A Framework for Non Intrusive Load Monitoring Using Bayesian Inference ». In : *2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing* (2013), p. 427–432. DOI : 10.1109/IMIS.2013.78.
- [55] Yu Hsiu LIN, Men Shen TSAI et Chin Sheng CHEN. « Applications of fuzzy classification with fuzzy c-means clustering and optimization strategies for load identification in NILM systems ». In : *IEEE International Conference on Fuzzy Systems* (2011), p. 859–866. DOI : 10.1109/FUZZY.2011.6007393.

- [56] Hugo GONÇALVES, Adrian OCNEANU et Mario BERGÉS. « Unsupervised disaggregation of appliances using aggregated consumption data ». In : *Workshop on Data Mining Applications in Sustainability (SIGKDD)*. San Diego, CA, USA, 2011.
- [57] A. COMINOLA, M. GIULIANI, D. PIGA et al. « A Hybrid Signature-based Iterative Disaggregation algorithm for Non-Intrusive Load Monitoring ». In : *Applied Energy* 185 (2017), p. 331–344. DOI : 10.1016/j.apenergy.2016.10.040.
- [58] Yi DU, Liang DU, Bin LU et al. « A review of identification and monitoring methods for electric loads in commercial and residential buildings ». In : *2010 IEEE Energy Conversion Congress and Exposition, ECCE 2010 - Proceedings* (2010), p. 4527–4533. DOI : 10.1109/ECCE.2010.5618423.
- [59] Hala NAJMEDDINE, Khalil EL KHAMLIHI DRISSI, Christophe PASQUIER et al. « State of art on load monitoring methods ». In : *2008 IEEE 2nd International Power and Energy Conference*. IEEE, 2008, p. 1256–1258. DOI : 10.1109/PECON.2008.4762669.
- [60] Jon FROEHLICH, Eric LARSON, Sidhant GUPTA et al. « Disaggregated End-Use Energy Sensing for the Smart Grid ». In : *IEEE Pervasive Computing* 10.1 (2011), p. 28–39. DOI : 10.1109/MPRV.2010.74.
- [61] J Zico KOLTER et Tommi JAAKKOLA. « Approximate Inference in Additive Factorial HMMs with Application to Energy Disaggregation ». In : *JMLR Proceedings* 22 (2012).
- [62] Robert A. JACOBS, Wenxin JIANG et Martin A. TANNER. « Factorial Hidden Markov Models and the Generalized Backfitting Algorithm ». In : *Neural Computation* 14.10 (2002), p. 2415–2437. DOI : 10.1162/08997660260293283.
- [63] Lee Min LEE. « Duration high-order hidden Markov models and training algorithms for speech recognition ». In : *Journal of Information Science and Engineering* 31.3 (2015), p. 799–820.

- [64] Uri HADAR et Hagit MESSER. « High-order Hidden Markov Models - estimation and implementation ». In : *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*. IEEE, 2009, p. 249–252. DOI : 10.1109/SSP.2009.5278591.
- [65] Master THESIS. « Semi-Markov modelling in a Gibbs sampling algorithm for NIALM ». In : (2013).
- [66] Mabrouka E L GUEDRI, Christian LAJAUNIE, Gilles FLEURY et al. *Modèle stochastique pour la séparation des composantes de la courbe de charge électrique*. Paris, 2009.
- [67] Mabrouka El GUEDRI, G D'URSO, Christian LAJAUNIE et al. « Time-Frequency characterisation for electric load monitoring ». In : *Proceedings of the 17th European Signal Processing Conference EUSIPCO*. T. 1. Glasgow, Scotland, 2009, p. 2278–2282.
- [68] Mabrouka El GUEDRI, G D'URSO, Christian LAJAUNIE et al. « RJMCMC point process sampler for single sensor source separation : An application to electric load monitoring ». In : *17th European Signal Processing Conference EUSIPCO*. Glasgow, Scotland : Ieee, 2009, p. 1062–1066.
- [69] Peter J GREEN et B S TW. « Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination ». In : *Biometrika* 82.4 (1995), p. 711–732. DOI : 10.2307/2337340.
- [70] Stephen MAKONIN. *Approaches to Non-Intrusive Load Monitoring (NILM) in the Home*. Rapp. tech. Burnaby : BTEch, British Columbia Institute of Technology, 2012.
- [71] Oliver KRAMER, O WILKEN, P BEENKEN et al. « On Ensemble Classifiers for Nonintrusive Appliance Load Monitoring ». In : *Hybrid Artificial Intelligent Systems*. Lecture Notes in Computer Science 7208 (2012), p. 322–331.
- [72] Oliver PARSON, Mark WEAL et Alex ROGERS. « A Scalable Non-intrusive Load Monitoring System for Fridge-Freezer Energy Efficiency Estimation ». In : *NILM Workshop 2014* (2014), p. 1–4. DOI : 10.3233/NRE-2009-0517.

- [73] O PARSON. « Unsupervised Training Methods for Non-intrusive Appliance Load Monitoring from Smart Meter Data ». Thèse de doct. University of Southampton, 2014.
- [74] Ian GOODFELLOW, Yoshua BENGIO et Aaron COURVILLE. *Deep Learning*. MIT Press, 2016.
- [75] G.D. FORNEY. « The viterbi algorithm ». In : *Proceedings of the IEEE* 61.3 (1973), p. 268–278. DOI : 10.1109/PROC.1973.9030.
- [76] James M. MCFARLAND, Thomas T. G. HAHN et Mayank R. MEHTA. « Explicit-Duration Hidden Markov Model Inference of UP-DOWN States from Continuous Signals ». In : *PLoS ONE* 6.6 (2011). Sous la dir. d’Enrico SCALAS, e21606. DOI : 10.1371/journal.pone.0021606.
- [77] J Zico KOLTER et Matthew J JOHNSON. « REDD : A Public Data Set for Energy Disaggregation Research ». In : *Proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*. 1. 2011, p. 1–6.
- [78] Kevin P MURPHY. *Hidden semi-Markov models (HSMMs)*. Rapp. tech. MIT, 2002.
- [79] Francesco CARTELLA, Jan LEMEIRE, Luca DIMICCOLI et al. « Hidden semi-markov models for predictive maintenance ». In : *Mathematical Problems in Engineering* 2015 (2015). DOI : 10.1155/2015/278120.
- [80] Michael DEWAR, Chris WIGGINS et Frank WOOD. « Inference in hidden Markov models with explicit state duration distributions ». In : *IEEE Signal Processing Letters* 19.4 (2012), p. 235–238. DOI : 10.1109/LSP.2012.2184795. arXiv : 1203.0038.
- [81] Vlad BARBU et Nikolaos LIMNIOS. « Empirical estimation for discrete-time semi-Markov processes with applications in reliability ». In : *Journal of Nonparametric Statistics* 18.7-8 (2006), p. 483–498. DOI : 10.1080/10485250701261913.
- [82] Gerold ALSMEYER. *Renewal, Recurrence and Regeneration*. 2012, p. 1–163.
- [83] Kai KAMMERS, Michel LANG, Jan G HENGSTLER et al. « Survival models with preclustered gene groups as covariates ». In : *BMC Bioinformatics* 12.1 (2011), p. 478. DOI : 10.1186/1471-2105-12-478.

- [84] Mark LUNT. « Survival Analysis ». In : July (2016). DOI : 10.1007/978-1-4757-2728-9. arXiv : arXiv:1504.04543.
- [85] Christian BECKEL, Wilhelm KLEIMINGER, Romano CICHETTI et al. « The ECO data set and the performance of non-intrusive load monitoring algorithms ». In : *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings - BuildSys '14*. 2014, p. 80–89. DOI : 10.1145/2674061.2674064.
- [86] M. AZIMI, P. NASIOPOULOS et R.K. WARD. « Offline and online identification of hidden semi-Markov models ». In : *IEEE Transactions on Signal Processing* 53.8 (2005), p. 2658–2663. DOI : 10.1109/TSP.2005.850344.
- [87] Ritendra DATTA, Jianying HU et Bonnie RAY. « On efficient Viterbi decoding for hidden semi-Markov models ». In : *2008 19th International Conference on Pattern Recognition* (2008), p. 1–4. DOI : 10.1109/ICPR.2008.4761926.
- [88] Stephen MAKONIN, IV BAJIC et Fred POPOWICH. « Efficient Sparse Matrix Processing for Nonintrusive Load Monitoring (NILM) ». In : *2nd International Workshop on Non-Intrusive Load Monitoring 1* (2014).
- [89] Anusha SANKARA. « Energy disaggregation in NIALM using hidden Markov models ». Thèse de doct. Missouri University of Science et Technology, 2015.
- [90] Dayar TUĞRUL. « Analyzing markov chains based on kronecker products ». In : *Computer* 4047.312 (2012), p. 279–300.
- [91] Jozef L. TEUGELS. « Some representations of the multivariate Bernoulli and binomial distributions ». In : *Journal of Multivariate Analysis* 32.2 (1990), p. 256–268. DOI : 10.1016/0047-259X(90)90084-U.
- [92] Wan HE et Ying CHAI. « An Empirical Study on Energy Disaggregation via Deep Learning ». In : 133 (2016), p. 338–342.
- [93] Jihyun KIM, Thi-thu-huong LE et Howon KIM. « Nonintrusive Load Monitoring Based on Advanced Deep Learning and Novel Signature ». In : *Computational Intelligence and Neuroscience 2017* (2017). DOI : 10.1155/2017/4216281.

- [94] Martín ABADI, Ashish AGARWAL, Paul BARHAM et al. « TensorFlow : Large-Scale Machine Learning on Heterogeneous Distributed Systems ». In : (2016). DOI : 10.1038/nm.3331. arXiv : 1603.04467.
- [95] Diederik P. KINGMA et Jimmy BA. « Adam : A Method for Stochastic Optimization ». In : (2014). arXiv : 1412.6980.
- [96] Nipun BATRA, Jack KELLY, Oliver PARSON et al. « NILMTK : An Open Source Toolkit for Non-intrusive Load Monitoring ». In : *ACM e-Energy*. Cambridge, UK, 2014.
- [97] Marc WENNINGER, Jochen SCHMIDT et Toni GOELLER. « Appliance Usage Prediction for the Smart Home with an Application to Energy Demand Side Management - And Why Accuracy is not a Good Performance Metric for this Problem ». In : *Proceedings of the 6th International Conference on Smart Cities and Green ICT Systems*. Smartgreens. SCITEPRESS - Science et Technology Publications, 2017, p. 143–150. DOI : 10.5220/0006264401430150.
- [98] Christian BECKEL. « Scalable and Personalized Energy Efficiency Services with Smart Meter Data ». Thèse de doct. Eberhard Karls University, 2015, p. 1–217. DOI : 10.3929/ethz-a-010578740.
- [99] Stephen MAKONIN et Fred POPOWICH. « Nonintrusive load monitoring (NILM) performance evaluation ». In : *Energy Efficiency* 8.4 (2015), p. 809–814. DOI : 10.1007/s12053-014-9306-2.
- [100] Armen E. ALLAHVERDYAN et Aram GALSTYAN. « Comparative Analysis of Viterbi Training and Maximum Likelihood Estimation for HMMs ». In : (2013). arXiv : 1312.4551.
- [101] Nilson HENAO, Kodjo AGBOSSOU, Souso KELOUWANI et al. « Power Estimation of Multiple Two-State Loads Using A Probabilistic Non-Intrusive Approach ». In : *Energies* 11.1 (2018), p. 88. DOI : 10.3390/en11010088.
- [102] Zoubin GHARAMANI et Michael I. JORDAN. « Factorial Hidden Markov Models ». In : *Machine Learning* 29.2/3 (1997), p. 245–273. DOI : 10.1023/A:1007425814087.

- [103] Stephen MAKONIN, Fred POPOWICH, Ivan V. BAJIC et al. « Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring ». In : *IEEE Transactions on Smart Grid* 7.6 (2016), p. 2575–2585. DOI : 10.1109/TSG.2015.2494592.
- [104] Lloyd R. WELCH. « Hidden Markov Models and the Baum-Welch Algorithm ». In : *IEEE Information Theory Society Newsletter* 53 No.4 (2003), p. 9–13.