



**HAL**  
open science

# Introduction de non linéarités et de non stationnarités dans les modèles de représentation de la demande électrique résidentielle

Arnaud Grandjean

► **To cite this version:**

Arnaud Grandjean. Introduction de non linéarités et de non stationnarités dans les modèles de représentation de la demande électrique résidentielle. Energie électrique. Ecole Nationale Supérieure des Mines de Paris, 2013. Français. NNT : 2013ENMP0002 . pastel-00817969

**HAL Id: pastel-00817969**

**<https://pastel.hal.science/pastel-00817969>**

Submitted on 25 Apr 2013

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

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

École doctorale n°432 : Sciences et Métiers de l'Ingénieur

**Doctorat ParisTech**

**T H È S E**

pour obtenir le grade de docteur délivré par

**l'École nationale supérieure des mines de Paris**

**Spécialité « Énergétique »**

*présentée et soutenue publiquement par*

**Arnaud GRANDJEAN**

le 10 janvier 2013

**Introduction de non linéarités et non stationnarités dans les modèles de représentation de la demande électrique résidentielle**

Directeur de thèse : **Jérôme ADNOT**  
Co-encadrement de la thèse : **Bruno DUPLESSIS**

**Jury**

**M. Bernard MULTON**, Professeur, Laboratoire SATIE, ENS Cachan  
**M. Stéphane Ploix**, Professeur, G-SCOP, Institut Polytechnique de Grenoble  
**M. Guillaume BINET**, Ingénieur de recherche, EDF R&D  
**M. Yvon HARADJI**, Chercheur Senior, EDF R&D  
**M. Jérôme ADNOT**, Professeur, CEP, Mines ParisTech  
**M. Bruno DUPLESSIS**, Docteur, CEP, Mines ParisTech  
**M. Pascal BONDON**, Directeur de recherche CNRS, UMR 8506, Supélec  
**M. Marc PETIT**, Professeur adjoint, Département Énergie, Supélec

Rapporteur  
Rapporteur  
Examinateur  
Examinateur  
Examinateur  
Examinateur  
Examinateur



*« Aussi laisserons-nous le lecteur répondre à cette question : qui est le plus heureux, l'homme qui aura bravé la tempête de la vie et vécu, ou celui qui sera resté en sécurité sur la berge et se sera contenté d'exister ? »*

Hunter S. Thompson



# Remerciements

Une thèse de doctorat c'est tout d'abord trois années passées au sein d'une équipe. Pour ma part, j'ai parfaitement été intégré au sein du groupe E12 de la R&D d'EDF, groupe constitué d'autant de personnalités singulières et atypiques que de personnes qui le constituent. Mes premiers remerciements iront donc à l'ensemble de ce groupe dans lequel j'ai eu la chance d'œuvrer.

Plus particulièrement, je voudrais souligner l'implication de Guillaume Binet, mon tuteur au sein de l'entreprise, qui a su me conseiller, aiguiller mon travail lorsque cela était nécessaire, réorganiser les choses quand je ne savais plus par quel côté du problème l'aborder et m'orienter vers les sources de données les plus pertinentes tout au long de ces trois ans. Ce fut également un partenaire d'entraînement et il s'est avéré très bon en tant que guide lors de nos pérégrinations : merci à toi Guillaume.

Je voudrais également mettre en avant Jean Bieret avec qui j'ai eu la chance de mener, en collaboration avec Guillaume, le travail d'analyse fonctionnelle. Cette étape de la thèse a été probablement celle la plus riche et la plus constructive sur un plan personnel. À cette dimension s'ajoute le fait que nous avons mené le travail dans une ambiance fort sympathique : je me rappellerai longtemps les nombreuses heures passées dans « notre » salle, tous les trois, chacun avec sa tasse et son breuvage. . . Jean qui, depuis cette période, a pris l'habitude de prendre les petits doctorants sous son aile et qui a assuré les conversations aussi diversifiées qu'intéressantes. De celles-ci, j'ai donc pu tirer deux enseignements majeurs : « Pour les systèmes audio, l'important c'est la reconstitution de l'image » et « Pour pouvoir monter/démonter des meubles à l'infini, il faut prendre de la visserie de qualité, de la visserie allemande ! ».

Bien entendu, la rédaction venant, il a su (merci Delphine de l'avoir incité en ce sens) minimiser l'impact parfois trop distrayant de ses interventions. Dans tous les cas, merci à toi Jean qui a contribué au fait que venir aux Renardières tous les matins était plaisant.

Je voudrais également témoigner de mon amitié aux docteurs et doctorants de la R&D que j'ai côtoyés durant ces trois années. À Benoit et Élie avec qui j'ai partagé notre « antre sacrée » placée sous la bénédiction du Totem Courage. À Maxime et Charlotte qui continuent sur la route de la thèse : que le force du Totem soit avec vous ! À Mathieu le tout jeune docteur et à Jean-Michel et Laurent qui ont atteint le grade tant convoité il y a déjà plus longtemps.

Bien évidemment je n'oublie pas mon stagiaire Hugo qui a travaillé à mes côtés, nos colocataires de bâtiment du groupe E11 avec qui j'ai échangé au quotidien, le groupe E14 et plus particulièrement les personnes que j'ai « harcelées » de questions concernant la programmation « dymolienne » ainsi que tou(te)s les collaboratrices/collaborateurs d'Enerbat et des autres départements de la R&D avec qui j'ai pu discuter au cours de ma thèse : merci à toutes et à tous.

Ma seconde « famille d'accueil » a été le Centre Énergétique et Procédés de Mines ParisTech. Bien évidemment, ma gratitude s'adresse en premier lieu à mes deux encadrants Jérôme Adnot, mon directeur de thèse, et Bruno Duplessis mon maître de thèse. Ceux-ci ont toujours répondu présents lorsque je les ai sollicités. Qu'ils soient remerciés ici pour leur disponibilité, la qualité de leurs conseils et de leurs remarques et enfin pour la transmission de leur expérience toujours intéressante. Au CEP, j'ai aussi côtoyé au cours de ces trois années bon nombre de personnes fort sympathiques : Dominique, Pascal, Charles, Thomas, Ozlem, Olivier, David, Marcello, Philippe, Sylvain et Toan. Merci à vous et bonne continuation.

Cette dernière énumération n'est pas complète et pour cause, c'est avec Cyril que j'ai partagé davantage tout au long de ce périple qu'est la thèse. En effet, nous avons tous deux vogué pendant trois ans sur le « frêle esquif » qu'est la modélisation de la courbe de charge résidentielle. Depuis notre première rencontre alors tous jeunes doctorants, c'est de front que nous avons cherché à progresser dans nos travaux, dans les modules de formation et les projets parfois associés et sur le long chemin de l'apprentissage de la science statistique. Nous avons finalement abouti à la mise en place de la « hotline »  $\text{\LaTeX}$  (Cyril, Charlotte et moi). Enfin, pour nous soutenir et nous encourager l'un l'autre, nous n'avons pas fait l'économie de nos cris de ralliement que sont l'emprunté<sup>1</sup> « Dis leur de mettre leurs [...] chaussures de golf! Impossible de marcher dans ce bourbier, on n'aura pas prise. » (pour toutes les fois où nous nous sentions comment dire... enlisés) et l'adapté<sup>2</sup> « In  $\text{\LaTeX}$  we trust » (devise de soutien et de croyance en ce bel outil qu'est celui qui a été utilisé pour rédiger ce que vous lisez, même si parfois...).

J'en viens finalement à remercier mes proches et à leur témoigner de mon affection. À Carine tout d'abord qui partage ma vie et qui a su m'épauler dans les moments de bourre et de doute. Elle a accepté les nombreuses semaines où ne nous voyions que furtivement voire pas du tout principalement du fait de mes activités sportives. Merci également à toi d'avoir assuré le réveil matinal à la fin de la thèse. Merci à mes parents, à mon frère et à ma sœur qui m'ont soutenu dans cette entreprise. Je leur dois beaucoup de ma réussite qu'elle soit scolaire, professionnelle ou sportive et de mon équilibre personnel. Ils ont en effet toujours été présents et à l'écoute en particulier lors des moments difficiles et ils ont mis en œuvre les moyens nécessaires à ce que j'atteigne mes objectifs.

Une autre partie importante de ma vie est faite de sport, plus particulièrement de cyclisme et je voudrais donc témoigner de ma grande amitié envers Julien, mon pilote de tandem avec qui nous avons presque tout connu en trois ans : des performances nationales et internationales prometteuses à notre double embrassade du sol romain en passant par un an de trêve forcée pour cause de blessure. Julien, tu as su être à mon écoute et faire le maximum avec Windy pour que je sois délesté au mieux des tâches annexes mais nécessaires liées à notre pratique du tandem : je vous suis très reconnaissant à tous les deux. J'espère que notre collaboration n'est pas finie et que le temps de notre meilleur niveau ensemble n'est pas révolu. En outre, pendant ces trois ans de pratique cycliste contemporaine à la thèse, nous avons été largement soutenu par Franck et Valérie. Je leur adresse ici mes sincères remerciements.

Enfin, je voudrais conclure ces remerciements par le cercle des Ami(e)s. Merci à toi Tamara d'entretenir le contact depuis le lycée et d'avoir toujours su être attentive et proposer ton aide pour la thèse même lorsque c'était toi qui en avais le plus besoin pour conclure la tienne.

Loïc, tu es toujours resté fidèle depuis que l'on se connaît et la distance qui nous a séparés (souviens toi jusqu'à Braunschweig-Melbourne) n'a pas altéré notre amitié. Merci beaucoup pour cela.

Antoine, toi sans qui mon expérience allemande n'aurait pas été aussi enrichissante sur le plan humain, je t'adresse un grand merci. Tu as réussi à faire disparaître le trop plein d'exigences qui était en moi et qui me coupait des autres. J'espère pouvoir célébrer mon futur titre de « Herr Doktor » avec toi autour d'une excellente fondue zurichoise!

Merci aussi aux autres « ex-Braunschweigiens » avec qui je garde contact en particulier Karin, Christophe, Charlotte, Quentin et les deux Alexandre B. et V..

Pour terminer j'adresse mes remerciements à toutes celles et ceux qui m'ont apporté une aide de près ou de loin, à un moment ou un autre de ce doctorat et dont il est difficile de dresser une liste exhaustive.

---

1. Tirade issue du film *Las Vegas Parano*

2. Inspiré de la devise nationale des États-Unis d'Amérique

# Résumé

La problématique développée dans la thèse est d'estimer, dans une **démarche prospective** et dans un but d'anticipation, les impacts en puissance induits par les **ruptures technologiques et comportementales** qui ne font pas aujourd'hui l'objet de **mesures** dans les panels. L'introduction d'une part de nouveaux usages tels que la mobilité électrique et/ou de nouveaux équipements (prolifération des appareils d'électricité spécifique, appareils thermiques haute performance) et l'avènement d'autre part, de nouveaux souhaits de confort et/ou de scénarios évolués de gestion de l'énergie au sein des bâtiments, constituent des exemples de ruptures d'origine technique/technologique et humaine.

Pour évaluer les modifications sur les appels de puissance du parc résidentiel engendrées par ces profondes transformations, un modèle **paramétrique, bottom-up, techno-explicite** et **agrégatif** est donc nécessaire. Celui-ci serait donc destiné à la reconstitution, de manière non tendancielle, de la **courbe de charge électrique** résidentielle. Il permettrait ainsi de conduire la **simulation** de différents scénarios d'évolution contrastés. L'élaboration d'un tel modèle constitue le sujet de ce doctorat.

Dans un premier temps, des recherches bibliographiques ont été menées pour identifier les modèles déjà existants dans la littérature. Ce travail a permis de mettre en place une classification des méthodes selon trois axes : **champ d'application** du modèle (quels sont les usages modélisés par la méthode ?), **résolution temporelle** (pas de temps de restitution des courbes de charge) et **technique de génération de diversité** (ou foisonnement).

À la suite de ce travail préliminaire, une phase d'**analyse fonctionnelle** a été conduite. Celle-ci a eu pour objectif de concevoir une méthode nouvelle de reconstitution de la courbe de charge résidentielle. Les réflexions que nous avons menées ont permis de **définir des concepts** qui ont été explicités en détails au sein d'un cahier des charges fonctionnel.

La **modélisation d'un certain nombre de concepts** a été ensuite réalisée : en effet, la méthode conceptuelle précédemment évoquée n'est pas directement utilisable pour réaliser des simulations. Ayant identifié l'explicitation du foisonnement des équipements comme un des points essentiels de la thèse, nous avons débuté le travail de modélisation en définissant un **algorithme stochastique** destiné au déclenchement journalier des appareils domestiques. Cet algorithme a pu être testé sur différents cas d'application.

Enfin, en parallèle de la définition de la méthode conceptuelle de reconstitution de la courbe de charge électrique résidentielle et des travaux de modélisation/simulation, nous avons jugé important d'améliorer notre connaissance des courbes de charge par usage. Dans un but d'identification de comportements réels face à l'usage des appareils domestiques, nous avons conduit des travaux **de classification des appels de puissance**. Ceux-ci ont en outre permis de mettre en lumière le besoin d'une **distance adéquate** pour comparer des courbes de charge. De fait, nous montrons dans le manuscrit que la distance euclidienne et les indicateurs classiques de la courbe de charge ne sont pas toujours suffisants pour quantifier la proximité ou la dissemblance entre les courbes d'appels de puissance. Finalement nous proposons une **méthode originale de visualisation et d'analyse du foisonnement** qui affecte les courbes de charge électriques à un niveau agrégé.



# Abstract

In this dissertation we focus on the estimation of the **impacts in terms of power demand** caused by the **technological and behavioural breaks** that will affect the **domestic sector** in the future. These deep changes are not measured in the existing panels and the estimation is required for prospective (long-term) studies.

On the one hand, new end-uses and appliances such as electric vehicles, specific electricity devices or high-performance heating systems are examples of technical/technological breaks. On the other, different comfort wishes and advanced energy management scenarios in the dwellings constitute some possible behavioural evolutions.

To evaluate the very likely modifications of the domestic power demand that will follow previous influences, a **bottom-up, technically-explicit** and **aggregative** model is needed. This one aims at reconstituting the **electric residential load curve** according to a non-trending manner. Thanks to it, various **evolution scenarios** can be **simulated**. The purpose of this PhD is the elaboration of such a model.

In a first step, a **bibliographical study** has been carried out to identify existing models in the literature. This leads to a **three-dimensional classification** according to the range of the model (what appliances can be simulated?), its temporal resolution (what is the time step of the generated load curves?) and the procedure to generate diversity (deterministic, probabilistic, data based...).

After this preliminary step, a **functional analysis** was led. Its role was the elaboration of a new method to reconstitute the domestic power demand. A series of **concepts** have been defined and explained in terms of **specifications**.

Then, we carried out a **modelling step** for some of the defined concepts. Effectively, raw specifications cannot be used in simulations. Since the clarifying of diversity represents one of the key points of our research, we decided to begin the modelling task with focus on it. More precisely, we elaborated a **stochastic algorithm** whose purpose is the realistic starting of domestic appliances. Some application cases have been tested.

Finally we led an experimental study on measured and simulated load curves. This step first aimed at identifying human behaviour concerning the use of appliances. In this frame, **load curve classifications** were carried out. Secondly, we studied the diversity affecting aggregated power demand. We propose a **new methodology to visualise and analyse it**. This method is based on a **distance adapted to the load curve**.



# De la bonne compréhension de ce manuscrit

Pour faciliter la lecture et la compréhension de ce manuscrit, nous avons décidé de mettre en place une nomenclature personnalisée.

Dans l'ensemble du mémoire, la première occurrence des termes importants – qui sont répertoriés dans le glossaire et l'index – est repérée par une astérisque (\*). En outre, la première apparition des acronymes utilisés fréquemment dans le manuscrit est également repérée de cette manière. En plus d'être précisée au moment de leur premier emploi, la signification des acronymes est aussi contenue dans le glossaire. Par ailleurs, un [lien hypertexte](#) (clic sur le mot ou l'expression) renvoie le lecteur à cette entrée du glossaire.

Les notations utilisées sont soulignées lors de leur première apparition dans le texte et sont listées au début du manuscrit.

Les paragraphes en *italique* décalés d'une indentation par rapport au corps du texte constituent les définitions essentielles.

Les paragraphes encadrés et centrés par rapport au reste de la page décrivent les opérations de plus « bas niveau » du modèle de reconstitution de courbe de charge. Celles-ci sont présentées comme tel de manière à faciliter leur identification au sein du manuscrit.

Les hypothèses principales et les conclusions sont doublement encadrées.

Les autres passages importants sont inscrits en **gras**.

Les éléments de l'encodage de la méthode (les variables et leurs modalités) sont repérés en fonte **Typewriter**.

Enfin, les citations que nous avons choisies d'inclure dans ce manuscrit ne se veulent en aucune manière être prétentieuses. Ce sont davantage des clins d'œil à prendre au second degré.



# Liste des notations

Symbole	Descriptif	Unité
<b>A</b>		
$a$	Indice générique des activités ou des tâches domestiques (y compris les macro et les micro-tâches)	-
$AD_{\iota}$	Autorisation de dépassement d'un intervalle $\iota$	-
$\alpha$	Indice générique des instants caractéristiques	-
<b>B</b>		
$\beta$	Indice générique des périodes définissant la temporalité de l'utilisation d'un équipement	-
$bt$	Bâtiment-type	-
<b>C</b>		
$c$	Compatibilité	-
$C$	Consommation d'énergie	kWh
$CDCMF$	Courbe De Charge Moyenne Foisonnée	-
$CDCU$	Courbe De Charge Unitaire	-
$c_{EE,i}(\iota)$	Compatibilité En Équipements de l'équipement $i$ par rapport à l'intervalle infra-journalier $\iota$	-
$C_{elec. sp\acute{e}. tot.}$	Consommation totale des équipements d'électricité spécifique	kWh
$C_{elec. sp\acute{e}. \acute{e}q. exp.}$	Consommation des équipements d'électricité spécifique explicités	kWh
$c_{ET,i}(\iota)$	Compatibilité En Temps de l'équipement $i$ par rapport à l'intervalle infra-journalier $\iota$	-
$c_i(\iota)$	Compatibilité d'un équipement $i$ par rapport à l'intervalle $\iota$	-
$c_{MTP i}(\iota)$	Compatibilité de la micro-tâche parente de l'équipement $i$ par rapport à l'intervalle $\iota$	-
$comp_{\iota}$	Compteur de balayage dans l'intervalle $\iota$	-
$comp_{Usage,i}$	Compteur de balayage dans l'usage de l'équipement $i$	-
$CUP$	Cycle Unitaire de Puissance	-
$inst CUP$	Version instantanée d'un Cycle Unitaire de Puissance	-
<b>D</b>		
$d$	Durée	s
$d_{AC}$	Durée à considérer	s
$d_{\beta}$	Périodes définissant la temporalité de l'utilisation d'un équipement	-
$d_{\beta, Disp}$	Niveau de disponibilité requis au sein des périodes définissant la temporalité de l'utilisation d'un équipement	-
$d_{\beta, Dur\acute{e}e}$	Durée des périodes définissant la temporalité de l'utilisation d'un équipement	s

$d_{PNDU}$	Durée de la partie non déplaçable de l'usage d'un équipement	s
$d_{Simulation}$	Durée de la simulation	s
$\delta$	Distance entre deux objets $o_1$ et $o_2$ issus d'un ensemble $\Omega$	[ $\Omega$ ]
$\Delta t_{Abs. Jour.}$	Durée séparant les instants caractéristiques $t_1$ et $t_2$	s
$\Delta t_{Départ}$	Durée séparant les instants caractéristiques $t_{Start}$ et $t_1$	s
$\Delta t_{P_{max}, CUP}$	Durée pendant laquelle la valeur $P_{max}$ du cycle unitaire est appelée	s
$\Delta t_{Prés. MÉR.}$	Durée séparant les instants caractéristiques $t_3$ et $t_4$	s
$d_{Tampon}$	Durée tampon de l'usage d'un équipement	s
$d_{Usage}$	Durée d'usage d'un équipement	s
<b>E</b>		
$E$	Quantité d'énergie	J
$ec$	Événement clé	-
$élec. spé.$	Électricité spécifique	-
eucl	Indice diminutif du terme euclidienne (pour la distance éponyme)	-
$\varepsilon$	Critère de convergence pour mesurer la proximité entre courbes de charge	[ $\varepsilon$ ]
<b>F</b>		
$f$	Fréquence d'un signal sinusoïdal	Hz
$F$	Facteur de puissance	-
$F_{P_{max}} (F_{n, P_{max}})$	Facteur de charge de la puissance maximale (pour $n$ consommateurs)	-
<b>G</b>		
$G$	Gradient thermique de consommation énergétique	W/ °C
$\gamma$	Distance cumulée dans l'algorithme de calcul du $DTW$	[série temporelle]
<b>I</b>		
$i$	Indice générique des équipements domestiques	-
$i$	Intensité d'un courant électrique sinusoïdal	A
$I$	Courant efficace d'un courant électrique	A
$IAC$	Intervalle À Considérer	-
$IE(\iota, i)$	Éligibilité de l'intervalle infra-journalier $\iota$ pour l'équipement $i$	-
$IL$	Intervalle Limite	-
$inf$	Borne inférieure	-
$\iota$	Indice générique des intervalles infra-journaliers	-
$\iota_{IAC}$	Indice de l'intervalle à considérer	-
$IT$	Intervalle Traité	-
<b>J</b>		
$j$	Indice générique de ménages ou de foyers	-
$jr$	Indice générique d'un jour considéré (simulé ou mesuré)	-
<b>K</b>		
$K$	Taille d'un chemin de déformation $W$	-
$\kappa$	Indice générique d'un élément du chemin de déformation $W$	-
$k_{f, n}$	Diversity factor (inverse du coefficient de foisonnement) pour un groupe de $n$ consommateurs	-
$K_{f, n}$	Coefficient de foisonnement (coincidence factor) d'un groupe de $n$ consommateurs	-

**L**

$l_{Intervalle}$  Longueur d'un intervalle infra-journalier s

**M**

$m$  Nombre entier -

$M$  Matrice de taille  $n \times m$  contenant les distances  $\delta$  estimées point par point entre deux séries temporelles dans l'algorithme *DTW* [ $\delta$ ]

$MAE$  Mean Absolute Error (erreur absolue moyenne entre deux séries temporelles d'une variable donnée) [variable]

$MAPE$  Mean Absolute Percentage Error (erreur quadratique moyenne entre deux séries temporelles d'une variable donnée) [%]

Manh Indice diminutif du terme Manhattan (pour la distance du même nom) -

Mink Indice diminutif du mot Minkowski (pour la distance éponyme) -

$MMAC$  Matrice de MAcro-Compatibilité -

$MMIC$  Matrice de MIcro-Compatibilité -

$mt$  Ménage-type -

**N**

$n$  Nombre entier -

$\mathbb{N}$  Ensemble des entiers naturels -

$N_{agr}$  Niveau d'agrégation d'une Courbe De Charge Moyenne Foisonnée -

$n_{CF}$  Nombre de cas favorables au placement d'un équipement -

$nu$  n-uplet -

$NVF$  Normalized Variation Factor -

**O**

$o$  Objet d'un ensemble  $\Omega$  -

$\omega$  Pulsation propre d'un signal sinusoïdal rad.s<sup>-1</sup>

$\Omega$  Ensemble d'objets -

**P**

$p$  Probabilité -

$P$  Puissance électrique W

$P$  Puissance active W

$\mathbf{P}$  Puissance instantanée W

$\bar{P}$  Puissance moyenne W

$P_{est}$  Puissance estimée (calculée ou simulée) W

$p_{\iota, i}$  Probabilité pour l'intervalle  $\iota$  d'accueillir l'instant de déclenchement de l'usage de l'équipement  $i$  -

$P_{max}$  Puissance maximale W

$P_{max, CUP}$  Puissance maximale d'un Cycle Unitaire de Puissance W

$P_{med}$  Puissance médiane W

$P_{mes}$  Puissance mesurée W

$P_{min}$  Puissance minimale W

$P_{nom}$  Puissance nominale d'un équipement domestique W

$P_r$  Puissance réduite (adimensionnelle) -

$P_{veille}$  Puissance de veille d'un équipement domestique W

$\phi$  Déphasage géométrique entre la tension et le courant électrique rad

	$\psi$	Indice d'un élément d'un vecteur	-
	$\varphi$	Dissimilarité entre deux objets $o_1$ et $o_2$ issus d'un ensemble $\Omega$	$[\Omega]$
<b>Q</b>	$qle$	Quartile d'un échantillon	-
	Q	Puissance réactive	VAR
	Q	Série temporelle	-
<b>R</b>	$\mathbb{R}^+$	Ensemble des nombres réels positifs	-
	$r_{EAC}$	Repère de l'équipement à considérer	s
	$RJT$	Résumé du Jour-Type	-
<b>S</b>	$S$	Série temporelle échantillonnée sur $n$ points	-
	<b>S</b>	Version segmentée de la série $S$	$[S]$
	<b>S</b>	Puissance apparente	VA
	$SFE$	Solde Fatal en Énergie	kWh
	$sup$	Borne supérieure	-
	$s_x$	Signature d'équipement	-
	$SXL$	Abscisse de l'extrémité gauche d'un segment de droite	-
	$SXR$	Abscisse de l'extrémité droite d'un segment de droite	-
	$SYL$	Ordonnée de l'extrémité gauche d'un segment de droite	-
	$SYR$	Ordonnée de l'extrémité droite d'un segment de droite	-
	$SW$	Poids éventuel associé à un segment de droite	-
<b>T</b>	$t$	Temps	s
	$T$	Température	°C
	<b>T</b>	Nombre de segments d'une série temporelle <b>S</b>	-
	$t_1$	Instant caractéristique séparant les intervalles Réveil et Matin	s
	$t_2$	Instant caractéristique séparant les intervalles Après-midi et Soir	s
	$t_3$	Instant caractéristique séparant les intervalles Matin et Midi	s
	$t_4$	Instant caractéristique séparant les intervalles Midi et Après-midi	s
	$t_\alpha$	Instants caractéristiques (appartenant au jeu initial d'instants caractéristiques)	s
	$t_{\alpha+}$	Instants caractéristiques appartenant aux jeux supplémentaires d'instants caractéristiques	s
	$t_{End}$	Instant de fin des activités domestiques pour le jour $n$	s
	$t_{Final}$	Instant de fin d'un intervalle infra-journalier	s
	$t_{Initial}$	Instant de début d'un intervalle infra-journalier	s
	$t_{Start}$	Instant de démarrage des activités domestiques pour le jour $n$	s
	$t_{StartBis}$	Instant de démarrage des activités domestiques pour le jour $n + 1$	s
	$\theta$	Indice d'un élément d'un vecteur	-
<b>U</b>	$U_{Bât}$	Coefficient de transmission surfacique	W/(m <sup>2</sup> .K)
<b>V</b>	$V$	Série temporelle	-
	<b>v</b>	Tension d'un courant électrique sinusoïdal	V

<b>V</b>	Tension efficace d'un courant électrique	V
$V_{nom}$	Tension nominale du réseau électrique	V
<b>W</b>		
$W$	Chemin de déformation dans la méthode de Dynamic Time Warping	-
<b>Z</b>		
$Z$	Zone géographique	-



# Table des matières

<b>Introduction générale</b>	<b>1</b>
<b>1 Bibliographical survey and analysis</b>	<b>13</b>
Introduction	15
1.1 Top-down models	18
1.1.1 Introduction	18
1.1.2 Deterministic statistical disaggregation models	18
1.2 Bottom-up models	21
1.2.1 Introduction	21
1.2.2 Statistical random model	22
1.2.3 Probabilistic empirical models	23
1.2.4 Time of use based models	30
1.3 Statistical-engineering model : Train et al.'s model	45
1.4 Cross analysis	47
1.4.1 Reading grid	47
1.4.2 Focused comparisons	50
1.4.3 Positioning the models on a pseudo 3D plot	51
Conclusion	53
1.5 Acknowledgment	53
<b>2 Analyse fonctionnelle</b>	<b>55</b>
Introduction du chapitre	55
2.1 Fonction « Modéliser le parc du territoire étudié (F 2) »	60
2.1.1 Définir les éléments-types du parc (F 2.1)	62
2.1.2 Définir la composition en n-uplets du parc habité année par année (F 2.2)	75
2.1.3 Contrôler la cohérence du parc (F 2.3)	84
2.2 Fonction « Scénariser et projeter l'évolution du contexte (F 1) »	86
2.2.1 Définir des événements-clés encadrant des périodes (F 1.1)	86
2.2.2 Définir les tendances d'évolution au sein des périodes (F 1.2)	87
2.3 Fonction « Calculer les courbes de charge (F 3) »	89
2.3.1 Construire le calendrier de simulation (F 3.1)	89
2.3.2 Calculer les courbes de charge unitaires par n-uplet (F 3.2)	91
2.3.3 Calculer les profils de charge par n-uplet (F 3.3)	96
2.3.4 Calculer les courbes de charge à l'échelle de la zone géographique (F 3.4)	98
2.3.5 Assurer la cohérence en énergie des courbes de charge (F 3.5)	99
2.3.6 Valider sur des données historiques les courbes de charge (F 3.6)	102
2.3.7 Calculer le solde fatal en énergie (F 3.7)	102
2.3.8 Répartir le solde fatal en énergie (F 3.8)	103
2.3.9 Sommer les courbes de charge usage par usage (F 3.9)	104
2.4 Fonction « Restituer les courbes de charge (F 4) »	105
2.4.1 Sélectionner le(s) format(s) de restitution (F 4.1)	105

2.4.2	Sélectionner les échelles spatiale(s) et temporelle(s) de restitution (F 4.2)	105
2.4.3	Calculer des indicateurs et construire des représentations caractéristiques (F 4.3)	105
	Conclusion du chapitre	107
<b>3</b>	<b>Modélisation des concepts et mise en œuvre par voie algorithmique</b>	<b>109</b>
	Introduction du chapitre	109
3.1	Modélisation des éléments conceptuels	110
3.1.1	Modélisation des jours-types	110
3.1.2	Modélisation des ménages-types	115
3.1.3	Modélisation des équipements domestiques d'électricité spécifique	116
3.2	Algorithme pour la génération de diversité d'origine comportementale	128
3.2.1	Distribution des tâches journalières	128
3.2.2	Création d'un profil basique de disponibilité	130
3.2.3	Méthode évoluée d'obtention des profils de disponibilité	133
3.2.4	Placement des équipements domestiques	136
3.3	Mise en perspective de l'algorithme	149
	Conclusion du chapitre	152
<b>4</b>	<b>Classification de courbes de charge et étude du foisonnement</b>	<b>153</b>
	Introduction du chapitre	153
4.1	Caractérisation et comparaison de courbes de charge	154
4.1.1	Indicateurs statistiques existants pour la comparaison de courbes de charge	154
4.1.2	Prise en compte de la dimension temporelle de la courbe de charge	161
4.1.3	Apports d'une méthode appliquée aux séries temporelles : le Dynamic Time Warping	162
4.2	Scénarios modélisés : étude des résultats	167
4.2.1	Construction de scénarios en vue de leur simulation	167
4.2.2	Étude des résultats du scénario principal	173
4.2.3	Étude des résultats en puissance du passage d'un cycle unitaire à un autre à iso-consommation	178
4.3	Mise en œuvre du <i>DTW</i> pour l'analyse de courbes de charge	181
4.3.1	Comparaison de deux courbes de charge	181
4.3.2	Comparaison de courbes de charge multiples	183
4.3.3	Relations entre la distance <i>DTW</i> et des indicateurs statistiques usuels	189
4.3.4	Décomposition des écarts entre courbes de charge moyennes foisonnées	191
4.3.5	Bilan de la mise en œuvre de l'algorithme <i>DTW</i> pour l'analyse de courbes de charge	195
4.4	Identification de comportements d'usage – Cas réel	196
4.4.1	Positionnement du problème et buts de l'étude	196
4.4.2	Présentation des données analysées	197
4.4.3	Principes des méthodes choisies d'analyse des données	199
4.4.4	ACP sur données brutes puis CAH sur les résultats de l'ACP	203
4.4.5	CAH sur les courbes de charge mesurées	211
4.4.6	Résumé des études supplémentaires de classification sur des courbes de charge normées de télévision	216
4.4.7	Bilan du travail de classification et perspectives	218
	Conclusion du chapitre	220
	<b>Conclusion générale et perspectives</b>	<b>221</b>
<b>A</b>	<b>Aide mémoire d'électricité</b>	<b>225</b>
<b>B</b>	<b>Détails de l'algorithme de placement des équipements</b>	<b>231</b>
<b>C</b>	<b>Étapes de simulation de notre modèle DYMOLA</b>	<b>235</b>

<b>D</b>	<b>Distance <math>DTW</math> et indicateurs sur <math>CDCMF</math></b>	<b>239</b>
<b>E</b>	<b>Étude des écarts d'origine temporelle et en puissance entre <math>CDCMF</math></b>	<b>243</b>
<b>F</b>	<b>Analyses complémentaires des courbes de charge de télévision</b>	<b>247</b>
	F.1 ACP puis CAH sur données normées . . . . .	247
	F.2 CAH sur les courbes de charge mesurées normées . . . . .	254
	<b>Bibliographie</b>	<b>271</b>
	<b>Liste des Figures</b>	<b>277</b>
	<b>Liste des Tableaux</b>	<b>281</b>
	<b>Glossaire</b>	<b>299</b>
	<b>Index</b>	<b>300</b>



# Introduction générale

« Chaque début d'écriture est un retour à la case départ. Et la case départ, c'est un endroit où l'on se sent très seul. Un endroit où aucun de vos accomplissements passés ne compte. »

Quentin Tarantino

## Contexte

LES limites et les insuffisances de la méthode expérimentale ont fait naître les activités de modélisation et de simulation nécessaires à l'Homme pour qu'il puisse comprendre les phénomènes qui l'entourent et solutionner les problèmes auxquels il doit faire face. Le terme de modélisation est relativement récent<sup>1</sup> dans le langage commun mais les travaux sur la théorie générale des systèmes, d'où résulte notamment celle de la modélisation, sont beaucoup plus anciens. J-L. Le Moigne dans son ouvrage, *La théorie du système général. Théorie de la modélisation* [LM94], a posé la problématique de façon très complète. Une vision plus générale<sup>2</sup> est proposée par E. Morin dans *La Méthode* [MOR04].

Parmi les principaux enjeux auxquels doit faire face le Monde à l'heure actuelle figure la problématique de l'approvisionnement énergétique d'une population mondiale en forte croissance (les Nations Unies prévoient que la Terre sera peuplée d'environ 9 milliards d'individus en 2050 [UNI04]). Assurer la fourniture en énergie dans un contexte de raréfaction des ressources fossiles et de lutte contre le réchauffement climatique soulève des problèmes nouveaux et complexes. Selon le point de vue adopté, des réponses technologiques, politiques, comportementales, sociétales... s'avèrent indispensables pour faire face à cet enjeu.

Dans le Monde en 2010, l'électricité représente 17,7 % de la consommation d'énergie finale [IEA12] et sa production, couplée à la fourniture de chaleur, furent responsables en 2009 de 41 % des émissions de CO<sub>2</sub> [IEA11]. Ce vecteur énergétique est donc en première ligne de l'enjeu planétaire précédemment discuté.

Dans cette thèse, nous cherchons à répondre à un objectif de modélisation qui vise à reconstituer la **courbe de charge**\* électrique résidentielle française. On appelle courbe de charge électrique, la représentation de la puissance électrique  $P$  exprimée en watt [W] en fonction du temps  $t$  mesuré en seconde [s] (les multiples ou sous-multiples de ces deux grandeurs peuvent être également utilisés). La courbe de charge traduit l'évolution au cours du temps de la consommation<sup>3</sup> d'électricité. Elle s'obtient par différentes méthodes [LB09] :

- via des **calculs** qui visent souvent à obtenir les appels de puissance théoriques d'un système ;
- par **extraction** : à partir de données de puissance agrégée (par exemple mesure à l'échelle d'un logement), des courbes de charge par usage ou par appareil sont recherchées par des méthodes statistiques, des techniques du traitement du signal<sup>4</sup>...

---

1. J. Giraud et al. dans [GPR74] définissent la modélisation comme étant « [l'] opération par laquelle on établit le modèle d'un système complexe, afin d'étudier plus commodément et de mesurer les effets sur ce système des variations de tel ou tel de ses éléments composants ».

2. E. Morin résume ainsi l'ambition de son œuvre majeure : « Nous avons besoin d'une méthode de connaissance qui traduise la complexité du réel, reconnaisse l'existence des êtres, approche le mystère des choses »

3. L'évolution dans le temps de la production d'électricité s'appelle **courbe de production**\*

4. Un exemple de décomposition de la courbe de charge résidentielle et tertiaire a été proposé par El Guedri dans [EL 09]

- via des **mesures** ce qui sous-entend une instrumentation préalable du périmètre étudié ;
- par **reconstitution** : cette dernière méthode fait appel à un modèle éponyme, modèle qui peut emprunter une ou plusieurs des trois techniques précédentes en plus de présenter ses propres caractéristiques.

L'intégrale sous la courbe de charge correspond à la quantité d'énergie consommée. Dans le contexte électrique, l'énergie est exprimée en wattheure [Wh] sur la période de temps considérée. Avant de poursuivre notre propos, nous indiquons au lecteur qu'un aide mémoire d'électricité est disponible au sein de l'annexe A. Celui-ci rappelle quelques notions de base sur l'énergie électrique. En outre, il nous amène à formuler l'hypothèse de base adoptée dans cette thèse qui est présentée ci-dessous.

Dans notre travail de modélisation des appels de puissance du secteur résidentiel, nous ne nous intéressons qu'à la composante active de la puissance électrique. Ainsi, dans l'ensemble du document (en dehors de l'annexe A), à chaque fois qu'il est question de puissance électrique, celle-ci correspond systématiquement à la **puissance active**\*

Nous proposons à la figure 1 une représentation graphique de la notion de courbe de charge qui est au cœur de notre travail de recherche. Soulignons également que des représentations sous forme de **puissance réduite**\* peuvent être employées notamment lorsqu'il est question de comparer diverses courbes de charge de manière relative<sup>5</sup>. La puissance réduite à l'instant  $t$ , notée  $P_r(t)$ , est une grandeur adimensionnelle égale au rapport entre la puissance instantanée à l'instant  $t$  et la valeur de la puissance prise pour référence (ce peut être par exemple la puissance moyenne ou maximale sur l'ensemble de la période étudiée).

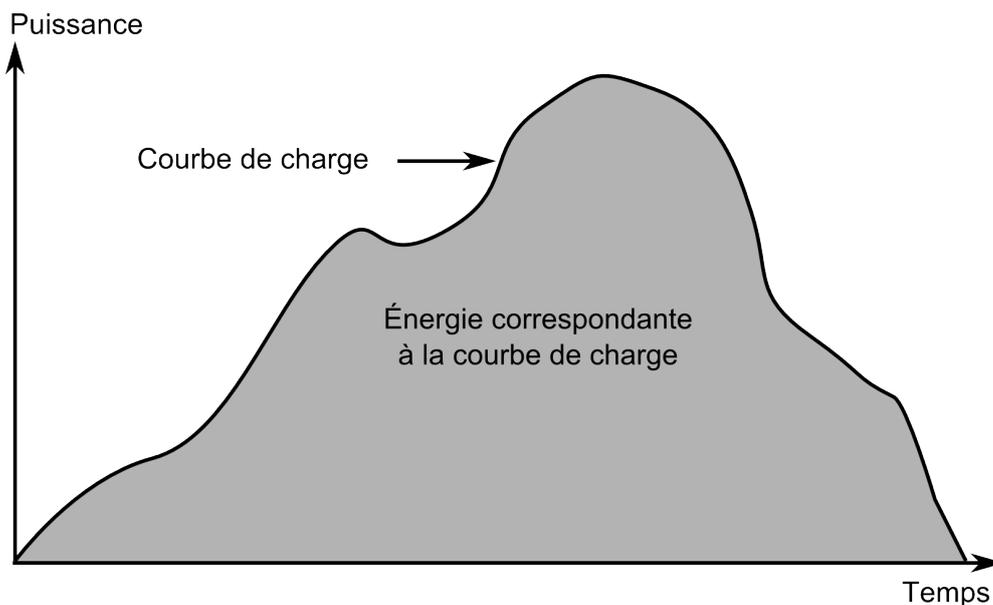


FIGURE 1 – Représentation graphique d'une courbe de charge et de l'énergie correspondante

Le but de la thèse est donc de reconstruire de façon dynamique les consommations instantanées en électricité des ménages français. La plus grande échelle géographique d'étude correspond à l'ensemble du parc hexagonal de logements.

En effet, contrairement aux secteurs industriel et tertiaire pour qui la demande en électricité est, à l'échelle journalière et même annuelle, relativement bien connue<sup>6</sup> et régulière<sup>7</sup> (les fluctuations peuvent

5. i. e. en ne s'intéressant qu'à leur forme respective sans juger les valeurs exactes d'appels de puissance

6. La majorité des clients tertiaires et industriels sont en effet soumis au comptage par courbe de charge depuis plusieurs années [REP00]

7. Régulière ne signifie pas que ces consommations d'électricité ne présentent aucune variation : il y a en effet une différence sensible de la courbe de charge de ces secteurs en fonction du type de jour ouvré/non ouvré par exemple

être considérées « lentes »), les besoins électriques des logements se caractérisent par une variabilité au quotidien et une **saisonnalité\*** importantes (causée notamment par les besoins de chauffage, de climatisation et d'éclairage). **Ces deux influences déforment nettement la courbe de charge électrique représentative de ce secteur.**

Nous illustrons par la figure 2 les variations que peut connaître la courbe de charge nationale tous secteurs, mesurée par le **Réseau de Transport d'Électricité (RTE)\***, et sur laquelle nous reviendrons par la suite.

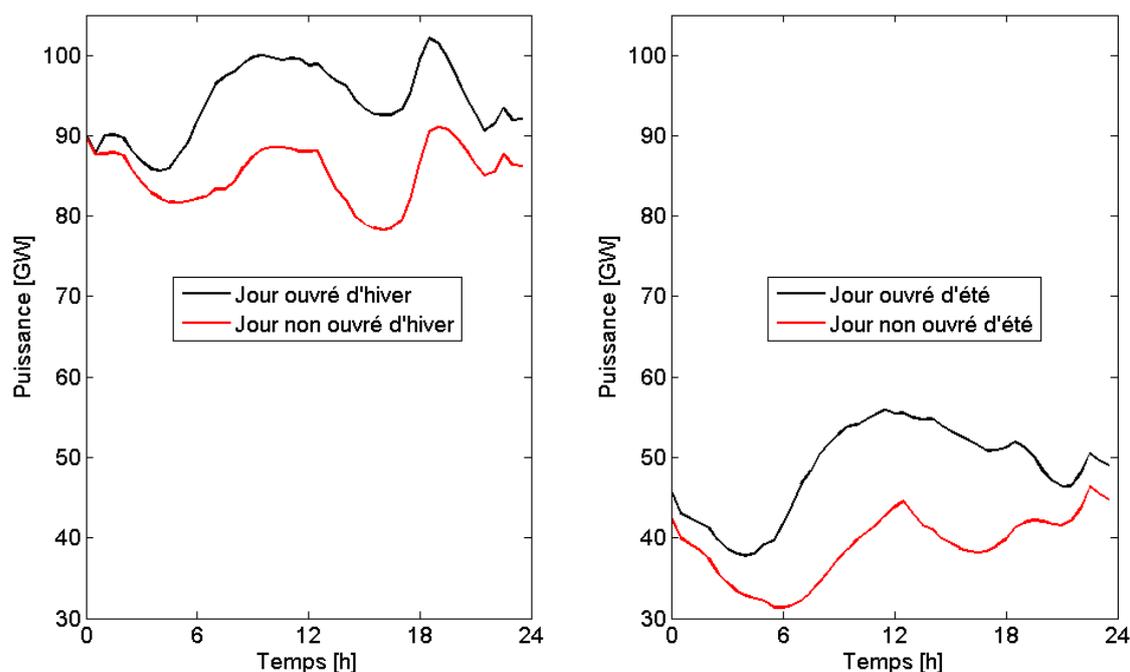


FIGURE 2 – Courbe de charge nationale pour 4 types de jour ouvré/non ouvré et hiver/été [RTE12c]

Les systèmes domestiques installés au sein des bâtiments résidentiels répondent aux besoins de leurs occupants. Traditionnellement dans ce secteur, ils se répartissent en trois grandes familles :

- les **usages thermiques** : chauffage, climatisation et **Eau Chaude Sanitaire (ECS)\*** ;
- les **usages liés à la cuisson alimentaire** ;
- les **usages dits d'électricité spécifique** : ceux-ci sont non concurrentiels de l'électricité c'est-à-dire que le service rendu ne peut l'être que par l'énergie électrique.

Jusqu'à présent en France, la consommation agrégée d'électricité spécifique croît sans cesse dans le secteur résidentiel (elle a augmenté de près de 40 % entre 1999 et 2010 – cf. figure 3). Plusieurs raisons expliquent ce phénomène. Tout d'abord, le nombre des ménages<sup>8</sup> croît dans la population française. Cette hausse est à la fois imputable à l'augmentation de la population et à la baisse du nombre moyen d'occupants par résidence principale [INS12c, INS12d]. Ensuite, un phénomène d'augmentation des **taux d'équipement\***<sup>9</sup> des ménages est notable<sup>10</sup>. Cette dynamique de croissance trouve sa source dans l'évolution des modes de vie ce qui engendre entre autres une tendance aux situations de **multi-équipement\***<sup>11</sup> pour certains

8. Au sens de l'INSEE, un ménage désigne l'ensemble des personnes habitant dans le même logement. De plus, il y a correspondance entre le nombre de ménages et celui des résidences principales

9. Indicateur en pourcentage de la part des foyers possédant un certain type d'appareils dans une population

10. Par exemple, le taux d'équipement en réfrigérateur est passé d'environ 10 % dans les années 1950 à près de 100 % aujourd'hui [INS10d] – d'autres exemples sont proposés dans le chapitre 2

11. Le multi-équipement désigne la « possession de deux appareils ou plus de même nature (soit dans la résidence principale, soit dans la résidence secondaire) » [INS11b]

appareils domestiques comme la télévision par exemple. L'augmentation de la consommation d'électricité des usages domestiques est également causée par la multiplication de ces appareils (intégration de nouveaux équipements d'électricité spécifique). Enfin, des usages nouveaux et en devenir utilisant l'électricité, notamment ceux qui concernent le transport individuel (**Véhicule Électrique (VE)\*** ou **Véhicule Hybride Rechargeable (VHR)\***), vont progressivement intégrer le parc d'équipements domestiques.

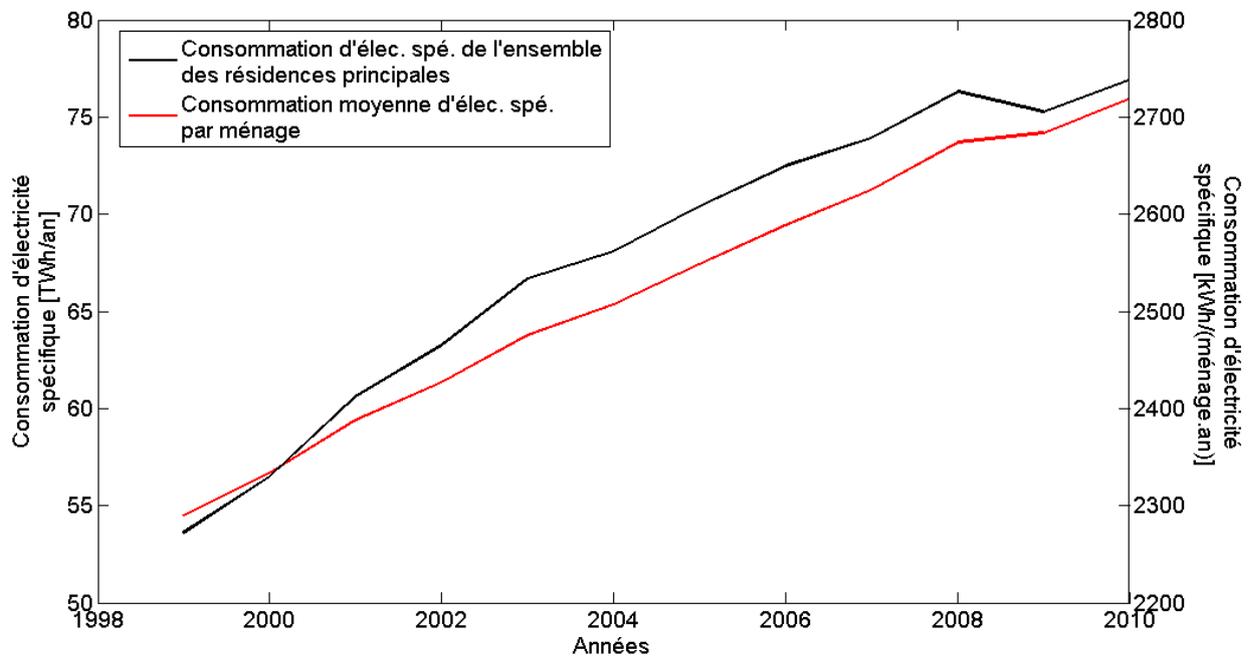


FIGURE 3 – Évolution de la consommation d'électricité spécifique des résidences principales françaises entre 1999 et 2010 [CER10]

Ces derniers usages vont être demandeurs d'une quantité importante d'électricité mais seront surtout sources d'appels de puissance particulièrement conséquents. Suivant la façon dont ils seront utilisés et la temporalité de la demande d'électricité ainsi engendrée, l'approvisionnement supplémentaire qui en découle pourrait accentuer les problématiques de **pointe de puissance\***<sup>12</sup>, c'est-à-dire d'appel maximal de puissance, sur le réseau électrique. À ce sujet, on constate une augmentation significative de l'appel annuel maximal de puissance en France qui est 2,5 fois plus rapide que la hausse de la consommation d'électricité lors des dix dernières années [RTE12b].

L'amplitude en puissance et la temporalité de ces besoins nouveaux en électricité sous-tendent de plus des impacts économiques et climatiques qui sont loin d'être anodins. En effet, les pointes de consommation d'électricité coûtent cher et ont un impact environnemental certain en matière de génération de CO<sub>2</sub>. Dans le cas du parc français, dont la composition en termes de capacités installées et dont la part de production par filière sont illustrées par les figures 4 et 5<sup>13</sup>, ce sont les moyens hydrauliques et thermiques à flamme (charbon, gaz et fioul) qui sont sollicités lors des appels de pointe sur le réseau électrique en plus des moyens utilisés tout au long de l'année (filiale nucléaire notamment) [UFE09a].

Plus concrètement, produire 1 kWh à la pointe d'un jour d'hiver est sensiblement plus onéreux, sur les deux plans précédemment cités, que ne l'est la fourniture d'une même quantité d'électricité au creux de la courbe de charge d'une journée d'été.

Selon la puissance consommée sur le réseau, différentes unités sont appelées pour assurer la fourniture

12. Rappelons ici que la pointe de puissance est un des principaux paramètres qui conditionne le dimensionnement des systèmes de production et du réseau de transport/distribution d'électricité [CRE12]

13. Sur ces figures, ENR\* est l'acronyme d'Énergies Renouvelables

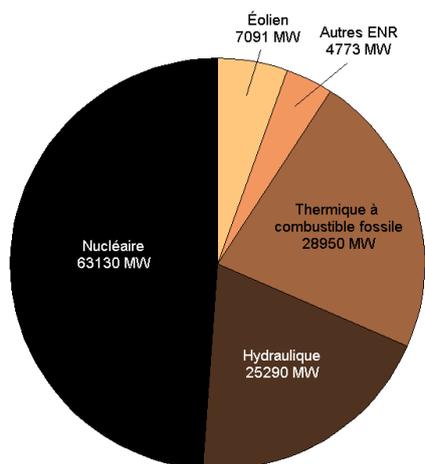


FIGURE 4 – Capacité installée en MW et par filière du parc français de production d'électricité en septembre 2012 [RTE12a]

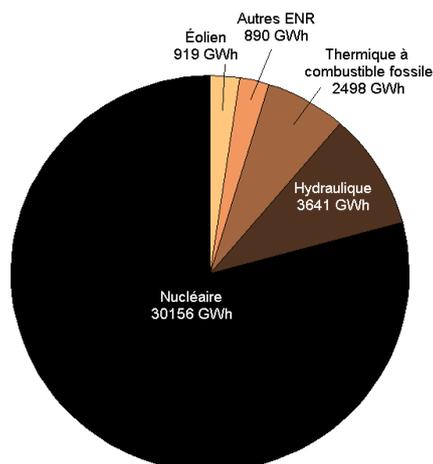


FIGURE 5 – Production française d'électricité par filière en septembre 2012 [RTE12a]

d'énergie électrique. Ainsi les 8760 heures d'une année de production sont classées par ordre décroissant du niveau de puissance. Cette représentation se nomme **monotone de puissance\*** et elle est illustrée à la figure 6<sup>14</sup>.

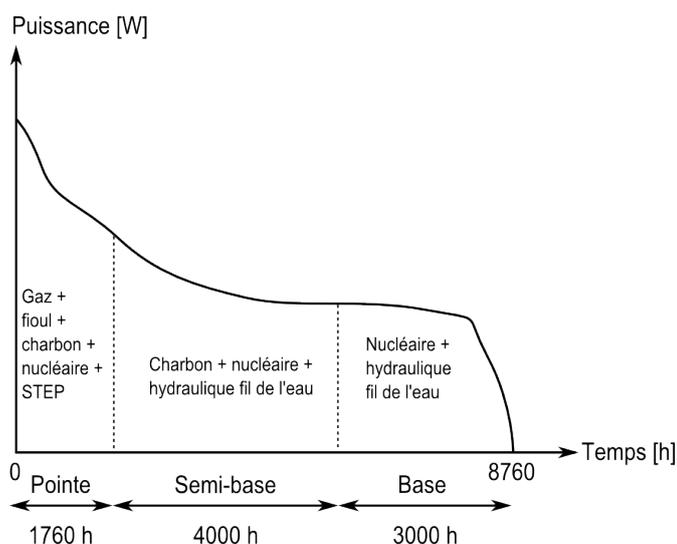


FIGURE 6 – Représentation graphique d'une monotone de puissance et principaux moyens de production associés par période [UFE09b]

Sur la monotone, l'évolution du coût de production de l'énergie fournie suit celle du niveau de puissance. D'autre part, l'échelle temporelle est découpée en trois périodes principales. La période de plus faible production dure 3000 h et est appelée base. Les heures de production intermédiaires forment la semi-base. Celle-ci a une durée de 4000 h. Enfin la dernière période, d'une durée de 1760 h constitue la pointe. Elle marque les heures durant lesquelles le réseau est le plus fortement sollicité [UFE09b].

Les contraintes précédemment identifiées pourraient s'amplifier du fait des ruptures qui affecteront le parc résidentiel. Or, les méthodes « classiques » de reconstitution de courbe de charge ne permettent pas de quantifier l'impact de ces contraintes sur les appels de puissance. Celles-ci s'appuient sur des données

14. Sur ce graphique, **STEP\*** désigne les Stations de Transfert d'Énergie par Pompage

passées et elles les déforment ou les ajustent suivant des procédures statistiques ou économétriques<sup>15</sup>. En effet, ce type de méthodologies n'est viable que pour de « petites » fluctuations autour des conditions de référence adoptées pour l'élaboration des modèles. Étant donné que ces usages récents et à venir ne sont pas encore identifiables sur la courbe de charge électrique du secteur résidentiel, leurs impacts sur cette dernière ne sont pas connus à ce jour. D'autre part, l'intégration dans le parc de logements de nouvelles technologies d'équipements, par introduction nette ou par substitution des technologies en place, s'accompagne le plus souvent d'une modification des besoins correspondants en électricité. Parmi les nombreux exemples existants, nous pouvons citer le cas du remplacement de chauffages à effet Joule par des pompes à chaleur performantes ou bien encore de la substitution des ampoules à incandescence par des dispositifs présentant une meilleure *efficacité énergétique*\* (les *Lampes Basse Consommation*\* par exemple).

Face à cet enjeu, les professionnels du secteur de l'électricité sont contraints d'anticiper les modifications identifiées ci-dessus afin d'adapter, dans une démarche prospective, l'ensemble de leurs moyens (production, transport, distribution, comptage...). Pour cela, ils doivent réaliser des simulations mettant en jeu différents scénarios d'évolution. Ils nécessitent alors des méthodes *techno-explicites*\* de modélisation (prise en compte de la dimension technique/technologique dans les procédures de calculs) de la courbe de charge, en particulier pour les appels de puissance émanant du secteur résidentiel. Actuellement en effet, **la courbe de charge électrique résidentielle n'est ni connue, ni mesurée à une échelle agrégée** en dehors des études sur les panels (en particulier ceux mis en place et/ou suivis par EDF\*). Ces modèles doivent en outre être capables de simuler des scénarios définis par un certain nombre de paramètres (méthodes paramétriques) et de restituer des résultats à différentes mailles géographiques ou plus généralement à différentes mailles d'*agrégation*\* (méthodes agrégatives). **L'élaboration d'un tel modèle est le sujet de cette thèse.**

## Données de courbes de charge disponibles à l'échelle nationale

Le RTE publie quotidiennement et en temps réel la courbe de charge d'électricité consommée (puissances au pas 15 min) sur le territoire national [RTE12e]. Celle-ci est estimée à partir de mesures :

- de la puissance injectée sur le réseau par les centrales de production et
- des transits aux frontières du réseau français.

À l'aide de ces mesures, la courbe d'appels de puissance point par point est déterminée grâce à l'équation 1.

$$\text{Consommation (pertes incluses)} = \text{Production} + \text{Imports} - \text{Exports} \quad (1)$$

La figure 7 présente les courbes de charge prévues<sup>16</sup> par RTE [RTE12e] et celle effectivement réalisée pour la journée du lundi 1<sup>er</sup> octobre 2012.

La courbe de charge de RTE, la seule publiquement disponible en plus de la *synchrone*\* d'ErDF\*<sup>17</sup> est celle de l'estimation de la consommation d'électricité en France continentale (i. e. hors Corse) et en omettant les puissances appelées pour les stations de pompage selon la formule 1. Elle s'obtient concrètement par cumul des données télé-relevées en intégrant les échanges avec l'international.

Les écarts à l'estimation sont imputables à l'imperfection des prévisions météorologiques ainsi qu'aux événements exceptionnels qui par nature sont difficilement prévisibles (voire impossible à anticiper) ou

15. C'est le cas par exemple dans [ASK84], [BFGL92] et [THW85]

16. RTE propose en effet deux estimations de la courbe de charge nationale journalière : une qui correspond à la prévision de la veille (à J-1) et l'autre est celle qui est calculée le jour J et mise en ligne à 14 h 30

17. Électricité Réseau Distribution de France – gestionnaire du réseau de distribution d'électricité – publie également à chaque début de mois, une synchrone c'est-à-dire la somme de courbes de charge d'un certain nombre de consommateurs ainsi que la courbe de production d'une sélection de fournisseurs. Dans ce cas, il s'agit d'un bilan des flux de puissance injectés sur le réseau de distribution par le réseau de transport d'électricité. Ces données s'appuient sur le meilleur historique connu mais ne sont pas utilisées pour reconstituer et valider les flux [ErD10a]

dont l'impact sur les consommations d'électricité n'est pas aisément quantifiable. La prévision de la courbe de charge est délivrée quotidiennement par le CNES (Centre National d'Exploitation du Système).

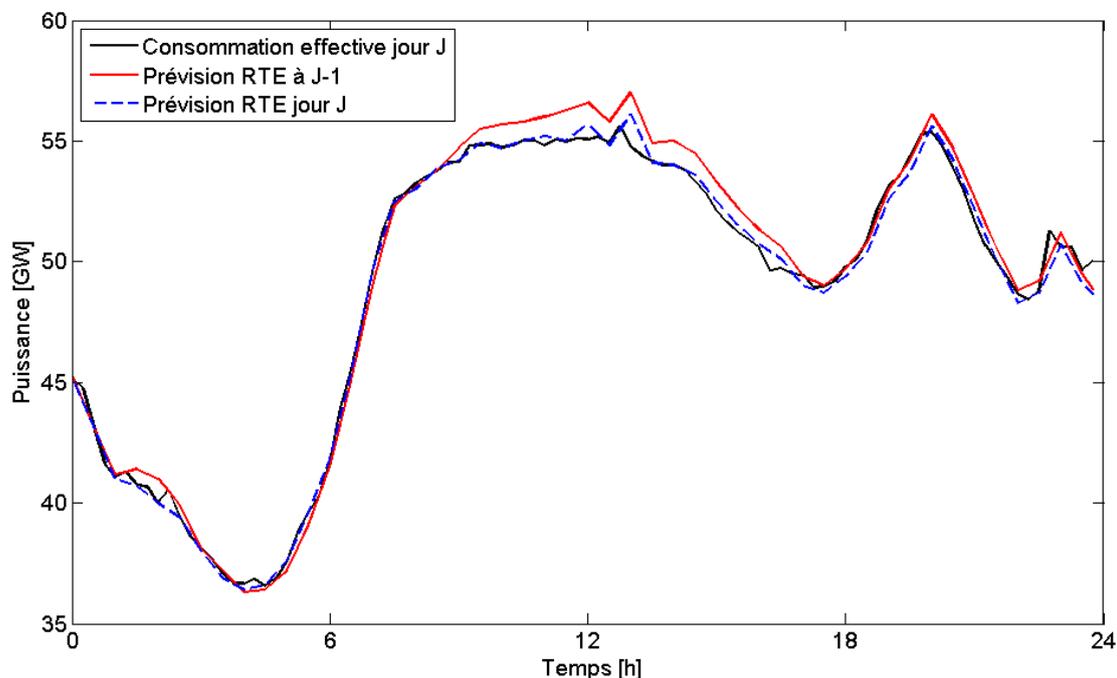


FIGURE 7 – Prévisions de RTE et réalisation d'une courbe de charge journalière nationale

En plus des données de consommation mises en ligne sur le site de RTE, ce dernier publie également un historique de production du parc de référence de l'Union Française de l'Électricité (UFE)\* ainsi que des données relatives à la disponibilité de celui-ci [RTE12d].

Il est intéressant de souligner que la puissance soutirée<sup>18</sup> sur le réseau de transport véritablement consommée (et non plus l'estimation de la courbe de charge) est également mesurée au niveau des postes sources (frontière entre réseau de répartition et réseau de distribution). Ces données ne sont cependant rendues disponibles qu'au travers des bilans annuels en énergie publiés par RTE.

Concernant les puissances injectées sur le réseau, les données disponibles sont des valeurs à la fois de réalisation (publiées « ex post ») et de prévision de disponibilité à la pointe (publiées « ex ante ») [RTE06] et qui sont agrégées par filière (nucléaire, charbon, gaz, fioul et moyens de pointe, hydraulique et autres). La délivrance d'informations par les producteurs n'est pas une obligation. À l'heure actuelle, cinq acteurs de ce secteur (EDF, GDF-SUEZ, E.ON, Poweo et ALPIQ) répondent à cette proposition de l'UFE [RTE12d] ce qui représente plus de 90 % du parc de production raccordé au réseau public de transport.

Enfin, le gestionnaire du réseau de transport d'électricité publie également tous les deux ans un bilan [RTE12b] qui établit un diagnostic prévisionnel à cinq ans de l'équilibre offre-demande du système électrique et qui élabore des scénarios prospectifs à long terme. Il sert aussi de support à la politique de Programmation Pluriannuelle des Investissements (PPI).

18. Remarquons qu'il ne s'agit pas forcément de la puissance consommée : c'est une sous-estimation de cette dernière grandeur s'il y a de la production en aval du poste source

## Enjeux de développement du modèle

Au commencement de la thèse, les modèles de la littérature capables de reconstituer la courbe de charge du secteur résidentiel qui étaient identifiés jusqu'alors, ne permettaient pas d'effectuer des projections non tendanciennes de ces appels de puissance. Leur limite se situe en particulier au niveau de la prise en compte d'évolutions structurelles du parc, qu'elles soient techniques ou comportementales. Les méthodes communément utilisées pour établir des prévisions font généralement l'usage de profils-types de consommation. Elles intègrent en outre les techniques et les comportements contemporains et passés.

Ces techniques prennent en compte, d'une manière plus ou moins fine, la variété des usages et/ou des équipements et le **foisonnement**\* qui en découle. Ce terme désigne le phénomène naturel qui se produit au sein d'un réseau électrique et qui traduit le degré de concomitance (et de non-concomitance) de la consommation d'électricité. En effet, il existe une forte diversité (« diversity » est d'ailleurs l'appellation anglo-saxonne du foisonnement) de paramètres et d'influences qui ont un impact sur les appels de puissance. Par exemple la courbe de charge d'un équipement domestique peut notamment être sujette aux éléments suivants :

- la composition du foyer dans lequel l'équipement est présent ;
- le scénario de présence journalière des occupants du logement ;
- la puissance souscrite et le tarif adopté qui en découle ;
- le revenu du ménage ;
- sa sensibilité aux prix ;
- sa consommation annuelle d'électricité ;
- les caractéristiques techniques de l'équipement ;
- le type d'habitat ;
- la zone climatique ;
- les conditions météorologiques ;
- ...

Le foisonnement est « la providence des électriciens » [LAN87] qui rend possible la fourniture de courant électrique à des coûts raisonnables. Grâce à lui, la pointe de consommation d'un groupe de clients alimentés par un réseau est toujours inférieure à la somme des puissances maximales de chacun des clients [LAN85, HUG83].

Du fait de sa complexité, le foisonnement est souvent inclus dans les données d'entrée des modèles de reconstitution de courbe de charge mais il n'est pas systématiquement explicité par ces derniers. En conséquence, ceux-ci ne permettent pas une prévision fiable de la courbe de charge lors de profonds changements et d'interrogations de type :

- quels impacts en puissance de la pénétration des équipements diffus d'électricité spécifique (appareils électroniques, recharge de téléphones portables, de tablettes tactiles...) qui font désormais partie de notre quotidien ?
- quels sont les impacts sur les appels instantanés en puissance de scénarios de rupture comme par exemple les mesures volontaristes de **Maîtrise de la Demande en Énergie**\*<sup>19</sup> (MDE\*) et de(s) l'éventuel(s) **effet(s) boomerang**\* associé(s) ?
- comment se déforme la courbe de charge suite à la pénétration de nouveaux usages n'ayant pas fait l'objet de panels de mesures : **Pompe À Chaleur (PAC)**\* avec ou sans stockage, VE et VHR, nouveaux **modes de gestion**\* et d'**effacement**\*<sup>20</sup> ?

19. Nous proposons la définition de Gellings [GEL96] reprise par Broc [BRO06] p. 18 : « La maîtrise de la demande en énergie désigne les actions conduites par les pouvoirs publics et par les producteurs et/ou distributeurs d'énergie, destinées à inciter et parfois à obliger les usagers d'un secteur d'activités à changer leur manière d'utiliser ou de consommer l'énergie. Dans ce cas, on va notamment chercher à modifier les appels de puissance des énergies non stockables et qui doivent être distribuées par réseau directement aux points de consommation »

20. Politique de MDE qui consiste à adapter temporairement le fonctionnement d'un certain type d'appareils, d'un usage – généralement le chauffage – ou d'une catégorie de consommateurs au cours de périodes infra-journalières

- quelle est la réponse en puissance de l'amélioration de l'efficacité énergétique des appareils et de l' (des) **effet(s) rebond\*** qui peut (peuvent) la compenser ?

Il y a donc des manques à différents niveaux du processus de modélisation de la courbe de charge résidentielle :

- traitement des consommations d'électricité spécifique en incluant les usages récents et nouveaux ;
- besoin d'explicitation du foisonnement des appels de puissance notamment par prise en compte des aspects comportementaux (rythmes et habitudes de vie, niveaux de confort, choix de gestion) ;
- prise en considération des ménages de manière individuelle, l'élément de base devant être l'équipement électrique domestique ;
- diversification statistiquement représentative des équipements (en prenant en considération l'aspect technologique) conformément à l'état du parc modélisé ;
- possibilité de réaliser un modèle intégrant une plus large multiplicité de facteurs d'influence (**gestion de la charge\***<sup>21</sup>, effacement, structure tarifaire...).

Il s'agit également d'établir un outil d'aide à la décision dans le cadre de la planification des moyens de production à long terme et de mieux contrôler les problèmes de pointe et d'aléas sur le réseau (notamment la déformation de la courbe de charge à échéance moyenne ou lointaine). Celui-ci pourrait être évalué par des données panel et en première approximation, il est envisageable de le « caler » (i. e. calibrer) de manière à ce que la courbe de charge résultante conduise à la consommation annuelle d'électricité effective et mesurée.

En résumé à l'heure d'entamer nos recherches, il n'existait pas à notre connaissance de modèle **paramétrique**, **bottom-up\*** (l'unité de base considérée est l'équipement électrique domestique au sein d'un ménage), **agrégatif** et **techno-explicite** suffisamment développé pour simuler et obtenir la courbe de charge du secteur résidentiel de façon non tendancielle, c'est-à-dire en intégrant les différentes ruptures techniques, technologiques, politiques, comportementales, sociétales... ainsi que les nouveaux usages.

## État des lieux des méthodes actuelles

Malgré le nombre et la diversité des modèles de courbe de charge identifiés et analysés dans le chapitre 1, aucun d'entre eux n'a pu être retenu comme méthode « clé en main » répondant à nos attentes en matière de modélisation. Plusieurs raisons sont à l'origine de cette insatisfaction :

- les spécificités du parc français sont peu abordées par les modèles ;
- ils sont en général moins riches en matière de diversité d'appareils domestiques pris en compte dans les simulations que celle correspondante à nos attentes ;
- la modélisation des appels de puissance pour un appareil ne tient compte que très rarement de sa technologie ;
- l'intégration des nouveaux usages n'est pas toujours possible en fonction de la méthodologie adoptée par les auteurs ;
- les ruptures technologiques, les nouveaux souhaits de confort, l'évolution des habitudes de vie et du comportement des occupants des logements ne sont en général pas modélisés de manière satisfaisante ;
- la modélisation des usages thermiques – le chauffage notamment – est effectuée par des méthodes statistiques ou des modèles thermiques. Bien souvent, ils simplifient trop la réalité et sont en outre difficilement réutilisables pour le cas de la France. L'Hexagone présente en effet la particularité d'un parc de production majoritairement basé sur la filière nucléaire lui permettant ainsi de mettre en avant l'électricité pour recouvrir aux besoins thermiques du secteur résidentiel entre autres. Ceci constitue une singularité<sup>22</sup> à l'international.

21. Il s'agit de l'« ensemble des actions prises par un distributeur d'énergie pour modifier la demande en fonction de contraintes financières et techniques reliées à la production et à la distribution de l'énergie » [Gra12]

22. La Norvège, grâce à son parc de production hydraulique, est dans un cas similaire à celui de la France dans ce domaine

En France plus particulièrement, la méthode jusqu'alors utilisée dans le secteur résidentiel pour estimer la courbe de charge de différentes catégories de consommateurs domestiques est appelée **profilage**\*. Celui-ci est employé par les gestionnaires du réseau public, ErDF en tête [ErD10b]. Le but du profilage est de calculer les consommations et/ou les productions des clients non sujets au comptage à courbe de charge [GE04]. Sont concernés à la fois des grands consommateurs d'électricité ou bien des consommateurs de masse. Ce dispositif a été mis en place sur une décision de la Commission de Régulation de l'Énergie (CRE) au moment de l'ouverture des marchés. Cet outil constitue en effet une aide pour les nouveaux entrants sur le marché de l'électricité (fournisseurs non historiques).

Le profilage abouti ainsi à des **profils de consommation**\* – courbes de charge représentatives d'un certain type de clients – obtenus à partir de télé-relevés d'index. Les « profils types » sont constitués par trois jeux de coefficients :

- 52 coefficients semaine qui tiennent compte de la variabilité de consommation hebdomadaire sur l'année ;
- 365 coefficients jour qui modélisent des fluctuations des besoins journaliers en électricité sur une semaine ;
- 17472 coefficients demi-horaire qui témoignent de variabilités infra-journalières.

L'aléa climatique est ainsi pris en compte dans l'élaboration de ces profils. La courbe de charge réalisée est mesurée de manière à améliorer les futures prévisions en « recalant » l'estimation sur la réalisation et ainsi affiner les coefficients correctifs. Au final, le gestionnaire du réseau de distribution cherche à obtenir une courbe de charge de l'ensemble des consommations avec une représentation statistiquement bonne des comportements de chacun.

Cette approche présente toutefois plusieurs limites détaillées ci-dessous :

- **vraisemblance plutôt qu'exactitude** : l'objectif de cet instrument est d'avoir une bonne approximation de la courbe de charge de chacun des fournisseurs et des responsables d'équilibre mis en jeu plutôt que de rechercher la courbe de charge exacte de chaque site ;
- **représentativité plutôt qu'exhaustivité** : le profilage s'appuie sur des échantillons représentatifs des clients non télé-relevés. Au sein des échantillons (i. e. sur un nombre limité de clients), les mesures de courbes de charge et la prise en compte de l'aléa climatique sont réalisées de manière exhaustive afin de pouvoir extrapoler les résultats à l'ensemble des consommateurs non télé-relevés.
- **extrapolation des consommations passées pour une prévision des besoins futurs** : tout ce qui n'est pas mesuré jusqu'alors ne peut pas être intégré dans les profils ;
- **les profils représentent la consommation électrique tous usages confondus** : il n'est donc pas possible de mesurer l'impact sur la courbe de charge lié à l'évolution d'un certain usage ;
- **le foisonnement des courbes de charge n'est pas explicité** puisqu'il provient des mesures ;
- **stricte dépendance des profils à la structure tarifaire actuelle** : il n'est alors pas possible d'envisager les courbes de charge de consommateurs domestiques selon d'autres modes de tarification non encore existants ;
- **restitution des courbes de charge au pas demi-horaire** alors que la dynamique réelle des équipements domestiques d'électricité spécifique, discutée ci-après, se situe bien en dessous de cette échelle de temps.

C'est en identifiant ces limites que nous nous sommes convaincus de la pertinence et de la nécessité d'une analyse fine du besoin en ce qui concerne le modèle de courbe de charge recherché dans cette thèse. Parmi les spécifications essentielles figurent l'explicitation du foisonnement ainsi que la restitution des courbes de charge à un pas de temps relativement fin.

Dans notre travail, nous avons développé un modèle de courbe de charge dont la spécificité est une capacité de générer du foisonnement. Notre champ d'étude fut celui des équipements d'électricité spécifique, ensemble qui représente une consommation importante d'énergie électrique (la

[IEA12]

figure 8 indique que la consommation électrique imputable aux usages spécifiques dépasse en 2010 celle des usages thermiques : chauffage et ECS). **Toutefois, notre modèle est tout à fait adaptable aux usages thermiques.**

Générer de la diversité d'usage des équipements domestiques implique de s'intéresser à la dynamique des temps de vie au sein des ménages. Concrètement, la temporalité de l'utilisation des appareils d'électricité spécifique est de l'ordre de la minute alors qu'elle est d'environ une demi-heure pour les usages thermiques. **Pour assurer une comparaison possible des résultats de notre modèle, nous avons choisi 10 min<sup>23</sup> comme pas de temps de restitution des courbes de charge.**

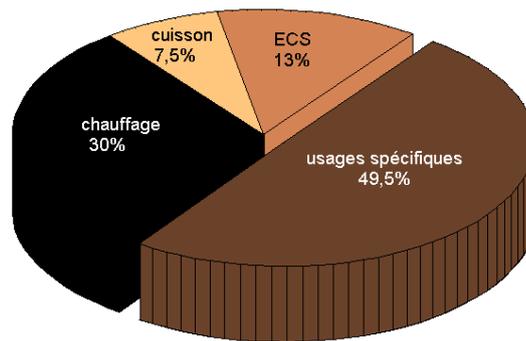


FIGURE 8 – Finalités des consommations d'électricité du secteur résidentiel français (153,1 TWh) en 2010 à climat normal [CER10]

## Démarche scientifique et plan de la thèse

Pour mener à bien nos recherches et plus particulièrement dans le but d'élaborer un nouveau modèle de reconstitution de courbe de charge suivant les contraintes que nous avons évoquées plus haut, nous avons décidé de mener une démarche en quatre étapes principales, chacune d'elles étant détaillée par un des chapitres qui constituent le présent manuscrit. L'ordre de ces chapitres correspond également à la chronologie générale de la progression dans notre travail, même si certains points présentés en fin de ce rapport ont pu être abordés avant d'autres détaillés au préalable.

**Le chapitre 1** synthétise les travaux de recherches et d'analyses bibliographiques que nous avons conduits. En premier lieu en effet, nous avons jugé indispensable d'établir un état de l'art relatif à notre sujet d'étude. Pour cela, nous nous sommes concentrés sur l'identification, l'analyse, la comparaison et la classification des modèles de la littérature qui, selon nous, se rapprochent le plus de nos exigences et des contraintes auxquelles nous devons répondre. L'étude de 12 modèles présentée dans ce premier chapitre nous a permis d'identifier les avantages et les limites de chacun d'eux. Ce travail a été également l'occasion pour nous de mettre en lumière un apparent manque dans l'espace des modèles pouvant restituer des courbes de charge résidentielles, pour l'ensemble des usages domestiques, à un pas de temps fin<sup>24</sup> et en explicitant le phénomène de foisonnement.

23. Ce pas de temps correspond en effet à la fréquence maximale de relève des compteurs domestiques Linky [CRE11] et à celle des mesures sur le panel 6000 d'EDF

24. Rappelons que le pas 10 min pour l'électricité spécifique et 30 min pour les usages thermiques constituent nos repères

**Le chapitre 2** présente la méthode conceptuelle de reconstitution de courbe de charge à laquelle nous avons abouti. À l'issue de la partie bibliographique et forts des conclusions que nous avons pu tirer de ces recherches, nous avons débuté nos réflexions en vue de la construction d'une méthode conceptuelle originale de génération des appels domestiques de puissance. Gardant en permanence à l'esprit les spécificités du modèle auquel nous voulions aboutir, nous avons choisi d'entreprendre une démarche d'analyse fonctionnelle. Cette intense phase de réflexion a conduit à la définition d'un certain nombre de concepts et à l'élaboration de procédures de calculs. L'ensemble de ce travail a ensuite été regroupé et résumé sous la forme d'un cahier des charges fonctionnel. Nous avons conservé dans ce chapitre le formalisme de cahier des charges, ceci pour deux raisons : premièrement, nous pensons qu'il s'agit de la manière la plus simple pour le lecteur de s'approprier cette méthode complexe. D'autre part et pour une raison moins essentielle, la forme de cahier des charges fonctionnel rappelle ainsi la méthodologie employée pour parvenir à ce résultat conceptuel.

**Le chapitre 3** expose le travail de modélisation que nous avons mené et décrit l'algorithme stochastique qui en découle. La méthode conceptuelle représente à nos yeux le principal apport de la thèse. Nous nous sommes attachés à la mettre en œuvre en vue de la rendre opérationnelle au travers de simulations aboutissant à des résultats exprimés en puissance électrique.

De ce fait, au sortir de la phase d'analyse fonctionnelle, nous avons entamé la construction du modèle de reconstitution de courbe de charge à proprement parler. Pour cela, nous avons engagé un travail de modélisation pour un certain nombre de concepts détaillés dans le second chapitre. Plus précisément, nous nous sommes concentrés sur les éléments et les procédures de la méthode nécessaires à la génération de diversité d'origine comportementale, celle-ci étant inhérente à l'utilisation de l'électricité au sein d'une population. Pour reproduire le foisonnement d'origine anthropologique, nous avons mis en place un algorithme stochastique chargé de simuler l'usage réaliste des appareils domestiques au sein des ménages.

**Le chapitre 4** se propose finalement d'examiner en détails trois thématiques de premier plan sous forme d'une étude expérimentale. Pour répondre au besoin de quantification du foisonnement, nous avons tout d'abord entrepris de synthétiser les éléments descriptifs des courbes d'appels de puissance en analysant les apports de chacun et en identifiant leurs manques respectifs. Simultanément nous avons réfléchi à la manière de mesurer l'écart entre courbes de charge. Une telle démarche n'est pas triviale du fait que ces dernières peuvent se différencier aussi bien en temporalité qu'en amplitude. Nous avons valorisé une distance adaptée à la courbe de charge pour aboutir à une méthodologie originale d'analyse du foisonnement affectant les appels de puissance à un niveau agrégé.

Cette méthodologie a été ensuite mise en œuvre sur les résultats en puissance issus de notre algorithme. Pour cela, nous avons défini et simulé plusieurs scénarios de modélisation.

Enfin nous avons cherché à identifier des typologies de comportements relatifs à l'usage de l'énergie électrique dans le secteur résidentiel. Un travail de classification a donc été mené sur les résultats d'une campagne de mesures de courbes de charge.

À l'issue de ces quatre chapitres, nous faisons le bilan de la thèse et nous suggérons des pistes de recherche pour poursuivre le travail présenté dans ce manuscrit.

# Chapitre 1

## A review and an analysis of the residential electric load curve models

*« Les articles de fond ne remontent jamais à la surface »*  
Boris Vian

### Résumé en français de l'article

Dans ce chapitre, que nous présentons sous la forme de l'article qui a été publié dans la revue *Renewable and Sustainable Energy Reviews* [GAB12], nous procédons à l'analyse et à la comparaison de modèles de reconstitution de courbe de charge du secteur résidentiel. Ceux-ci sont décrits de manière plus ou moins détaillée au travers d'articles de journaux scientifiques ou dans des manuscrits de thèse. Au sein de la littérature, nous avons identifié de nombreuses méthodes statistiques, mathématiques, physiques ou encore mixtes (ce terme désigne les méthodes qui valorisent plusieurs des précédentes approches) qui permettent d'évaluer les besoins domestiques en puissance d'un certain nombre de logements. Les modèles que nous avons retenus dans ce chapitre présentent des similarités notables avec les objectifs et/ou la mise en œuvre de la thèse (modélisation au niveau de l'équipement électrique individuel, restitution de courbes de charge journalières à différents [niveaux d'agrégation](#)<sup>\*</sup>, prise en compte de la problématique de génération de diversité) sans pour autant coïncider totalement avec eux. Ainsi, nous n'avons pas trouvé de méthode « clé en main » satisfaisant totalement les spécifications de modélisation qui étaient établies au commencement de la thèse. Toutefois, ces recherches bibliographiques nous ont permis de saisir les particularités et les limites des différentes approches qui découlent à la fois de l'objectif en matière de résultats à atteindre (sous forme de courbes de charge) mais aussi de la nature, de la richesse et du volume de données que les auteurs ont eu à disposition. En outre, la rédaction et la publication d'une telle revue bibliographique semble combler le manque d'articles proposant ce type de recherches.

L'étude des modèles a été conduite en gardant à l'esprit notre propre objectif de modélisation c'est-à-dire que nous avons mené notre analyse en regard d'un modèle idéal, inexistant à notre connaissance au début de nos recherches, et qui satisfait strictement l'ensemble des exigences de la thèse. Les objectifs de celle-ci sont d'aboutir à une méthode de reconstitution des appels de puissance de type bottom-up, techno-explicite, agrégative et évolutive capable de simuler tous types d'usage domestique, a fortiori les usages thermiques, ceux relatifs à la cuisson et les équipements d'électricité spécifique. De ce fait, nous avons principalement axé nos efforts sur deux sortes d'approches : en priorité les méthodes ascendantes ou bottom-up mais les approches descendantes ou top-down n'ont pas été ignorées. Les caractéristiques générales de chacune d'elles sont abordées plus en détails dans la partie 1 de l'article mais aussi dans l'introduction de chacune des sous-parties dédiées à ces deux méthodologies (paragraphes 1.1.1 et 1.2.1). Nous verrons également au sein de la section 1.3 que des méthodes hybrides, qui cherchent à concilier les avantages des deux approches précédentes, ont été proposées.

Jugeant relativement faible cette distinction traditionnelle entre les méthodes, nous avons défini une classification nouvelle basée sur la nature du processus permettant aux modèles de recréer de la diversité en matière d'appels de puissance. Nous caractérisons dans l'introduction de l'article (paragraphe 1) les cinq catégories de modèles que nous suggérons. Dans le cœur de l'article, nous décrivons les approches de la littérature conformément à l'ordre établi de ces modalités dans la partie 1 (cf. paragraphes 1.1.2, 1.2.2, 1.2.3, 1.2.4 et 1.3). La diversité des appels de puissance est une caractéristique essentielle puisque c'est elle qui permet la rentabilité de l'activité de fourniture d'électricité à une population d'individus. Plus concrètement, elle est responsable du phénomène de foisonnement des consommateurs d'électricité.

La manière avec laquelle est générée la diversité au sein des modèles n'est qu'un des éléments qui sont étudiés pour chacun d'eux. Plus précisément, nous avons opéré une lecture approfondie et une analyse systématique des méthodes suivant un processus bien déterminé. Ceci nous a permis d'aboutir à la carte d'identité de chaque approche : celle-ci comporte notamment une représentation du modèle sous forme de schéma-blocs qui identifie les données d'entrée et de sortie de la méthode ainsi que ses éléments opérationnels. Le caractère uniforme du descriptif des modèles a facilité leur comparaison (voir la partie 1.4) et la mise en place, au sein du paragraphe 1.4.1, d'une grille de lecture qui reprend les principales caractéristiques de chacun d'eux. À la suite de ces tableaux récapitulatifs, nous opérons une comparaison plus approfondie des modèles sur deux points précis que sont la méthode de génération de diversité et la façon de scénariser l'utilisation journalière des équipements domestiques (voir la section 1.4.2). Enfin, au sein du paragraphe 1.4.3, nous proposons une illustration de la classification des modèles à laquelle nous avons abouti sous forme de représentation pseudo tridimensionnelle. Les trois critères qui ont été sélectionnés pour ce graphique sont le type de méthode employée pour générer du foisonnement, la résolution temporelle des modèles et leur champ d'application.

En définitive, ce travail d'analyse bibliographique identifie un **manque de modèles** capables de générer des courbes de charge pour l'ensemble des usages domestiques, à un pas de temps fin et en explicitant le foisonnement qui impacte les appels de puissance à un niveau agrégé.

## Abstract

Due to the growth of electric end-uses, the management of the variations in time of the electric power demand has become essential, especially in the residential sector. According to this issue, the anticipation of the power demand is of great interest. This implies a better knowledge of the electric load curve of the household stock. Papers about understanding and forecasting energy demand are numerous but studies on building's load curves are rare. In this paper we propose a cross analysis of some existing methods capable of building up a residential electric load curve. Two main types of load curve models have been identified in the literature : top-down and bottom-up methods. Even if the review presents two existing top-down approaches, the authors focused the further analysis on bottom-up models. For each of them we first identify its functional characteristics : finality and scope, input data required, output format, modeled appliances and end-uses covered, generation of the diversity and validation of the model. Secondly, we establish a bloc diagram representing its architecture with focus on the mathematical model chosen. Finally, the authors list the limits of the model in view of the criteria needed to build up an ideal, bottom-up and technically explicit load curve model for the residential sector.

A cross reading of the different methodologies is proposed through a global table that characterizes and sums up the analyzed models. Moreover, a graphical representation of the models studied is proposed, according to three criteria (range of application, modelling of the diversity, time scale accuracy), that allows us to compare them « at a glance ». To conclude, an identification of the gaps among existing models is proposed. It consists of listing the end-uses, the appliances and other behavioural correlations which may affect the load curve and that are not included in the reviewed methodologies and need further research.

Keywords : load curve modelling, literature survey, classification, residential sector

## Introduction

Modelling the electricity consumption of the residential sector is not a recent research topic. For a few decades authors have published their methods developed to anticipate the temporal evolution of the electric energy demand with different time and space scale considerations. Among the motivations for developing such a model, the main one is the need of a tool to better predict, quantify and plan the future requirements in terms of power plants. Other reasons are the electric network control and management or the impacts of various evolution scenarios that focus on technical, demographic, behavioural or economical aspect.

For such purposes the authors generally focused on the annual electric energy demand and the chosen methodologies are different depending on the finality, the scope, the input data and the output format of each model.

Historically, the methods have been divided into two categories : top-down and bottom-up approaches. Swan and Ugursal indicate that this classification « is with reference to the hierarchal position of data inputs as compared to the housing sector as a whole » [SU09]. Models operate according to one of the processes represented in figure 1.1. However, some authors tried to benefit from the advantages of both approaches and thus develop hybrid models.

Swan and Ugursal [SU09] have classified the domestic electricity demand models of the literature with regard to the method used. They established a finer segmentation of the available methods than the simple differentiation between top-down and bottom-up methods. In this paper, we carried out a similar, global and transverse analysis of the methodologies for modelling the electric power demand.

The power demand issue is a more recent research topic than the study of electric energy consumption. However, many references exploring the load curve research field have been published since the 1940's : Hamilton [HAM42, HAM44] and Bary [BAR45] respectively focused on the properties of aggregated load curves and on the mathematical relationships between load curve characteristic factors (coincidence factor, load factor). In Europe, UNPEDE [UNI73] published the first load curve manual to clarify the

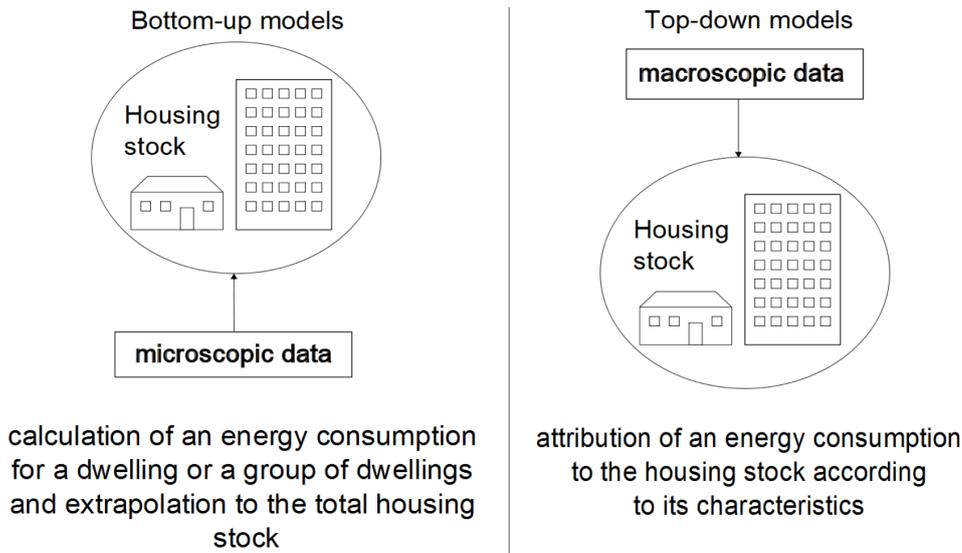


FIGURE 1.1 – Top-down vs bottom-up models

load curve concepts and terminology. Rehberg [REH46] proposed a publication about the study of the residential load curve. In the United States, the Association of Edison Illuminating Companies (AEIC) [AEI51, AEI61, AEI01] published several versions of the procedures to carry out load curve surveys.

In fact power demand is more complicated to predict than energy demand because of its random nature and its acute fluctuating aspect. On top of that, even if customers can have the same general characteristics (e.g. two domestic customers with the same household size, type of dwelling, building characteristics, appliances ownership...) there is every chance that their corresponding power demand, so their load curves for selected days, are completely different. Diversity that represents both the non-coincidence in energy use and an unlimited variety of customers' characteristics is responsible for this effect.

The influence of the human behaviour on the domestic power demand is so important that there is every chance for instance that two households with the same daily energy consumption will not show a similar load curve. That is the reason why modelling the power demand in a specific area for a specified target is also an hard task.

The power demand issue was explored to ensure the safety of electricity supply, to better predict the peak power demand, to optimize the network control techniques and to analyze the impact of Demand Side Management (DSM) strategies or the modification of the network load flows after integration of renewable energy sources.

The aim of this paper was the establishment of a new and improved classification of the residential load curve models described in the recent publications. This segmentation can be seen in figure 1.2.

We decided to analyze these models with reference to an ideal and to our knowledge non-existing model. To our opinion such an ideal model should present the following characteristics :

- it has to be parametric in order to simulate various scenarios ;
- it has to be technically explicit i. e. the different specificities of the simulated elements (equipments, buildings...) must specifically impact the load curve calculations and results ;
- it has to be evolutive, i. e. new elements can be introduced in the model so as to be simulated ;
- it has to be aggregative so that results can be obtained at different levels (household, area, city, region...);
- all domestic end-uses must be considered in the load curve calculations : Heating, Ventilation and

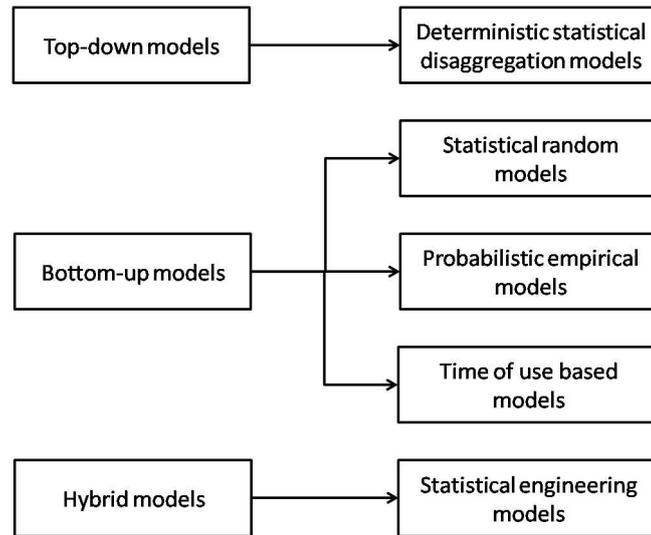


FIGURE 1.2 – Proposed classification of the load curve models

Air Conditionning (HVAC), Domestic Hot Water (DHW), cooking devices and specific electricity appliances.

According to the manner diversity is gotten, we define five types of models :

- deterministic statistical disaggregation model : this approach consists of disaggregating measured load profiles to identify various appliances. Diversity is not modelled at all because it is embedded, in a deterministic way, in the measured data ;
- statistical random model : in such a model, the reconstitution of the diversity makes use of statistical data. To generate variations for a given scenario, a random procedure is applicated ;
- probabilistic empirical model : from real collected data concerning, among other things, domestic habits of people, probabilistic procedures are defined and applied to generate a diversity of results ;
- Time of use (TOU) based model : diversity is constructed thanks to data coming from time of use surveys (real and precise data concerning the behaviour of people) ;
- statistical engineering model : in this case, the diversity is partly embedded in the measured data that serve as input (dwelling characteristics, weather data, penetration rates. . .). On top of that, diversity is embedded in the statistical coefficients that adjust the original results. These coefficients are calculated with the help of measured load curves and socio-economic data.

To carry out our analysis, we decided to follow a generic reading process for each model. We will focus on a series of criteria in order to build a standardized ID for each approach. This will help us to set the cross reading table at the end of this paper, all models considered.

In section 1.1, after a short introduction about the top-down approach, the authors review two of them that use a deterministic statistical disaggregation method. Section 1.2 first provides a short description of the bottom-up methodology, then an analysis of the corresponding models. These methodologies are classified according to the proposed segmentation. In the following, section 1.3 exposes a model which is not sortable into the two main families : it's an hybrid method. At the end of this paper, a comparison of all methods, following a cross analysis procedure, is carried out. The main properties of the models are put together into a synthetic table and a graphical representation using three criteria is drawn.

## 1.1 Top-down models

### 1.1.1 Introduction

Top-down methodologies consider situation on a whole (making use of national energy statistics for instance) and try to attribute an electricity consumption to the studied household stock with regard to its characteristics. The input data for this type of model are very general information such as :

- Gross Domestic Product (GDP) ;
- unemployment rate ;
- present statistics on the targeted population with possibly predicted evolution ;
- appliance saturation rates. . .

Generally these methodologies consist of a mathematical identification of the past and present electricity demand for a specified area. On this basis, the model may be used to generate trend evolutions : that can be well predicted by this type of model. The advantages of such a method is its simplicity, the use of generally well available and aggregated data and the possibility to calibrate the simulation tool with real electricity consumption values. Yet these models are unable to simulate the impacts of technical breaks or the appearance of new end-uses which have not been measured.

### 1.1.2 Deterministic statistical disaggregation models

In this section we present two methodologies whose aim is the analysis of residential load curve. In fact, total load curve (i. e. the aggregated power demand of a dwelling, all end-uses being considered) measurements are used as input data and the models try then to differentiate the individual contribution of the observed end-uses. This work corresponds to the load curve analysis and it makes use of mathematical techniques such as regression, statistical or econometric methods. In this approach, the authors try to explain the domestic electric power demand (dependent variable) with the help of some various and independent variables.

#### 1.1.2.1 Aigner et al.'s model

Aigner et al. present in [ASK84] the methodology they developed in order to get the hourly load curves of selected domestic end-uses for households located in areas next to Los Angeles. This analysis technique is in fact the same that Parti and Parti [PP80] developed for the domestic energy consumption analysis : this is called Conditional Demand Analysis (often mentioned as CDA). Aigner et al. reuse the same mathematical development made in [PP80] (the base CDA) but the authors improved it so that the variable which is explained becomes the hourly electric power demand.

As input data, the authors used on the one hand total load curves measured at a 15 min time step and on the other hand appliance penetration rates for the selected equipments.

CDA model uses twenty-four regression equations (one equation for each hour of a day). Each mathematical relation tries to determine the hourly variation of the customers group's load curve compared to the mean daily power demand value<sup>1</sup>. The variables which explain the power demand are :

- the desired temperature in the home (two scalar variables) : they are set with taking into account the outdoor temperature of the geographical area<sup>2</sup> ;
- the size of the dwelling (three scalar variables informing the number of rooms) ;
- nine dummy variables representing the presence or absence of the observed appliances<sup>3</sup> in a specified household (binary variables).

Running the regression model over the 24 h of a day without taking into consideration relationships between hours (notably in terms of consumption habits) may give incoherent results. That is the reason

1. This last figure is calculated over an entire month

2. The three measured samples correspond to so many climatic zones

3. Central air conditioner, room air conditioner, water heater, dishwasher, washing-machine, tumble-dryer, range, freezer and swimming pool pump

why the authors include in their method some simplifications and apply use restrictions. These constitute what the authors call a Seemingly Unrelated Regression system.

Moreover the authors introduce in the mathematical model an indicator to differentiate the power demand according to the house typology : detached house or flat. The functional architecture of Aigner et al.'s model is shown in figure 1.3.

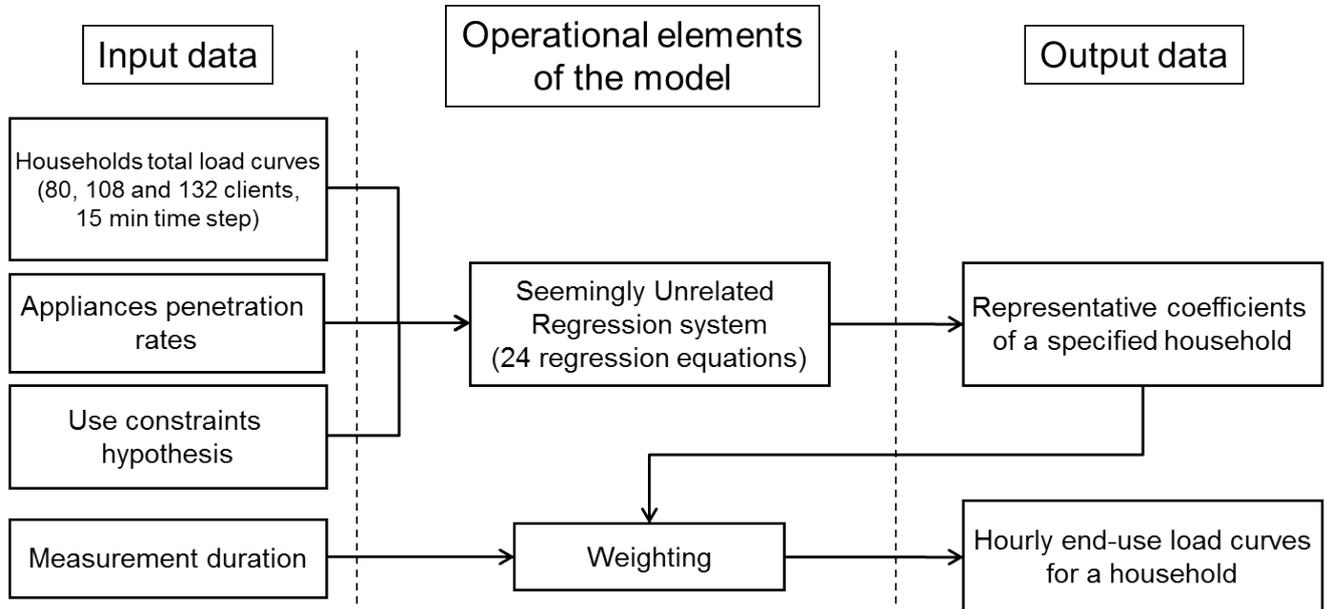


FIGURE 1.3 – Bloc diagram of Aigner et al.'s model [ASK84]

Aigner et al. present in [ASK84] the results obtained when each regression equation is solved separately with three different mathematical techniques :

- ordinary least square method without use restrictions ;
- ordinary least square method with use restrictions ;
- generalized least square method with use restrictions.

The obtained results with the three calculation ways are in the same range. The authors notice that the restriction use is useful in order to improve the real aspect of the daily load curves. Yet there is no comparison in the paper of the obtained end-uses daily load curves with other measurements for each appliance. So the level of the regression results are questionable because of this absence of validation.

To conclude with this model, we can say that Aigner et al. propose a simple way to get hourly load curve for some electric appliances in a sample of households compared to the direct metering of each equipment. This version of the model can be improved if input data concerning the real daily use of the domestic appliances is used. In this way, uses of equipments can be defined more accurately. Nevertheless the model is limited when considering the accuracy of the results. There is no distinction in terms of unitary load profile between the same type of electric device and the use scenarios between the households. Another limit of the model is its disability to supply the load curve of appliances which are in a large majority of dwellings : the refrigerator is the best example. On the other hand, CDA technique presents a good ability to identify the load curve of appliances with an high energy consumption.

### 1.1.2.2 Bartels et al.'s model

Bartels et al. present in [BFGL92] the load curve model called DELMOD they develop on the basis of the CDA technique. In fact, CDA is part of its functional procedure but it's not the only one as figure 1.4 shows.

The aim of the authors was to obtain a tool able to simulate the impact of various scenarios on the regional power demand of the household stock.

DELMOD is made of two modules :

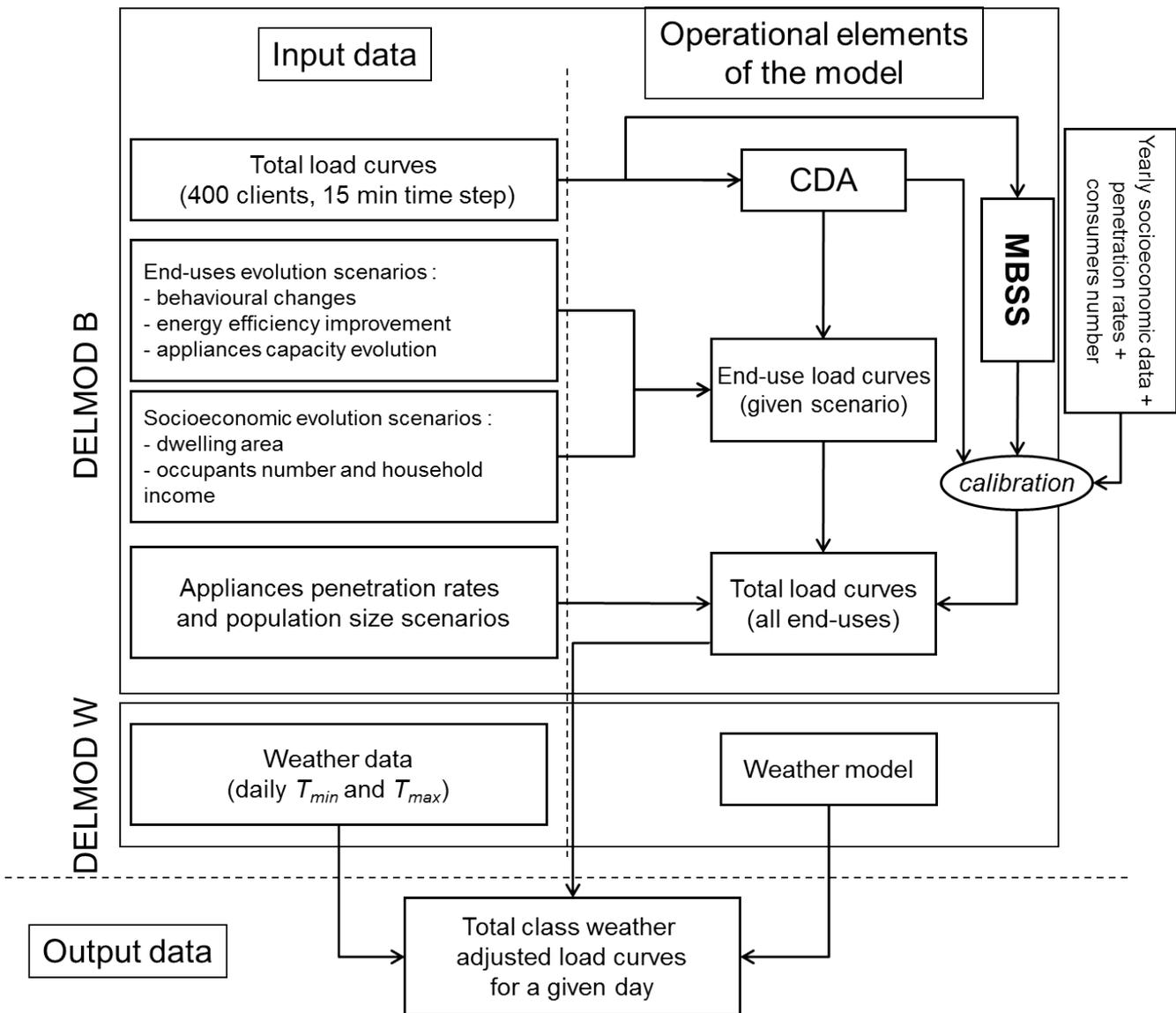


FIGURE 1.4 – Bloc diagram of Bartels et al.'s model [BFGL92]

- DELMOD B (base module) calculates the load curve of the residential sector for a mean working day and a specified month ;
- DELMOD W (weather module) quantifies the climatic influence on the power demand and modifies the results of the base module.

Results of this model are hourly load curves.

As input data, this method requires much more information as Aigner et al.'s model. In fact, on top of measurements of total load curve (400 consumers, 15 min time step) and information concerning the saturation rates of appliances (which are the same type of input data as last model), DELMOD makes use of weather data (outdoor temperature knowledge over the load curve measurement period), socio-economical and technical evolution scenarios (expressed through the growth or the decline of economical, demographical and social indexes). This feature enables the use of DELMOD so as to do demand forecasts. In this model, the authors selected nine different appliances<sup>4</sup>. An other variable was included in the model so as to take into consideration the domestic appliances with a large power demand. The remaining power corresponds to the in-the-model unspecified appliances.

4. Air conditioner, freezer, automatic defrost fridge, oven, dishwasher, tumble-dryer, main and/or secondary heater, water heater (off peak and main tariff) and pool pump

Some coefficients are used in order to take into account the evolution of the technical and/or behavioural aspect which may affect the load curve of each end-use. The capacity and the efficiency of the appliances are the selected factors. Concerning the sociological and economical evolutions three influences are used : dwelling size, number of occupants per household and income.

So in order to take into consideration the meteorological influences, Bartels et al. introduce in the model the minimal and maximal daily outdoor temperature of the considered area (New South Wales, Australia). From these different sources of information various scenarios can be established. All what concerns the non weather influences constitutes evolution trends which are constructed in two modules of DELMOD B. The weather scenarios are built in DELMOD W. There are temperature profiles for the simulation period constructed from the measured chronicle on which deviations around the calculated mean daily temperature are implemented.

Bartels et al. construct the hourly load curves of the selected end-uses with the help of the CDA technique whose results are calibrated using the Model Based Statistical Sampling (MBSS) method – developed in [WRI84] – enabling then to obtain better results than those from the « simple » Conditional Demand Analysis. End-uses load curves are then obtained for two types of day : with or without evening shopping. Bartels et al. simulate the day when the peak load has been reached (the « peak day ») and analyze the results of DELMOD. Total load curves for customers were calculated after aggregation of their different end-uses. Apparently the model enables a good representation of the load curve for this day. Other comparisons were made for a working day of each month of the year. In this case, DELMOD provides not so fine results as for the peak day. As explanation for this observation, the authors say that the considered end-uses are much more weather dependent in cold months.

DELMOD is a load curve model which can be used for power demand forecasts what constitutes a good improvement comparing with the CDA technique. Various evolution sources can be then taken into account. Yet the model suffers the disadvantages of the method developed by Aigner et al. : end-uses owned by a majority of households can not be affected with a corresponding load curve. Moreover although the hourly power demand for the heating system<sup>5</sup> is calculated by the model, there is no differentiation in terms of thermal characterization of the envelope of the dwellings in the simulated area. Finally, the results of DELMOD have been compared on an aggregated basis. That is to say that end-use load curves have not been validated individually and thus do not have any proven validity for industrial use (e.g. impact after retrofitting a dwelling).

## 1.2 Bottom-up models

### 1.2.1 Introduction

Bottom-up models compute electric demand for a few modelled households which are either representative of a larger space scale or are simply target consumers. The unitary results are then extrapolated to obtain the electricity consumption for the entire studied geographical scale. Here the required input data can be :

- the individual consumption of the selected domestic appliances ;
- their technical properties ;
- the geometrical and thermal properties of the modelled dwellings ;
- weather information ;
- electricity bills of households ;
- human behaviour . . .

This methodology presents the advantage of a high level of details because energy demand can be calculated for a couple of households and even for each end-use. With such data, energy demand evolutions may be more precisely quantified since the participation in the global demand of individual domestic appliances is known. Energy demand measurements are not absolutely necessary and thus it is possible to simulate and to obtain results for non-trend scenarios. Among other capabilities, this enables the

---

5. This constitutes a singularity among the presented load curve models

simulation of the impact of technical breaks. Higher precision is correlated with higher complexity of the modelling and with the requirement of a large amount of input data.

In this section we present methodologies whose aim is the synthesis of residential load curve. With the help of accurate data, the authors try to build up the load curve of households when considering a set of selected appliances and with applying a series of assumptions. On top of that, diversity affecting electricity consumption has to be generated by the model. Various methodologies exist to obtain it. To our opinion, some are relatively simple, other are well elaborated and complex. In the following, the models are classified thanks to the authors' adopted scale of complexity for the generation of diversity. The models appear according to a subjective and increasing complexity level of diversity.

### 1.2.2 Statistical random model

Yao and Steemers present in [YS05] the bottom-up model they built up. Called Simple Method of formulating Load Profile (SMLP), this method produces end-use load curves with consideration of various human occupation scenarios. SMLP was established in order to simulate the impact of the integration of decentralized energy production, especially with Photovoltaic (PV) modules, on the British electricity grid. The aim of this tool is notably to predict the peak load curves and to identify the network configurations when the system stability can be at fault. That is why the authors simulate the worst possible case that is to say that all appliances in each dwelling are supposed to run through each simulated day. Results of the method are total load curves at 1, 5, 15 or 30 min time steps. The model architecture is shown in figure 1.5.

Yao and Steemers make use of different input data :

- individual DHW profiles ;
- per capita daily electricity consumption for each end-use ;
- heating system energy consumption from simulation results of Yao et al.'s thermal model [BY02, YBM02] ;
- statistical results from national population surveys ;
- appliance saturation rates for a series of domestic electric devices ;
- constructed occupation scenarios in the dwellings.

The (15) appliances taken into consideration are listed below : TV, VCR, fridge, freezer, fridge-freezer, hob, oven, kettle, micro-wave oven, dishwasher, washing-machine, tumble-dryer, vacuum cleaner, iron and lighting.

In order to generate the total load curve (all end-uses considered) of a selected area and according to the parametrized simulation scenario, Yao and Steemers's model proceeds in three steps :

1. at the household level, generation of one daily load curve for a selected appliance and occupants' presence scenario : the start time is selected randomly with the help of a Random Number Generator Technique (RNGT) ;
2. start time diversity is obtained when repeating previous step 20 times ;
3. aggregation for the specified presence scenario of the daily load curves obtained at step 2 : the resulting load curve is naturally smoother than the individual ones ;
4. repetition of the steps 1-3 for each appliance according to the considered presence scenario ;
5. repetition of the steps 1 to 4 for each presence scenario considered ;
6. aggregation of previous total daily load curves according to the characteristics of the population area (composition of households, type of dwellings).

So as to validate their model, Yao and Steemers used a measured load curve coming from the data base of the UK Electricity Associate Load Research Team. The chosen load curve is a three-person household's, measured during a winter working day. The SMLP simulation to obtain a comparable load curve consisted of simulating a 100 household community, to calculate its load curve and to divide the result by 100 in order to get the load curve of the « mean » household.

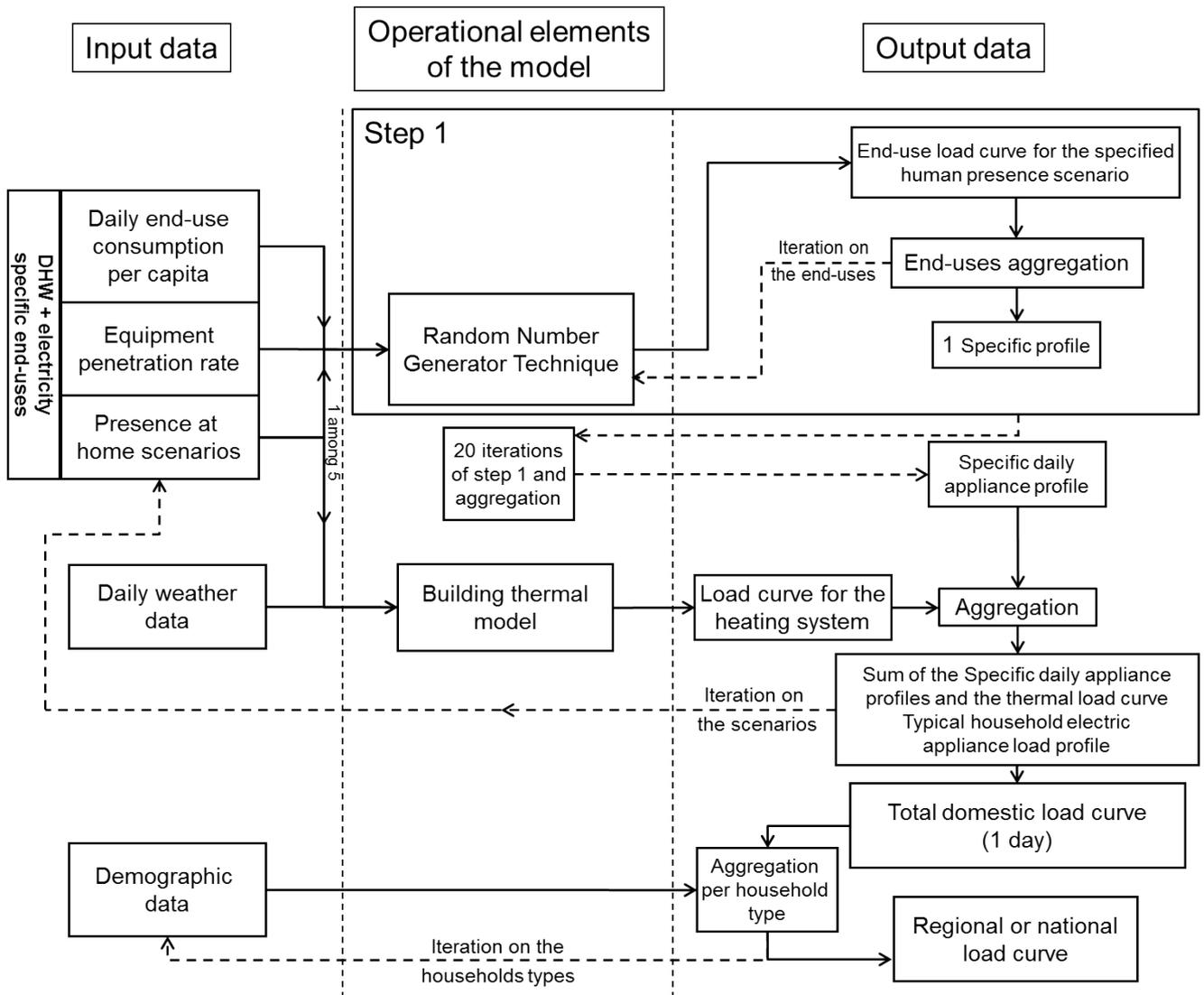


FIGURE 1.5 – Bloc diagram of Yao and Steemers's model [YS05]

The used method to give a validity to the model is very questionable because of the chosen compared things : individual and sample mean load curve are two very different elements.

To conclude, this model presents the advantage of the simplicity, the availability of the input data and its deeply bottom-up character (several aggregation steps are needed). Yet the occupation scenarios that are proposed by the authors seems to be very simple to represent the reality with a good fit. Moreover the introduction of the diversity is done with a pure mathematical procedure without considering for instance domestic habits. On top of that, the thermal model which is used corresponds to a one air-node building modelling which allows a limited daily regulation of the heating system in a home. There is no differentiation and/or no indication in terms of building envelope characteristics and appliance efficiency. That is the reason why technical evolution scenarios can not be simulated with this model.

### 1.2.3 Probabilistic empirical models

#### 1.2.3.1 Stokes's model

Stokes developed a model applicable to the domestic sector and presented in [STO05]. It is able to generate three types of aggregated load curves :

1. 30 min demand profile for an average household ( $1^{st}$  level),

2. 30 min demand profile for a specified dwelling (**2<sup>nd</sup> level**) and
3. 1 min demand profile for a selected domestic customer (**3<sup>rd</sup> level**)

with a fine spatial resolution that is the individual domestic appliance.

To get this gradual precision, the model is structured into three main modules; results of the level  $n$  serve as inputs of the level  $n + 1$ . Figure 1.6 schematically illustrates this architecture.

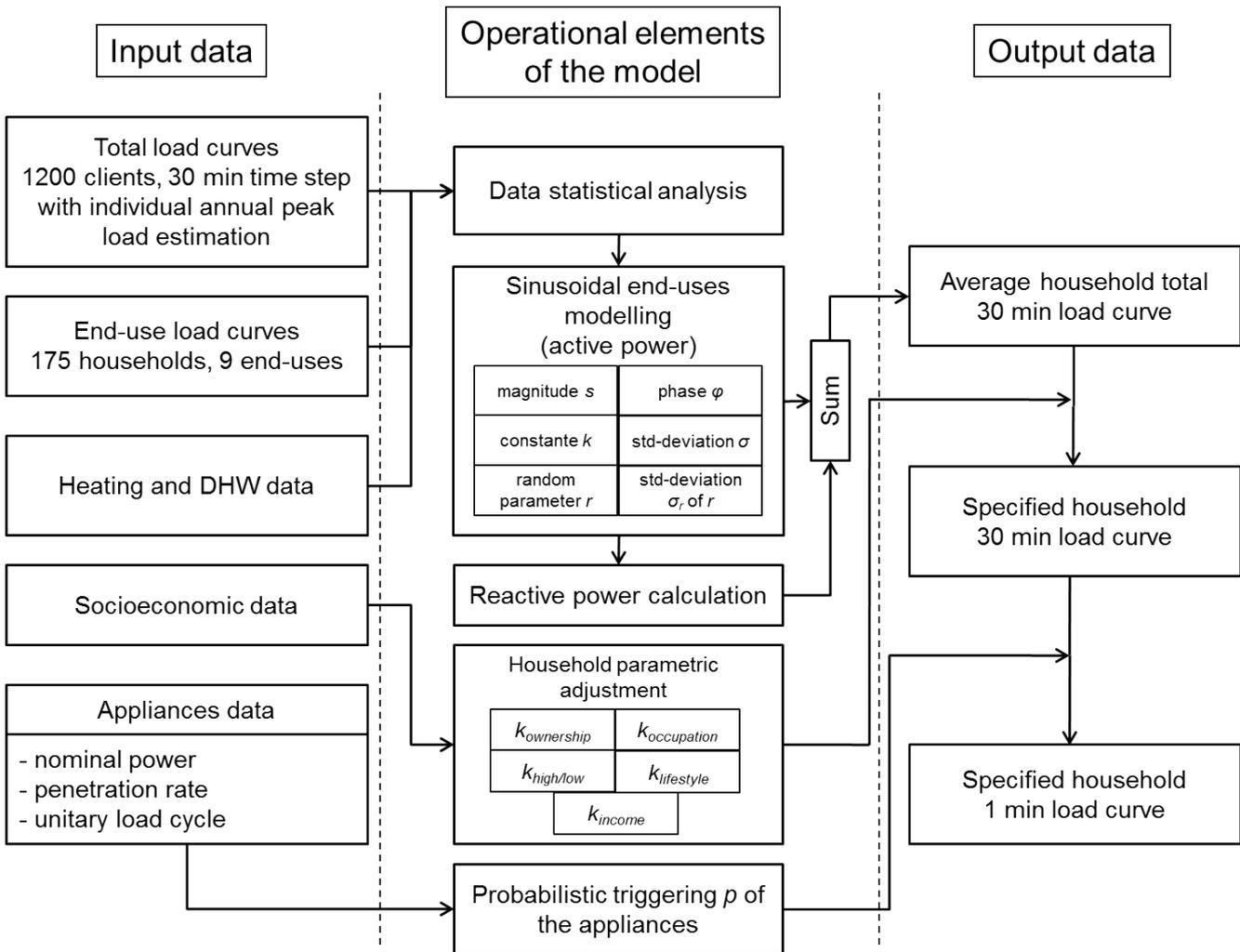


FIGURE 1.6 – Bloc diagram of Stokes's model [STO05]

Because of the lack of 1 min measured load curves which can be used as (calibration) input data and in order to avoid modelling precision discrepancies between appliances, Stokes built up a unique, clear and homogeneous modelling framework that is applied to all end-uses. This way, the level of details of the modelling is the same whatever the domestic device. On top of that, Stokes's model calculates the load factor of each equipment at each time step enabling then to take into account the reactive power demand. According to the author, this part of the power is often neglected in load curve models despite its interest as for instance for load flow analysis purposes.

In the following, it will be considered the input data for each level, each computing phase with focus on the assumed simplifications, the results of each level and finally the validation of them.

### Input data

In order to calculate the 30 min demand profile for an average household, the 1<sup>st</sup> level of Stokes's model uses 30 min load curves that have been measured in a sample of 1200 English households. Moreover data at the device level (individual load curves for nine appliances) and knowledge about production

and stocking of DHW : this set is differentiated in terms of day-type (three modalities) and tariff (two options). On top of that, information about socio-economic characteristics of the households on the one hand and concerning the dwelling on the other, was available through answers of questionnaires. With regard to the national average, different discrepancies have been identified in the sample by the author. On top of 30 min demand profile for an average household, the 2<sup>nd</sup> level of the model, whose aim is the calculation of 30 min demand profile for a specified dwelling, receives information about the penetration rate for each appliance, the occupation of the households (number of occupants according to the dwelling area and presence at home scenarios) and an estimation of its annual maximal power demand value. The type of dwelling, its surface area, its orientation. . . are other diversity elements that have been considered out of Stokes's work because of the lack of so precise data.

So as to step down at the 1 min time resolution and to provide corresponding load curves for a specified household (load curve that is derived from the 30 min load curve calculated by the 2<sup>nd</sup> level), the 3<sup>rd</sup> level gets information concerning the habits of appliance use (e.g. duration and frequency of the cycles), measured unitary load cycles and estimations of maximal power demand per appliance. According to these data, four parameters are defined to characterize the unitary load cycles at the 1 min time resolution : scale of demand, event duration, use frequency and timing.

### Modelling details

Stokes's purpose is the modelling of the domestic load curve for average British dwellings. Her model has to be able to provide demand profiles for contemporary situation as future one : so load curves must be adjustable. That is the reason why she decided to assume following statement :

« Variations that affect end-use peak demand and its annual consumption are identical in the time ».

Concretely, reduced (normalized) power demand,  $P_{r,i}$ , for each end-use has been calculated thanks to the Load Research Group<sup>6</sup> (LRG) data set according to the equation 1.1

$$P_{model,i} = \frac{P_{mes,i}}{P_{max_a,i}} = P_{r,i} \quad (1.1)$$

where  $i$  is a specified end-use,  $P_{model,i}$  is the end-use power data used in the model,  $P_{mes,i}$  is the measured 30 min power demand for end-use  $i$  and  $P_{max_a,i}$  its annual maximal power demand.

From the LRG data set, Stokes tried to build up a mathematical approximation for 30 min power demand fluctuations on the year. According to her research, the best compromise between accuracy and simplicity is the sinusoidal modelling of the 30 min normalized power demand –  $P_{sinusoidal}$  – that is reported in equation 1.2.

$$P_{sinusoidal} = s \cdot \sin \left[ \left( 2\pi \cdot \frac{N_d}{N_y} \right) - \Phi \right] + k + r \quad (1.2)$$

$s$  represents the sinus function amplitude,  $\Phi$  its phase,  $N_d$  corresponds to the considered day's number in a year with  $N_y$  days.  $k$  is a constant and  $r$  is a random number generated with an end-use specific Laplace–Gauss density function  $[0, \sigma_{sinusoidal}]$ .

All parameters derive from the LRG data set. They intrinsically include weather influences : that is the reason why Stokes's model does not make use of « traditional » weather file as input data.

Moreover because of the independence between the Laplace–Gauss density functions, linked end-uses are not considered in Stokes's model.

For certain end-uses (e.g. lighting), minimal and maximal power reduced demand levels have been defined and are applied when equation 1.2 is not relevant or gives negative values. In these cases, 30 min reduced power demand  $P_r$  is calculated according to equation 1.3.

$$P_r = \begin{cases} P_{r_{min}} + r_{min} & \text{if } P_r \leq P_{r_{min}} \\ P_{r_{max}} + r_{max} & \text{if } P_r \geq P_{r_{max}} \end{cases} \quad (1.3)$$

6. See [STO05] p. 6

$r_{min}$  (respectively  $r_{max}$ ) is a random number generated with a  $[0, \sigma_{min}]$  (respectively  $[0, \sigma_{max}]$ ) Laplace–Gauss distribution.

Long term consumption trends have been provided by the Market transformation Program [DEF01] for each end-use  $i$ . From these, trend estimations of end-use maximal power demand  $P_{trend,max,i}$  have been calculated enabling then the scaling up of the normalized demand ( $P_{r,i}$ ) generated by equations 1.2 and 1.3. Thus households' group average 30 min (non reduced) power demand –  $P_{group,i}$  – is calculated with the help of relation 1.4.

$$P_{group,i} = P_{r,i} \cdot P_{trend,max,i} \quad (1.4)$$

In her dissertation [STO05], Stokes adapted the previous modelling approach for various individual end-uses listed in table 1.1. Further details about their modelling and the taken assumptions can be directly read in the referenced document.

End-use family	End-uses considered
domestic cold	fridge
	freezer
	combined fridge/freezer
heating	storage heater
domestic hot water	immersion heater
cooking	hot plate
	oven
	kettle
	micro-wave oven
wet appliances	washing-machine
	tumble-dryer
	combined washing-machine/tumble-dryer
	dish-washer
lighting	« standard » bulb
miscellaneous	other appliances

TABLEAU 1.1 – End-use families and appliances considered in Stokes's model [STO05]

The results of this level have been validated on measures conducted by the LRG. Comparisons in terms of 30 min aggregated consumption (all end-uses considered) showed the good ability of the model in terms of consumed energy.

The second level of Stokes's model introduces some influences responsible for diversity in order to produce the 30 min power demand load curve for a specified household.

Income – in relation to the English mean salary – and lifestyle are two elements that influence the use of electricity because household's type of dwelling and corresponding room number are correlated to them : that is why they were integrated into the model. However they are only applied for lighting and miscellaneous appliances on the one hand, they influence the equipment set of the households on the other.

According to the dwelling surface area, that is linked to the number of rooms, the household size (number of members) is calculated. Then household's consumption for each end-use is adjusted according to its size notably with the help of results from Boardman et al. study [BFMH+94].

However the presence at home daily profiles and household's occupants' age are two variation sources that are not taken into account in Stokes's model. Nevertheless correlations between lifestyles, penetration rates and daily appliance use were identified in [MNP96] and applied when possible.

In order to calculate 30 min (Half Hourly) load curve for a kind of household (not a precise one) and an end-use  $i$ ,  $P_{HH,spe,i}$ , the second level of Stokes's model declines equation 1.5

$$P_{HH,spe,i} = P_{HH,group,i} \cdot k_{ownership,i} \cdot k_{occupation,i} \quad (1.5)$$

where  $P_{HH,group,i}$  is the calculated half hourly power demand for the end-use  $i$  into a group of customers,  $k_{ownership,i}$  is a boolean factor indicating the presence (or absence) of the end-use  $i$  in the considered household and  $k_{occupation,i}$  is a scalar factor that modifies the value of the calculated power demand according to the size of the regarded household.

For some end-uses  $i$ , another coefficient ( $k_{lifestyle,i}$ ) multiplies previous equation. It is aimed to take into account the lifestyle (more or less the social status) of the simulated household. For the domestic hot water,  $k_{lifestyle}$  is replaced by a coefficient (four modes : very low, low, average and high) whose mode is randomly selected that adjusts the water usage when simulating a sample of households.

The output of the second level of Stokes's model is calculated when adding all the previously determined 30 min load curves for each end-use. However this level does not provide results for a specified household because all dwellings with the same characteristics (household size, lifestyle, domestic equipment set. . .) will be affected with the same load curves. What introduces variations between them is the third level of the model. It takes the previous calculated power demand profiles as input and considers them as probability distributions to start the end-uses within a day.

In the third level of her model, Stokes tries to explain in detail the diversity of the domestic energy consumption in order to provide the calculation of the 1 min domestic load curve for a specified household. According to Stokes, power demand can not be modelled without consideration of the chance influence. That is the reason why the end-use launching is determined by comparison between a generated  $[0; 1]$  random number and a constructed ratio  $p$  calculated as follows :

$$p(\text{event occurring}) = \frac{P_{HH, spe, i}(hh, d)}{P_{HH, app, i}} \quad (1.6)$$

where  $p(\text{event occurring})$  is the probability that an event (the launching of an end-use) occurs,  $P_{HH, spe, i}(hh, d)$  is the half hourly power demand value for the end-use  $i$  calculated for a specific household at the day  $d$  and the half-hour  $hh$  and  $P_{HH, app, i}$  is the power demand required for a « standard » appliance  $app$  that serves the end-use  $i$ .

The more the value of the power demand calculated at the second level is high, the more the probability  $p$  to start at least one appliance is important. On top of that, the model is able to start several equipments that are related to the same end-use (e.g. lighting) on the one hand, it takes into account the different power demand levels when relevant on the other hand (e.g. hot plates).

Concerning the time start selection of the end-uses, they are chosen with considering individually each half hour. Apart from equipments whose functioning cycle(s) is (are) longer than 30 min, the appliances start times are chosen so that the end-use required energy demand is satisfied into the current half-hour. The end-use starting frequency and the duration of each cycle are two elements that make the model adaptable. Other assumptions have been made by Stokes for each kind of end-use.

To validate her model, the author makes use of two data sets :

- measures on a bungalow district at the 5 min resolution ;
- measures on a sample of 13 households at the 1 min time step.

She considers different comparison criteria : energy demand, maximal and mean power demand value (power factor), daily profiles and load distribution density function.

On the first sample, Stokes's model succeeds in estimating the energy demand on the simulation period. However, it seems to be the result of various overestimations and underestimates depending on the considered end-use. In terms of power demand, generated daily profiles are relevant but measured maximal power demand value is always greater than the simulated one. According to the author, it seems to be due to the too low generated diversity.

On the second sample, the same conclusions can be said but on top of that the model gives a general underestimate of the energy consumption and is quite unable to predict the morning and evening peak loads of the domestic customers.

To conclude, Stokes established a very detailed model split into three levels. She adopted an homogeneous modelling framework that makes easier its reuse and modification. Differentiation between end-uses and

appliances is carefully taken into consideration. Moreover, the author tried to implement in her model a large set of power demand determinants.

However as we said in the validation paragraph, the generation of diversity reaches its limits because of the use of random coefficients that do not really represent the existing end-use relations. On top of that, the 1 min load curve result for a specified household is partly dependent on the group of measures used for the calculation (because of the load curve « break down » process). Moreover Stokes's model does not explicitly take into account the daily occupation scenarios of the occupants for the end-uses starting algorithm (it is supposed to be integrated in the input data). Finally, the heating power demand and associated energy consumption is only dependent on the dwelling size area : yet the building characteristics, especially the insulation level of the house, plays a determinant role for the thermal appliances.

### 1.2.3.2 Paatero and Lund's model

In [PL06], Paatero and Lund explain the model they have set up. Its aim is to provide 1 h load curve for domestic customers (from few dwellings to a relative large set – 1000 clients). Its architecture is represented in figure 1.7.

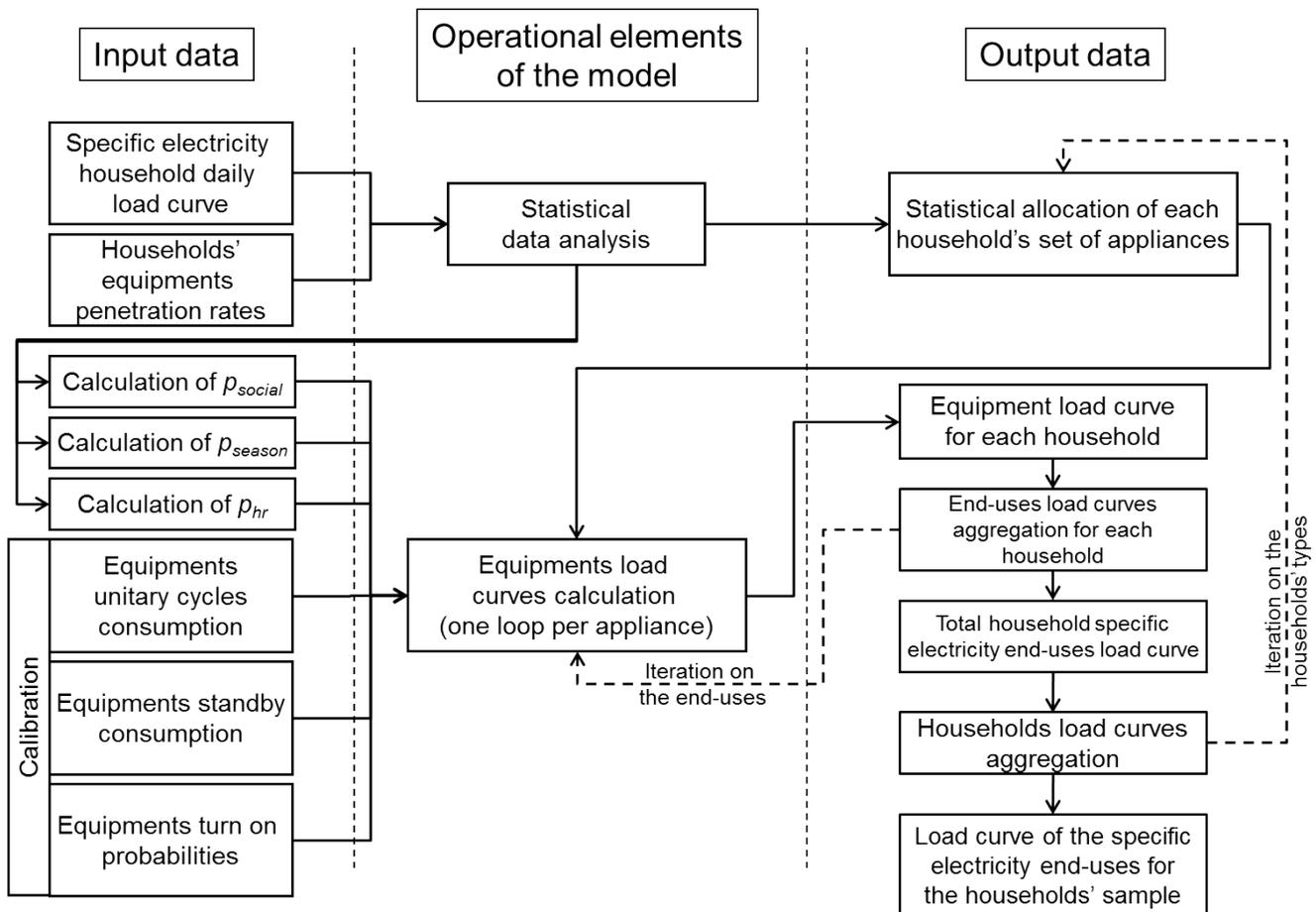


FIGURE 1.7 – Bloc diagram of Paatero and Lund's model [PL06]

The authors made use of two main data sets as inputs. In fact, there are 1 h power demand measurements that were led on Finnish domestic customers living in flats (respectively during 1 year on 702 households and during 143 days on 1082 dwellings). Because of the absence in the measured apartments of electric heaters, coolers and immersion heaters for domestic hot water supply, these end-uses are not simulated by the model. On top of that, Paatero and Lund used various studies to construct a data base of unitary load cycles and standby power consumption for the simulated appliances. Moreover, they gather national statistics in order to build up the domestic equipment sets.

In order to develop the model, Paatero and Lund studied a series of load curve measures. This way they noticed that power demand fluctuates according to different periods especially the season, the day and the hour. Exogenous influences like outdoor temperature and daylight availability can be theoretically responsible for them. However in this dataset, these influences can not explain such fluctuations because of the absence of thermal appliances. Domestic habits and social behaviour are responsible for this variability but they did not have been taken into account because it constitutes a too high level of details according to the authors. Moreover Paatero and Lund identify as well as Capasso et al. the deeply interactions between the presence at home of the households' members and conducting domestic activities.

Paatero and Lund constructed a two-level model : the first level defines both the domestic equipments set for each household and the general load fluctuation trend. The second level simulates each appliance in each household with the help of end-use starting probabilities.

To generate diversity, the first level calculates the density function of the domestic daily consumption with the help of sample data. The authors have reduced the previous distribution in order to get a dimensionless curve that is centered on the daily mean consumption. Then for each simulated day, a value is selected according to the distribution : it is called  $p_{social}$  (social random factor) because it is supposed to capture the social diversity of the demand and it is applied identically for each simulated household. The equipment set in each dwelling is constructed according to the ownership statistics.

The second level of the model provides the hourly load curve for each equipment with considering each household separately. To compute this, two elements are required : the unitary load cycle of each appliance and its starting probability  $p_{start}$  calculated according to equation 1.7.

$$p_{start}(i, we, \Delta t, \sigma_{ap}, hr, d) = p_{season}(i, we) \cdot p_{hr}(i, hr, d) \cdot f(i, d) \cdot p_{step}(\Delta t) \cdot p_{social}(\sigma_{ap}) \quad (1.7)$$

with :

- $i$  is an equipment or a group of equipments ;
- $hr$  is the hour of the day ;
- $d$  is the day of the week ;
- $we$  is the week of the year ;
- $p_{social}$  is the social random factor whose standard deviation is  $\sigma_{ap}$  ( $ap$  is for apartment) ;
- $p_{season}$  is the seasonal factor (models the weather influences) ;
- $p_{hr}$  is the hourly factor (that takes into account the changes in terms of activity levels during the day) ;
- $f$  is the mean daily starting frequency for each end-use  $i$  ;
- $\Delta t$  is the working time step and
- $p_{step}$  is a scaling factor that adjusts the probability values according to  $\Delta t$ .

According to the sign of the comparison between  $p_{start}$  and a  $[0; 1]$  generated random number, the appliance  $i$  is started or not. If the equipment is effectively launched, its load cycle is added to the currently calculated household total load curve. Apart from « cumulative » appliances (such as light bulbs) the calculation of  $p_{start}$  is stopped for the equipments during their functioning. As soon as the end of the cycle is reached, the appliance returns then in off-mode and the starting probability can be evaluated again.

Concerning the standby consumption, it is taken into consideration with adding a constant load at each time step.

As a exercise, Paatero and Lund tested their model to calculate the yearly energy consumption of a 10 000 households sample living in flats. It is not a validation of the load curve reconstitution method because no comparison is carried out on the generated power demand profiles. They only conclude that their model underestimates the mean daily energy consumption and that the distribution of the calculated daily energy use is more spread as the measured data.

Paatero and Lund developed a bottom-up model that is able to provide load curve that are used for instance to quantify the impact of DSM measures. The strength of this method is that it relies on public

data and its computational process is relatively simple.

Yet the model is limited in different ways. First of all, because of the used input data, the method is only able to calculate load curve for households living in flats : the data sample is highly non representative of the entire housing stock. Then the model can not really step back from the past reality because the random influences are constructed on the input data set. On top of that, the model is not capable of evaluating the power demand impact that follows the introduction of new end-uses, appliances or behaviour.

#### 1.2.4 Time of use based models

Time of use data – whose concept was originally developed by Pratt et al. in [PCR<sup>+</sup>89] – generally correspond to daily frequency curves of use for each end-use/domestic activity. In order to construct this kind of information two main ways are available :

- a monitoring campaign is realized in a sample of households and during the study time each household's member has to write down in diaries the conducted domestic activities at the time step resolution that is chosen for the campaign ;
- the other way is to focus the study on different general domestic activities (breakfast, lunch and dinner time, get up and bedtime time, daily transport duration. . .) and to reconstitute an incomplete diary with mean times of each activity with a density function around these moments.

Then the compilation of the results enables one to obtain frequency curves for the studied activities by comparing for each time step the number of households where a specified activity is being conducted with the total number of studied households. The segmentation of the results can be done for each end-use with the same methodology if the study has been conducted on the domestic equipment basis. Moreover it is possible to get curves for different types of day (it is often divided between weekday and weekend day). This concept enables the modelling of domestic activities and it provides realistic sequences of electric equipments use in a sample of households.

##### 1.2.4.1 Walker and Pokoski's model

Walker and Pokoski [WP85] are the first authors which use the concept of time of use with the aim of constructing a load curve model. More precisely they are the first who take into account human behaviour to reconstruct load curves for a specified household or a set of different dwellings. The model was developed so as to help the planning of new power plants and it gives total load curves at 15 min time step.

The basis of the human behaviour modelling is the use of two kinds of probability functions respectively called availability and proclivity which will be then generalized in the model of Capasso et al. (cf. section 1.2.4.2). The first one gives the household's occupants probability to be at home and so « available » to use an electric domestic appliance. Concretely it is implemented in the model with  $96 \times 1$  daily vectors (the time resolution is 15 min –  $96 \times 15 = 1440$  min = 24 h) containing 0 and 1 indicating respectively on the one hand the absence or the asleep state of a person and, on the other hand, the presence at home of a specified household member in an awake state. Theoretically in order to represent the exact reality, the input data required for this function is one occupation diary per studied person. Consequently these data-gathering can be very expensive. That is the reason why Walker and Pokoski only make use of one presence scenario and they improve it with variations sources thanks to different density functions implemented at the start times of the main activities. In this way they get various « mathematically constructed » occupation daily scenarios. Yet the authors had to take into account the different ways people live : typically the time to go to work and to come back from work generally in the evening. Delay probabilities were defined in accordance with the estimated proportion of each « way of life » in the simulated area. All things considered, the availability function is constructed with the help of 13 sub-functions, each one gives a unitary kind of information. Moreover Walker and Pokoski take into consideration the absence from home because of non professional reasons (hobbies, shopping. . .) : they model this influence with the help of  $P_{away}$ , a probability function which makes dependent on it the availability function.  $P_{away}$  is

compared with a random number generated with a  $[0; 1]$  uniform density function. Two ways are then adopted in order to calculate the number of occupants at home for each time step according to whether  $P_{away}$  is bigger or smaller than the generated random number.

The availability function gives not enough information in order to reconstruct the load curve of a set of households. That is the reason why the authors make use of various proclivity functions which model domestic habits and more precisely time periods when they are likely to be conducted<sup>7</sup>. The proclivity for a certain activity is equivalent to the tendency to do it during a day. These preferential time periods come from standards of living, conventional ways of life or even people habits. In fact there are two kinds of proclivity functions :

1. the first one indicates the likelihood for the use of a specific individual electric appliance ;
2. the second type considers a domestic activity which may require the use of various equipments.

In order to construct the first proclivity function (the so called basic proclivity function), Walker and Pokoski make use of different results coming from various surveys whose aim was the construction of a daily diary for a selection of domestic appliances so as to know the periods when they are likely to be used and the weekly frequency of turn on events.

The second type of proclivity functions represents more general activities such as the daily mealtimes. Potentially this activity involves the use of a few domestic equipments such as the fridge, the range and/or the water heater because of the very possible draw of water for this activity. With the help of results from various studies, Walker and Pokoski define density probability functions for each mealtime of the day. This way each period in the day is affected with a probability that meal can be taken during this time period. Such probabilities can be defined for other activities.

With the help of the availability and the proclivity functions, the authors have the knowledge concerning the moments when the electric appliances are turned on during the simulation time. To get the corresponding load curve for each end-use they have to « stick » on the identified on/off events, unitary load curve patterns that are specific for each appliance. In the case of the appliance which run without the need of a human intervention (for instance, they are turned on or off according to the information which is provided by sensors), Walker and Pokoski use what they call « normal-cycle » functions. It's notably the case of the fridge which runs all the time with a cycling on/off alternation.

The architecture of the model is presented in figure 1.8.

Figure 1.8 shows the chosen aggregation method. In fact, the simulation tool provides individual end-use load curves (one curve per electric appliance). These raw results are then aggregated according to two steps :

1. first a summation is done at the household level achieving then the total load curve for a specified dwelling ;
2. secondly the total load curves are aggregated in order to get the global load curve of the simulated area.

This modular architecture enables to consider the influence of the power demand of an individual appliance with regard to the total demand of the household. Another advantage of this technique is the possibility to add a new appliance in the model without reconstructing its whole structure and with conserving the calculation method. Moreover the simulation provides the load curve of an area which is not calculated as an individual household's mean load curve multiplied by the number of simulated dwellings. Thus real behaviour with potentially huge differences between them can be simulated. Furthermore it gives the opportunity to simulate a lot of dwellings with different sets of domestic equipments.

In [WP85] only the water heater is presented in details but the entire model is described in [WAL82] and at the date of the publication the model did not contain thermal load appliance (such as heater). However Walker and Pokoski did not consider this feature much more complicated than the water heater because, according to the authors, it only requires another function which takes into account, on top of the availability and proclivity information, the in- and out- heat flows affecting the dwelling and economic considerations concerning the occupants.

---

7. Without excluding their respective probability at any other time in a day

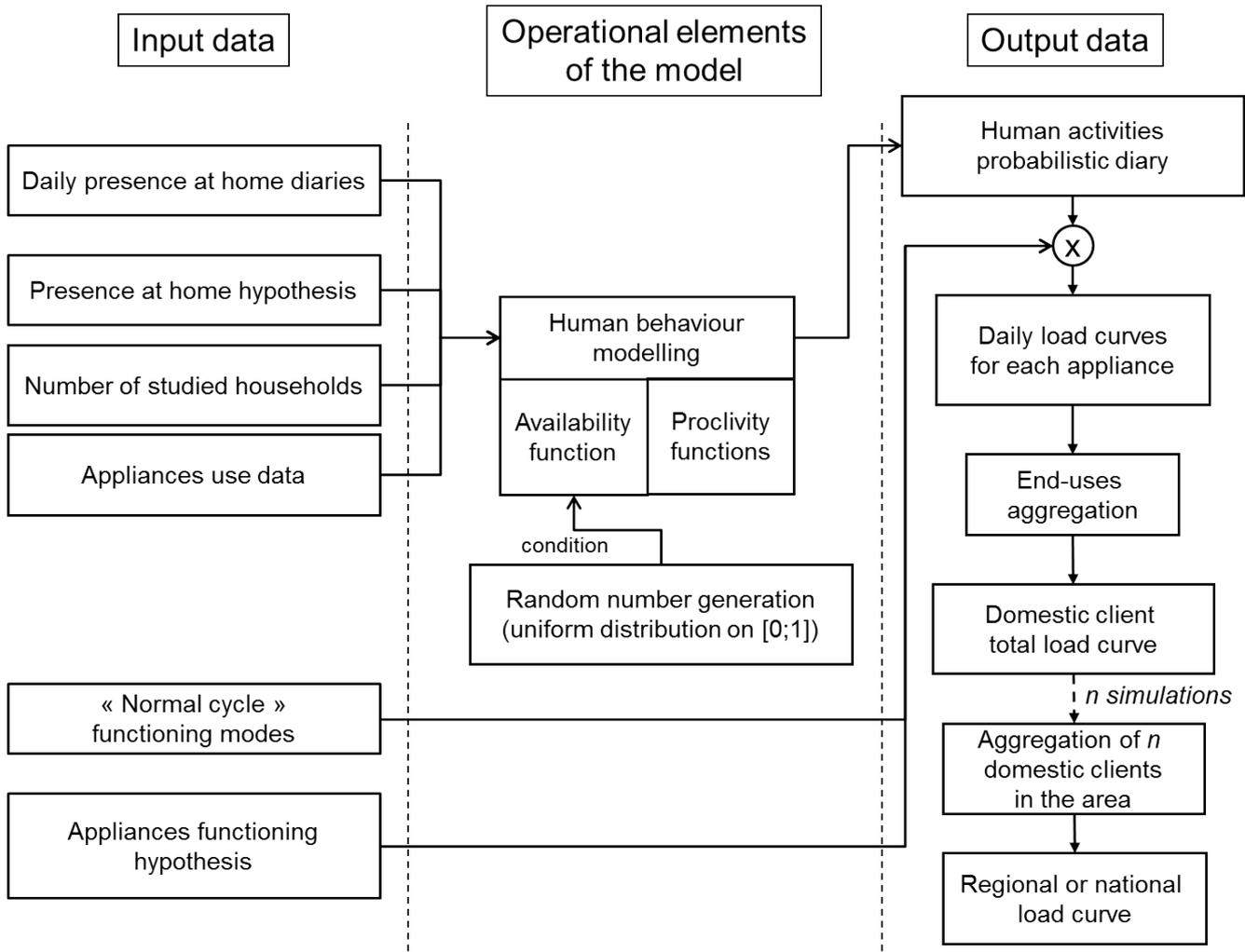


FIGURE 1.8 – Bloc diagram of Walker and Pokoski's model [WP85]

To validate their model the authors carried out two simulations. The first one had the aim to obtain the daily load curve for two single-family homes. The second test concerned two groups of residences. Simulation results were compared with data from a load curve measurement campaign. The chosen criteria for the comparison was the **Normalized Variation Factor (NVF)**\* defined as follows :

$$NVF = \frac{\frac{1}{n} \cdot \sum_{t=1}^n (P_{est}(t) - P_{mes}(t))^2}{\overline{P_{mes}}^2} = \frac{\frac{1}{n} \cdot \sum_{t=1}^n (P_{est}(t) - P_{mes}(t))^2}{\left(\frac{1}{n} \cdot \sum_{t=1}^n P_{mes}(t)\right)^2} \quad (1.8)$$

with  $P_{est}(t)$  the estimated power demand at time  $t$ ,  $P_{mes}(t)$  the measured power demand at the same time and  $n$  the number of simulation time steps. Conceptually the *NVF* corresponds to a « pseudo-variance » of the estimated power demand calculated with reference to the measured load curve that is normalized by the square of the mean measured power demand value. So *NVF* represents the dispersion of the estimated power values compared to the measured ones. *NVF* is reused by Capasso et al. [CGLP94] and Widén et al. [WLV<sup>+</sup>09].

With this indicator, the authors tried to know if their results are included in the fluctuation range existing between two daily measured load curves. In fact, it is possible to calculate the *NVF* for the same household (or group of households) between different days or for the same day and between different households. They got good *NVF* values (i. e. small figures) confirming the fact that the simulated load

curves fit the measured ones.

#### 1.2.4.2 Capasso et al.'s model

Capasso et al. present in [CGLP94] the model they developed which is called ARGOS. Its aim is to reconstruct the load curve for a specified household up to a set of dwellings, in a defined area, in order to quantify the impacts of DSM measures. These are taken namely so as to provide a solution for one of the major problems affecting the Italian electricity network : the winter peak load occurring during worked days. Another purpose of this methodology is to be a tool for optimizing the production of the power plants or even to « drive » the planning of new energy supply sites. The main characteristic of this approach is the fact that it takes into account the major interactions between the electricity energy demand and the behavioural factors. The results are load curves with a 15 min time resolution.

From results of various surveys, the authors constructed some probability functions. They notably make use of the Walker and Pokoski's availability and proclivity functions enabling them to obtain mean profiles for the basic proclivity function.

ARGOS simulates the load curve of a residential customer. This is built up with a set of domestic appliances and a specified number of occupants. According to Capasso et al. the real issue for calculating domestic load curves is the modelling of the interactions between the set of equipments and the household's members.

So as to set a realistic modelling frame, rules and constraints are programmed in ARGOS under a probabilistic form : there are behavioural and engineering functions.

On the one hand, the first ones define the appliances set for each household according to its characteristics. On the other hand, they establish the specified household corresponding time structure for the use of the equipments. That is to say that the domestic activities, which are grouped together in four main families (cooking, housework, leisure and personal hygiene), are not randomly conducted by the occupants through a day, the starting of the equipments too. On the contrary, they are both spread in the simulated day thanks to time of use density functions. These are calibrated so that simulations lead to appliances daily consumption that are included within a confidence interval around values coming from measurements. Simultaneously each household member is characterized with a presence at home scenario and a specific proclivity for each domestic activity.

The engineering functions give various information : first they precisely characterize the functioning of each domestic equipment according to its properties, secondly they set each household subscribed maximal electric power and finally they set the technological penetration in the simulated population.

The previous described data and probability functions (behavioural and engineering ones) are stored in a file which actually contains all required inputs of the model.

From the parameters that describe the scenario, a Monte Carlo extraction is carried out to choose the simulated type of household. The corresponding total load curve is calculated in ARGOS by addition of each end-use specific load curve.

The total load curve for a complete area is simply calculated by addition of the individual consumer's load curves. The architecture of ARGOS is shown in figure 1.9.

As model validation task, Capasso et al. tried to reconstruct the aggregated load curve of 180 apartments which are representative of the Italian dwellings located in the main cities' suburbs. Among them, only 95 households have been kept as comparison references because they answered a questionnaire whose results built up the ARGOS input data file.

As we previously said, this model was established especially to predict the residential load curve during peak load days. That is the reason why, the authors focused the comparison results on restricted time periods. For these critical hours, modelled and measured load curves are very similar. For other infra-day periods, comparisons are less favourable.

In a second step, the comparison period was extended to a complete simulated day. For this task, Capasso et al. estimated and compared the *NVF* between measured and modelled load curves on the one hand and between two consecutive daily measured load curves on the other. Results satisfied the confidence

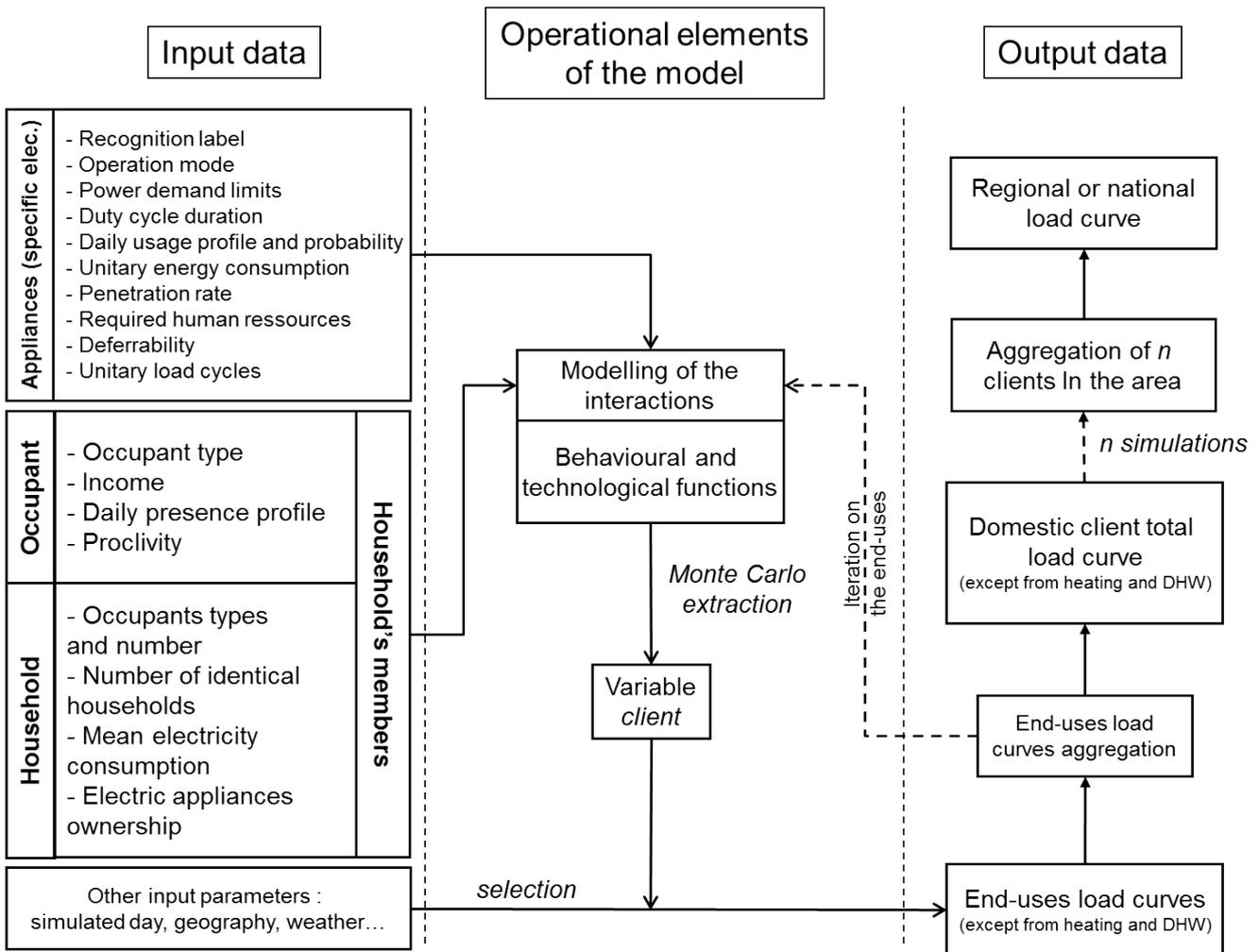


FIGURE 1.9 – Bloc diagram of Capasso et al.'s model [CGLP94]

interval criterion indicating the good approximation ability of the model.

ARGOS constitutes a noticeable improvement of the Walker and Pokoski's model and in the same time a comparison reference for the models developed after it, especially for the human behaviour modelling. Finally, this is notably well reproduced because of the use of real elements instead of standardized occupation profiles as in Yao and Steemers's model. Another advantage of ARGOS is its ability to provide the load curve results under various aggregation levels enabling then the calculation at different scales of load curve specific coefficients. Moreover, new appliances can be implemented in the model, rules and constraints are changed and adapted by the user.

Yet, these possibilities require a lot of input data that are difficult to get at a large geographic scale (i. e. the simulation of the national domestic load curve seems to be illusory). On top of that, the needed information for the elaboration of the behavioural functions have been systematically adapted into Laplace–Gauss distribution. This can lead to non optimal adjustments in all cases. Some significant distinctions in setting the parameters of the simulation have not been taken into account : worked day/not worked day, winter/summer... Nothing about building diversity (for instance thermal characteristics) has been implemented in ARGOS because of the relative insignificance of electric heating in Italy. Finally, the influence of the electricity tariffs is not supported by this model.

After the first model implementation in FORTRAN, research has been resumed and continued by Prudenzi, Falvo and Silvestri. This work leads to a new software called DELOS (see [PSR09, PSL+09]). However no significant evolutions have been added in this model compared with the initial one (AR-

GOS) : in particular the concepts remain the same.

### 1.2.4.3 Armstrong et al.'s model

In the framework of Annex 42 of IEA (International Energy Agency) « The Simulation of Building-Integrated Fuel Cell and Other Cogeneration Systems (COGEN-SIM) », Armstrong et al. [ASR+09] established a model which reconstructs the load curve of domestic specific electricity appliances with a 5 min time resolution. The aim of the modelling task is to obtain various typical consumption profiles for the selected appliances so as to determine the way micro combined heat and power ( $\mu$ -CHP) can be integrated in Canadian dwellings and how they can provide the required power.

In this frame, three target households have been simulated. They differ in terms of dwelling surface area and intensity of energy consumption.

Armstrong et al. make use of information and notably :

- estimated target household annual energy consumption for each selected appliance<sup>8</sup> thanks to three sets of use factors<sup>9</sup> that compare actual with national average appliance energy consumption ;
- domestic equipment set with their characteristics (inclusive penetration rates<sup>10</sup>) and
- occupants' use profiles.

This wide set of information was established thanks to surveys, statistical and appliances manufacturers data ; it is representative of the situation of the Canadian domestic appliances stock in 2003.

On top of that, information about unitary load cycles, cycles duration, estimated month equipment use and appliances typical nominal power have been used by the authors.

Moreover time of use data have been implemented in the model. Because of the lack of these kind of information, Armstrong et al. reused the TOU profiles defined by Pratt et al. [PCR+89] for some appliances (range, dish-washer, washing-machine) without any modification. For other devices, they performed some adaptations on the available profiles.

The load curve construction methodology, which is schematically illustrated in figure 1.10, relies on simple principles and assumptions regarding on the one hand the timing and on the other hand the magnitude of the power demand.

The management of the appliances starting is performed with the help of the TOU curves and an appliance specific chance factor  $c$  that provides the estimation of the hourly probability  $p$  of a start event according to the equation 1.9 :

$$p = \frac{f}{c} \quad (1.9)$$

where  $f$  is the hourly fraction of the daily use (hourly TOU value). For each household's appliance,  $c$  value was set by iteration so as to meet the target consumption via the annual cycles number.

With regard to the magnitude of the load curve construction and despite the individual simulation of each device, less modelling methods have been developed than the number of different appliances. So it means that a same method was applied to several devices whose functioning showed similarities.

In fact, simplified modelling approaches have been chosen by the authors :

- assignment of a constant power demand during the functioning of washing and cooking appliances ;
- the power demand of domestic cold appliances is directly constructed from measures that are adapted by the authors to match the consumption targets ;
- the load curve of lighting and other appliances is constructed from the TOU curves, a duration use that is randomly selected between chosen lower and upper bounds and nominal power values ;
- a 65 W base load represents the stand-by power demand of the domestic equipments in the household.

8. Fridge, freezer, dish-washer, washing-machine, tumble-dryer, range, lighting and other appliances

9. Non validated coefficients

10. Because of the choice of the target dwellings, the authors adapted the national penetration rates. Concretely they systematically assigned for each household some selected appliances (fridge, dish-washer, washing-machine, range and tumble-dryer) and they applied target household specific hypothesis

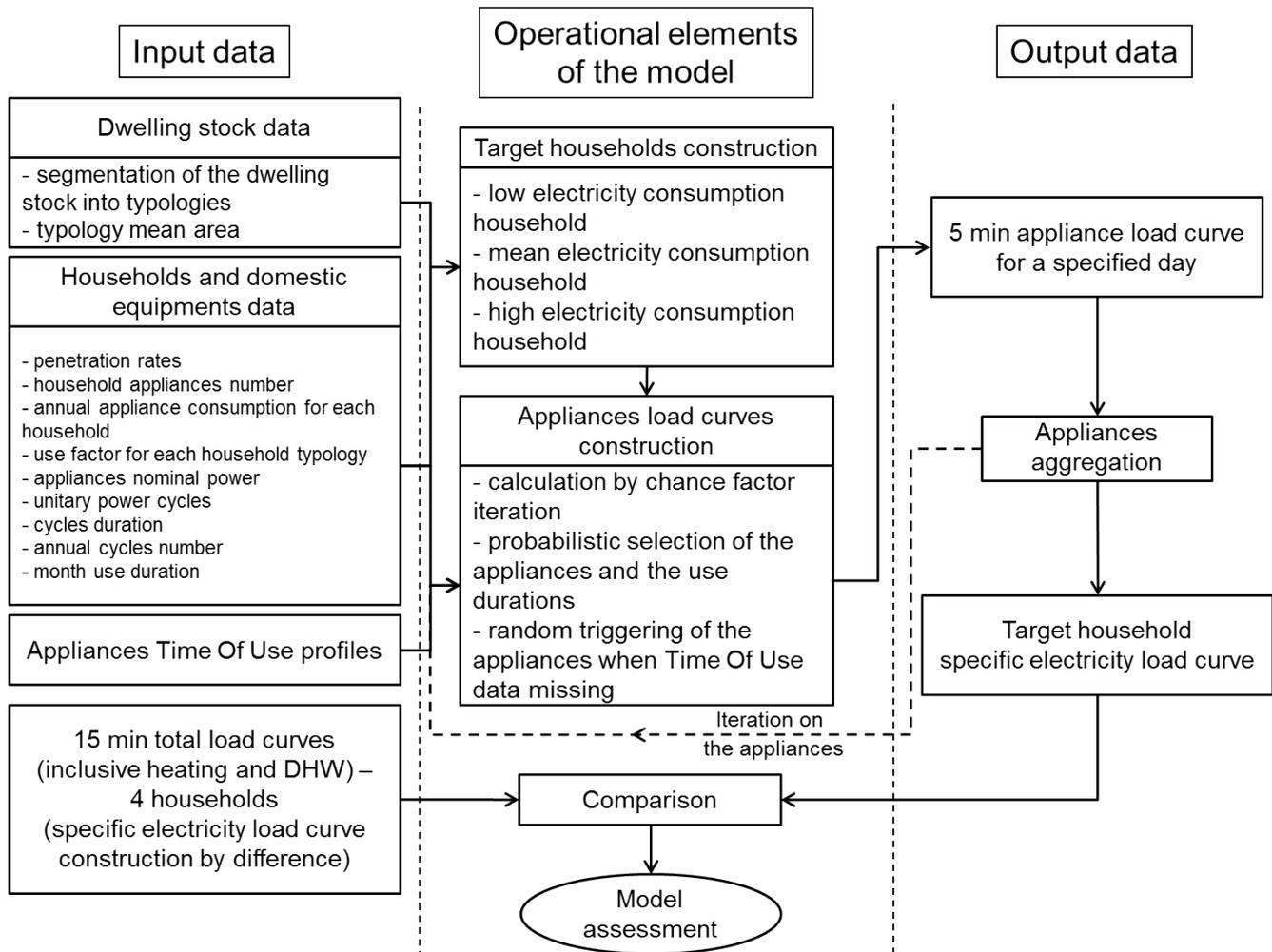


FIGURE 1.10 – Bloc diagram of Armstrong et al.'s model [ASR+09]

Then Armstrong et al. proceed with an aggregation step to get the daily load curve for a specified household : the eight previous calculated loads (one per appliance) are simply added up. The yearly load curve for each dwelling type is obtained when performing multiple simulations.

In order to evaluate the goodness of fit of their model, the authors carried out comparisons between simulation results and load curve measurements. These have been collected by Hydro Quebec in the 1990's during a campaign focused on electric heated detached houses. With this campaign total, heating and DWH loads have been simultaneously measured at a 15 min time resolution. By difference, specific electricity load curves were obtained. Four houses from this campaign have been selected to proceed the evaluation task.

First, the visual comparison showed the good ability of the model to be near the reality in terms of peak loads, mean loads and annual energy consumptions.

Secondly, Armstrong et al. notice that the model produces less repetitive results than what is seen in the measurements. For the authors it does not constitute a weakness of their model since the generated loads will be used as inputs for the simulation of  $\mu$ -CHP : variable demand profiles lead to stress on the CHP that are more difficult to comply with.

Thirdly a statistical comparison (study of the repartition of the values in 100 W ranges) has been proceeded leading to two main observations :

- the selected power level for the modelling of stand-by modes (65 W) is not optimal ;
- the second observation is more trivial : in measures, two 100 W-ranges of power demand are empty what is not the case of simulated results. This can be explained by the domestic appliances sets in the

selected households.

This load curve reconstitution tool shows a series of advantages : first its relative simplicity, secondly the reasonable required amount of input data compared with the other models and thirdly its ability to generate various specific electricity demand profiles. Yet some weaknesses can be found : the used TOU curves probably have few validity because they are not recent (1989) : domestic habits can have evolve significantly. Moreover, apart from the used load curves for refrigerators and dishwashers that come from measures, the other unitary load cycles are built up when assuming a constant power demand during the functioning what is a real simplification of the reality. Finally the model does not make the difference between the types of the days (weekday, weekend days for instance) and the influence of the weather only concerns the lighting and it is very simplified in this case.

#### 1.2.4.4 Widén et al.'s model

Widén et al. expose in [WLV<sup>+</sup>09] the load curve generation model they built up. This TOU data based model makes use of simplified end-use to power demand conversion schemes, natural daylight density function and water consumption ; moreover it defines and simulates « mean<sup>11</sup> » domestic equipments. The main purpose of this model is to generate electric load curves for domestic appliances that will then serve as inputs of other simulation tools. Among them are modelling programs of small scale decentralized electricity generation systems or domestic water production equipments based on solar energy. Another purpose of this simulation tool consists in studying the modifications on the domestic load curve in a prospective point of view when considering the changes in terms of behaviour and energy efficiency of the appliances.

In the authors' opinion, a load curve generation model can be classified according to two criteria : its spatial and temporal resolution.

The spatial resolution corresponds to the smallest simulation level where a tool is able to provide a curve of need (electricity, heating, water...) : from a unitary domestic equipment to a household type to the national residential field. The temporal resolution of a model is linked with the time step results are returned : from seconds to hour.

In terms of input data, the main requirements of this model are time of use information with a 5 min time resolution that have been collected during a series of campaigns. Concretely, it is constituted by empirical presence at home scenarios and domestic activity sequences for a weekday and a weekend day. During them, each occupant of the selected households had to write down his activities and the corresponding start and stop times of them. That notably includes the instants when getting up and to go to the bed and when moving out and to home. Thus the model provides a load curve for each household's member and can construct load curves for a mean domestic customer.

On top of time of use information, end-use load curves measurements and domestic water profiles have been used for the elaboration of this model.

In the whole, five surveys have been used as input data sources, comparison elements data bases and validation references.

The original characteristic of Widén et al.'s model, whose architecture is described in figure 1.11, is the use of end-use to power demand conversion schemes. This leads to a simple modelling way rather than a very precise one complying with the authors' wish that is a satisfactory load curve representation at a 1 h resolution.

Concretely, for each domestic activity that implies an electricity and/or a water consumption the authors assign a standard end-use pattern ; each of them is described by a series of parameters. These were set so that a realistic average consumption is got when they are applied to the simulated households. Nominal power and use duration were estimated thanks to a campaign of tests. Activities whose electricity and/or water needs were insignificant have been neglected. This modelling approach constitutes an essential feature of Widén et al.'s model.

Thus, five possible schemes have been built up :

---

11. In terms of performance

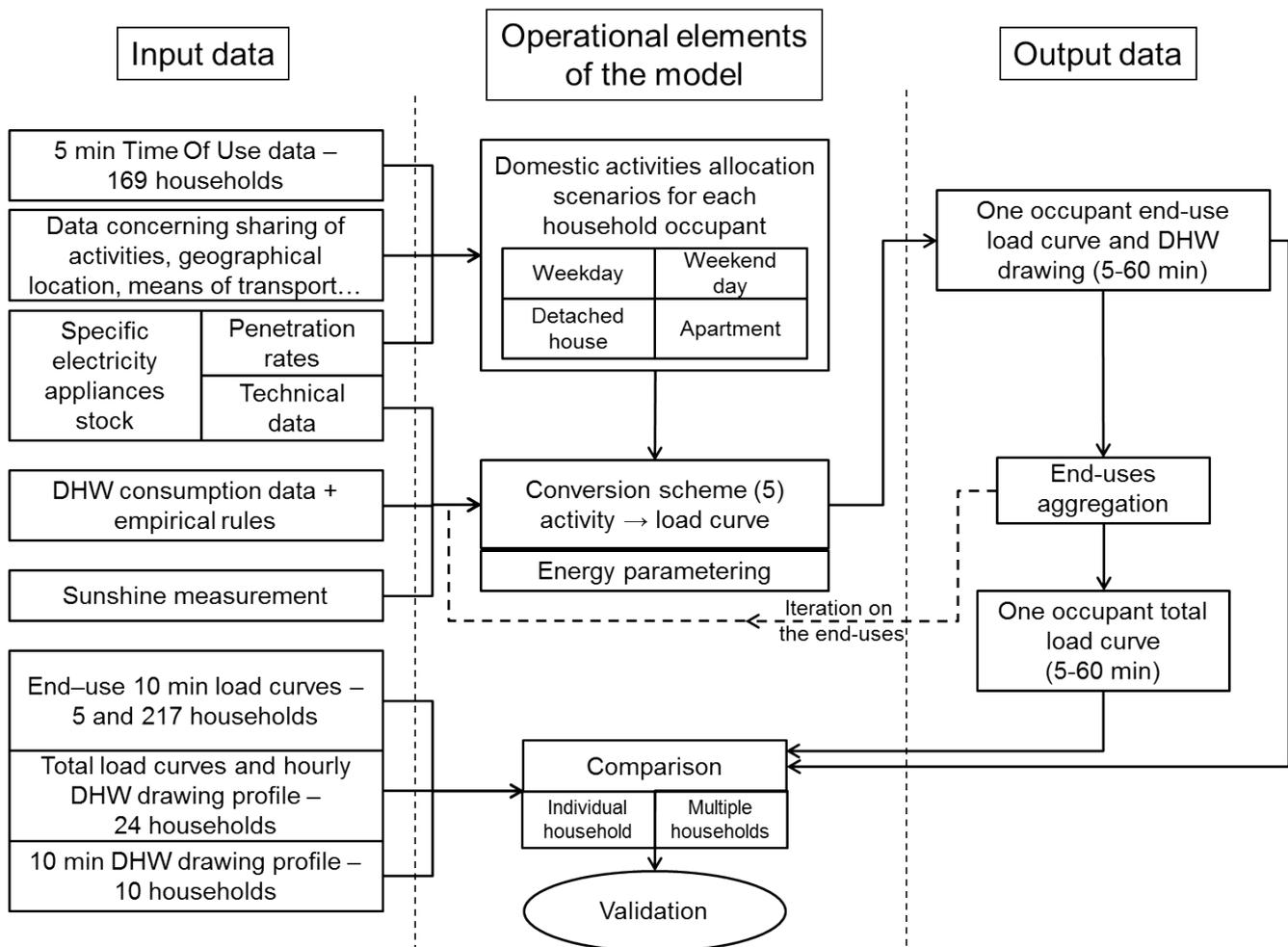


FIGURE 1.11 – Bloc diagram of Widén et al.'s model [WLV<sup>+</sup>09]

1. either power demand is disconnected from the activity. That is the case of cold appliances whose modelling is a constant power demand in functioning mode ;
2. or power demand is constant during the activity. This scheme is applied to a lot of domestic activities from cooking to ironing and to TV watching. Thus it constitutes a strong assumption in the sense that it is applied to very different appliances ;
3. or power demand is constant after the activity. Dishwasher, washing-machine and tumble-dryer are concerned by this scheme. In fact, the corresponding domestic activity is the filling up of the washing appliance. Although power demand effectively is required immediately after the end of the domestic activity, the authors admit that they widely simplify the reality ;
4. or power demand is constant during the activity with addition of a temporal constraint. This scheme is applied to the majority of activities that imply domestic hot water especially when drawings-off occur in the first part of the task. The constraint restrains the reuse of already drawn hot water for the following activities ;
5. or finally the domestic activity requires a fluctuating power demand. The modelling of lighting follows this assumption. More precisely, the power demand fluctuates according to the available daylight : if somebody is at home in an aware state (i. e. not sleeping) the maximum power demand level  $P_{max}$  is applied when the daylight availability  $L$  is low (under a  $L_{min}$  level) and the other way around ( $P_{min}$  when  $L > L_{max}$ ).

Because of the lack of information concerning the use of DHW, the authors had to make use of empirical rules.

Widén et al.'s model was developed in MATLAB. Load curve time resolution can be selected by the program user (from 5 min to 1 h). Distinction between detached house and flat on the one hand and between weekday and weekend day on the other is made by the simulation tool.

As validation, the authors simulate the power demand requirements of the households described in the 1996's statistical survey (that contains among other things time of use data and equipments characteristics) and compare them with the results coming from the latest (2006) available load curve measuring campaign. Thus, two assumptions are conceded : domestic appliances and TOU data have not changed significantly between 1996 and 2007 enabling the possible use of the input dataset.

To some extent, the results (individual households' total load curves, i. e. all simulated end-uses aggregated) are in accordance with the comparison references notably in terms of load curve shape and load peak prediction (high consumption days and households).

Yet, Widén et al. notice that the relevance of the results becomes weaker when considering the end-uses separately. For instance, the power demand modelling for end-uses that are simulated according to the second conversion scheme is more satisfactory than these when power requirement and activities are disconnected.

Moreover, the authors carried out other simulations in order to make comparisons at an aggregated scale especially on a households sample that represents the Swedish population. The model shows a good ability to differentiate the average power demand shape that characterizes a detached house from an apartment on the one hand, a weekday from a weekend day on the other. To go further in the comparison, the authors calculated the  $NVF$  for some end-uses : this shows similarities between measures and simulations. Yet when observing more precisely the outputs of the simulation tool, some discrepancies can be shown as for example an undervaluation of the power demand for a detached house during the night. Widén et al. suggest that this is due to the miscellaneous end-uses that are not considered by the model. Other divergences can be seen as for instance temporal move of the modelled peak load for some end-uses compared with measurements.

From TOU data and information about the stock of domestic appliances, the authors built up a simulation tool that reconstructs load curves for different households' and days' typologies. With the help of a restricted number of domestic activity to power demand conversion schemes, Widén et al. get satisfactory results when considering the load curves at aggregated scales (wide simulated sample, 1 h time resolution). However, the tool is directly modelled on the available time of use data, that is to say that it represents the 1996's situation (non scalable modelling). On top of that, these data have been collected in diaries that were filled in by the selected households' occupants. The use of the collected data needs a good quality of them what is not ensured as underlined by the authors. Moreover, there is no modelling of electricity consumption diversity because it is embodied in the input data. As a consequence, the integration of a new end-use seems to be difficult because it implies the implementation of the corresponding time of use scenario in the available diaries. The model corresponds to the Swedish housing stock where electric heating appears as an exception. That is the reason why this end-use (and the thermal issue in general) is not considered in the model.

At the end of [WLV<sup>+</sup>09], Widén et al. list the possible improvements of their model : they notably announce the use of a Markov-chain and of density functions to generate diversity in terms of human behaviour on the one hand, its changes in time on the other. These evolutions have been implemented later in Widén and Wäckelgård's model.

#### 1.2.4.5 Widén and Wäckelgård's model

In [WW09], Widén and Wäckelgård expose the model they developed : this is an high time and spatial resolution tool that generates domestic activities scenarios. These are then converted into electric load curve.

The main improvement of this model compared with Widén et al.'s is the use of a non homogeneous Markov-chain that generates diverse domestic activities sequences with high resolution levels (1 min time

steps, modelling of each household's member). Only the activities that imply electric power demand are taken into account by the model : heating and domestic hot water supply<sup>12</sup> are notably excluded. Load curve synthesis follows the same two-step process as in Widén et al.'s model :

1. domestic activity patterns are allocated to each dwelling's occupant ;
2. activities are converted into power demand thanks to conversion schemes.

The functional architecture of the model is schematically represented in figure 1.12.

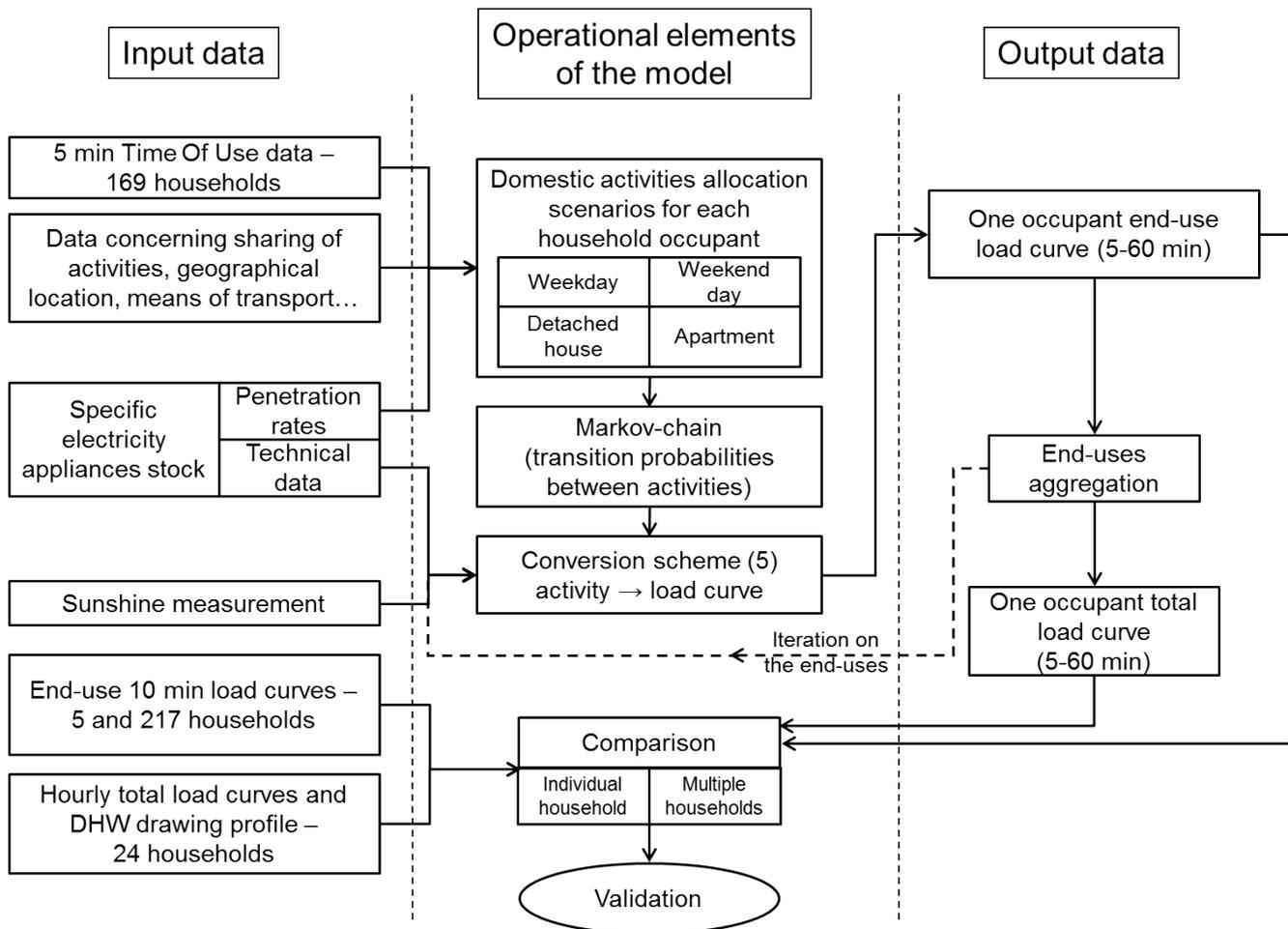


FIGURE 1.12 – Bloc diagram of Widén and Wäckelgård's model [WW09]

As expected, input data of this model is the same as the previous model's : time of use for domestic activities coming from statistical surveys and weather data. From these sources of information, a Markov-chain (firstly applied to the lighting and fully detailed in [WNW09]) has been built up : this is the main feature of Widén and Wäckelgård's modelling approach. The model assumes that at any simulated time  $t$ , each occupant of a household must be in one of the three following availability states :

- absent : the person is not at home ;
- present and inactive : the occupant is at home and sleeps ;
- present and active : the dwelling's member is at home and can carry out each kind of domestic activity.

Nine domestic activities/end-uses<sup>13</sup> are dependent on these availability states especially the « present and active » that makes them possible : in the Markov-chain, they are represented by a code (number). Domestic cold and lighting are independent of last availability state and are not driven by the Markov-chain.

12. Very little electricity demand is dedicated to these end-uses in the Swedish housing stock

13. Away, sleeping, cooking, dish-washing, washing, TV, computer, audio appliances and other devices

Starting with the carrying out of the domestic activity  $a_1$  at time  $t$ , the stochastic process calculates the transition probability to change to activity  $a_2$  ( $a_1 = a_2$  or  $a_1 \neq a_2$ ) at  $t + 1$  with the help of the input data. Transition probabilities have been calculated at a 1 h time resolution. Even if load curves can be computed at a better resolution, activities' transition probabilities stay constant over one simulated hour. However the probabilities are not fixed along the day (non homogeneous Markov-chain), so the daily fluctuations of the activities are taken into account by the process. At each time step, these calculated figures are compared with a random number generated by a  $[0; 1]$  uniform density function. According to the sign of the difference, activities are launched or not.

The detailed calculation process is described in [WW09]. For few cases, the authors adapted it to ensure its coherence (especially for night periods).

As in Widén et al.'s model, domestic activity to load curve conversion schemes have been built up with the help of measured load curves. Moreover these have been refined compared with those in [WLV<sup>+</sup>09] especially in three points :

1. appliances' power demand can occur before, during, just after the corresponding domestic activity or even totally disconnected from it ;
2. unitary power demand pattern and stand-by power level are characteristic of each appliance ;
3. appliance sharing between household's occupants is taken into consideration.

Concerning this last point, the authors underline that this kind of data are not recorded in the time of use diaries. Thus a supplementary assumption is made by Widén and Wäckelgård : a few appliances are totally shared by all dwelling's members (e.g. the fridge) whereas some other are systematically used by a single occupant. According to the device status, the related domestic activity can be carried out by several people at the same time (leading to a single power demand cycle) or not (the activity can only be conducted by a single person and it systematically leads to a power demand pattern).

In [WW09], more details are given about the modelling of each activity related appliances. In fact, assumptions should have be done in order to adapt the modelling method to each domestic device (respectively each appliance related to the same domestic activity).

So as to validate their model, Widén and Wäckelgård proceeded with two kinds of modelling tasks and comparisons taken the Sweden Energy Agency's data as references : first they tested the model on a small scale housing stock (14 dwellings) in order to evaluate the ability of the model to generate diversity. Then they compared the generated load curves at a more aggregated scale (two hundreds simulations for a similar dwelling type during a week – five weekdays and two weekend days, two dwelling types considered detached house/flat) so as to know if the model provides a good approximation of the energy consumption with regard to the statistical national energy consumption data.

To compare simulated and measured load curves, the authors chose the two following criteria :

- the load factor  $\underline{F_{n, Pmax}}$  related to the maximal power demand for a set of  $n$  clients during an interval  $\Delta t$  :

$$F_{n, Pmax}(\Delta t) = \frac{\overline{P_n(\Delta t)}}{P_{max, n}(\Delta t)} \quad (1.10)$$

with  $\overline{P_n}$  and  $P_{max, n}$  respectively the mean and maximal power demand required by the  $n$  consumers over the time period  $\Delta t$  ;

- the diversity factor  $\underline{k_{f, n}}$  related to  $n$  electricity users and the time slot  $\Delta t$  :

$$k_{f, n}(\Delta t) = \frac{\sum_{j=1}^n P_{max, j}(\Delta t)}{P_{max, n}(\Delta t)} \quad (1.11)$$

where  $P_{max, j}$  is the maximal individual power demand value for the client  $j$  and  $P_{max, n}$  the maximum demand of the  $n$  consumers considered in the whole.

Both equations give a summarized analysis when considering different power demand profiles and thus explain their frequent use in such a study.

In the whole, comparisons provide the same kind of conclusions that have been made for Widén et al.'s model : the simulation tool gives satisfactory estimations of domestic load curves, generates sufficient differences in terms of timing and magnitude of the power demand between weekdays and weekend days on the one hand, between house and flat on the other hand.

However some improvements can be carried out such as a better modelling of some end-uses, a refinement in terms of domestic activities and a more suitable distribution of the domestic equipments according to the dwellings' properties. Carrying out these evolutions necessarily requires more detailed input data. These come from surveys and their precision directly depends on the attention of the selected people sample. Moreover, the model is only relevant for short term forecasting since time of use data always represents the past situation. This is notably what causes the little ability of the model to simulate the Information and Communication Technologies (ICT) load curve (the computer is the best example). Thus to maintain a satisfactory precision of the model, updated time of use data have to feed the simulation tool imposing regular to continuous surveys. Although heating is not considered in this model, the authors underline that its integration is feasible. Yet, without apparent help of a building simulation tool, we can wonder about the precision of this end-use modelling.

#### 1.2.4.6 Richardson et al.'s model

In order to quantify the impact of « low carbon measures », such as demand response and CHP, on a local electricity distribution network, Richardson et al. developed a model presented in [RTIC10].

The authors adopted a « deep » bottom-up approach since the simulated elements are the specified household's domestic equipments and the households' members whose individual behaviour modelling is based on activity diaries built up from surveys. Thus Richardson et al. are able to reconstruct the 1 min electric load curve corresponding to a few number of dwellings located in the same area. According to the authors, this simulation time resolution has been considered as the best compromise between the amount of produced results and the required precision level at a small spatial scale.

The structure of the simulation tool is schematically drawn in figure 1.13.

As figure 1.13 underlines, Richardson et al. made use of the results of a time of use survey in order to get availability scenarios (presence at home and awake state). According to the authors, the occupant's state « at home and active » is the most interesting because in this case electric devices can be run. In the model, this kind of information is coded by a binary variable which equals 1 when an occupant is at home and active, 0 otherwise. An aggregated availability profile is obtained when adding each occupant's availability variable.

Yet the active occupation only corresponds to the necessary condition to use an equipment, the next step to construct a load curve is to know the kind of device that is run, the use duration and the start times. That is the reason why Richardson et al. have made use of daily activity profiles (probability density over the day at the diaries' time scale) elaborated from the TOU survey for seven different domestic activities. Here we recognize the same input material as in Widén and Wäckelgård's method [WW09]. However these elements differ from the Swedish model because they are « static » : in Richardson et al.'s model, daily activity profiles are only functions of the day-type (weekday/weekend day) and the number of occupants in the simulated household (from 1 to 5). Using such kind of data enables a realistic human behaviour modelling on condition that first diaries have been written down carefully and secondly the people sample is large enough to « smooth » domestic habits.

So as to provide a satisfactory device modelling, Richardson et al. used statistical information (penetration rates, devices' and households' annual consumption) and some measured power demand patterns – all input data are available in [RT10]. When this last type of information was unavailable for the appliances, constant power demand has been assumed. They sometimes had to adjust them in order to reach the average East Midlands's household that constitutes the target annual domestic consumption.

Domestic activities can imply the need of electricity without forcing it. That is the reason why Richardson et al. had to make connections between activities and appliances : according to the authors' classification of the considered domestic activities (watching TV, cooking, laundry, washing/clothing, ironing, cleaning, other activities), some links are obvious because one activity corresponds to a single device (watching

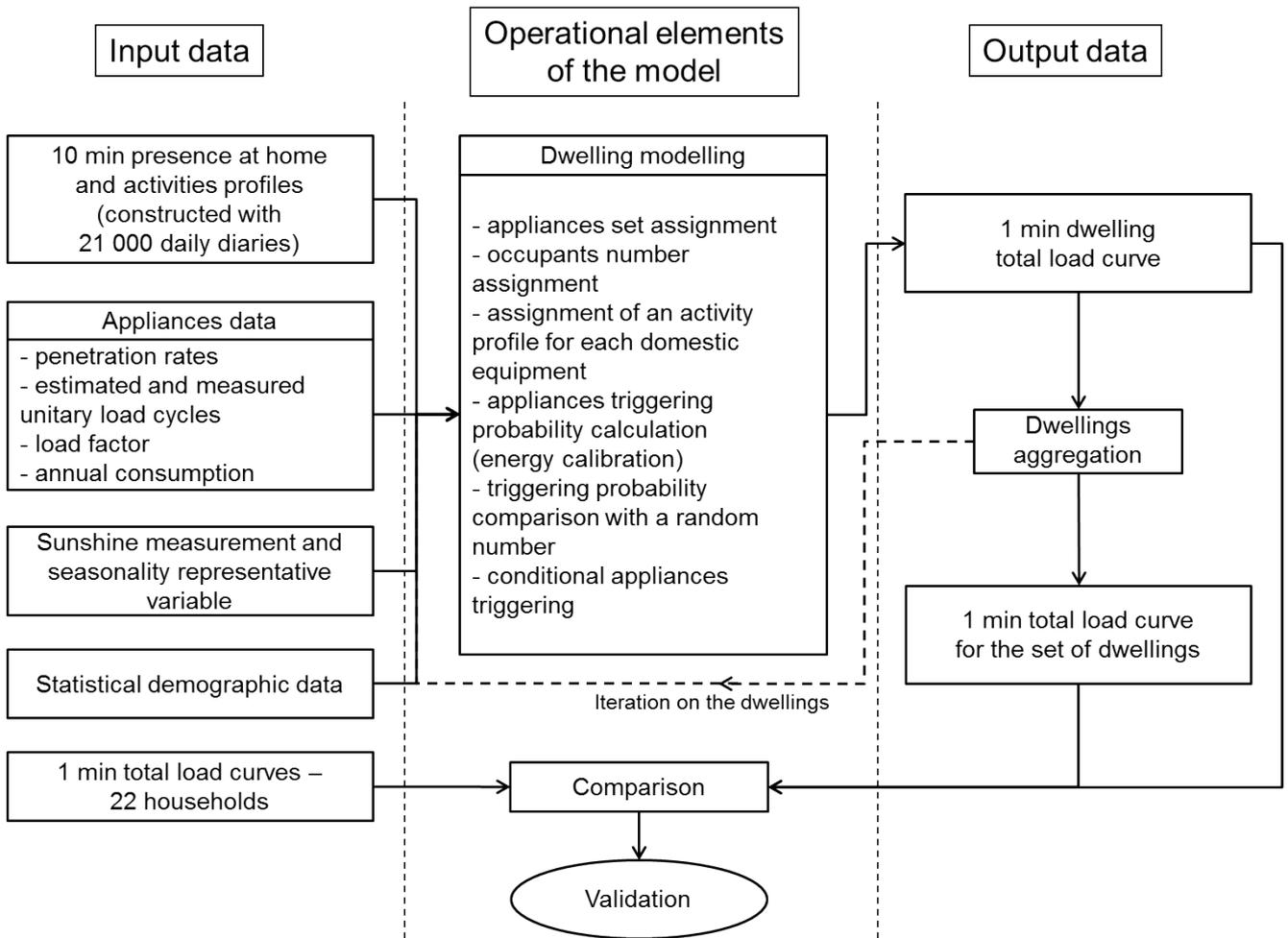


FIGURE 1.13 – Bloc diagram of Richardson et al.'s model [RTIC10]

TV  $\Leftrightarrow$  television, 1-1 relation), other are more « ambiguous » (cleaning  $\Leftrightarrow$ ?, 1- $n$  link). However, last case does not mean that when carrying out an activity linked with  $n$  electric devices, all  $n$  appliances are turned on as long as the activity is conducted. It only allows a possible simultaneous use of the same activity related appliances.

« Other activities » related appliances have been put together by the authors because :

- either any of the other activity profiles did not match with them (e.g. telephone). Thus they are only dependent on the human presence at home,
- or their functioning is continuous (e.g. cold appliances) and their modelling is not function of the availability at home,
- or finally their use is much more linked with an exogenous parameter such as the outer temperature what is the case of heating.

When simulating a sample of households, their set of domestic appliances has been randomly assigned with respect of the national penetration rates (multi-equipment is taken into account). During a simulation, an electric device can be in two different states : turned on and turned off. For some appliances, last state imply a standby consumption which is taken into consideration in the load curve modelling. Moreover, in Richardson et al.'s opinion, sharing of equipment and use of linked devices (e.g. washing-machine/tumble-dryer) can be « read » directly from the time of use data that have been arranged in the adequate form to serve as input of the model.

To calculate a load curve at each time point, four steps are required :

1. according to the selected appliance, the number of active people at the simulated time step and the type of the day, the corresponding daily activity profile is chosen ;

2. the calculated time point value is read on the activity profile ;
3. then the device starting probability has to be calculated. It is done by multiplying the value obtained in step 2 with a scalar factor. In fact, this last figure ensures the energy accuracy of the model because it relies on the mean annual number of uses for each equipment (device energy consumption/mean energy consumption per cycle) ;
4. finally, the device launching probability is compared to a random number generated thanks to a  $[0; 1]$  uniform density function. If the probability is greater than the random number, the appliance is launched.

In order to validate the results of their model, Richardson et al. conducted an in-depth multi-criteria study. Comparisons have been carried out considering following aspects :

- annual and mean daily energy consumption ;
- daily load curves ;
- 1 min power demand volatility ;
- diversity of the power demand ;
- load duration curve ;
- load factor.

As reference data, Richardson et al. make use of 1 min load curves measured in a sample of 22 dwellings located in the East Midlands area. These validation data is totally independent of the values that serve as model inputs.

The following points sum up the main conclusions made by the authors :

- electric heating has been excluded from the validation study because of a too low penetration rate ;
- one year simulation for 22 households shows that calculated energy consumption matched well with the validation data in average but its standard deviation was less than the real one. Assumptions of the model and input data biases can be the causes ;
- in terms of average daily consumption, inter-month variations are correctly modelled but seasonality influence is under-evaluated. Lack of precision in the diaries and non modelling of the heating auxiliaries can explain this observation ;
- when considering the mean daily calculated load curve, the model under-estimates power demand during the night and the morning peak occurs later than in the reality. The fact that the measured households sample does not match the national case can be responsible for this difference ;
- the model shows a little ability to represents little and large power demand transitions<sup>14</sup>. To the authors' opinion, modelling simplifications are largely responsible for these observations ;
- diversity factor and After Diversity Maximum Demand (ADMD) values from the model are close to the real indicators ;
- load duration curves from reality and model are quite similar.

As a conclusion, Richardson et al.'s model provides high spatial and time resolution load curves thanks to detailed time of use data. Its strengths and weaknesses have been identified by the authors and mostly listed in the previous paragraphs. We would underline the unknown concerning the modelling of heating. In [RTIC10], Richardson et al. do not mention the use of a building simulation tool which seems to us necessary to build up a relevant load curve at a fine time step. Concerning the appliances modelling and more precisely the nominal power values of the equipments, only mean values have been used in the model. A statistical distribution can have generate more variations and diversity between the simulated households. Moreover, the model is restricted to short-term forecasting because new end-uses can not be integrated. Finally, households' socio-economic characteristics, price of energy and DSM measures are load curves influences that are not included in Richardson et al.'s model.

---

14. Transition =  $|P(n+1) - P(n)|$  where  $P$  is the power demand value at the time step  $n$

### 1.3 Statistical-engineering model : Train et al.'s model

Previous models either follow an upward (bottom-up) or downward (top-down) process that both present advantages and limits. From this assessment, the main motivation to develop an other kind of model is to put to use all benefits of each approach. That is what Train et al. [THW85] tried to do in building up a Statistical Adjusted Engineering (SAE) model.

The methodology corrects the end-use load curves simulated with an engineering approach (so outside the model) thanks to statistical coefficients obtained through measured load profiles. In terms of input data, Train et al. made notably use of a PG&E metering campaign [TRA92] that provided them total and by end-use load curves for about 800 domestic clients. During this survey information concerning the households (income, dwelling size area, number of occupants, appliances penetration rates...) has been collected. Moreover engineering loads have been calculated by ADM Associates using the method described in [EPR91].

Thus engineering power demand values correspond to the explanatory variables of a conditional demand model into which they are put. The general functioning of Train et al.'s model is shown under a bloc diagram in figure 1.14. Results are given at the 1 h time resolution.

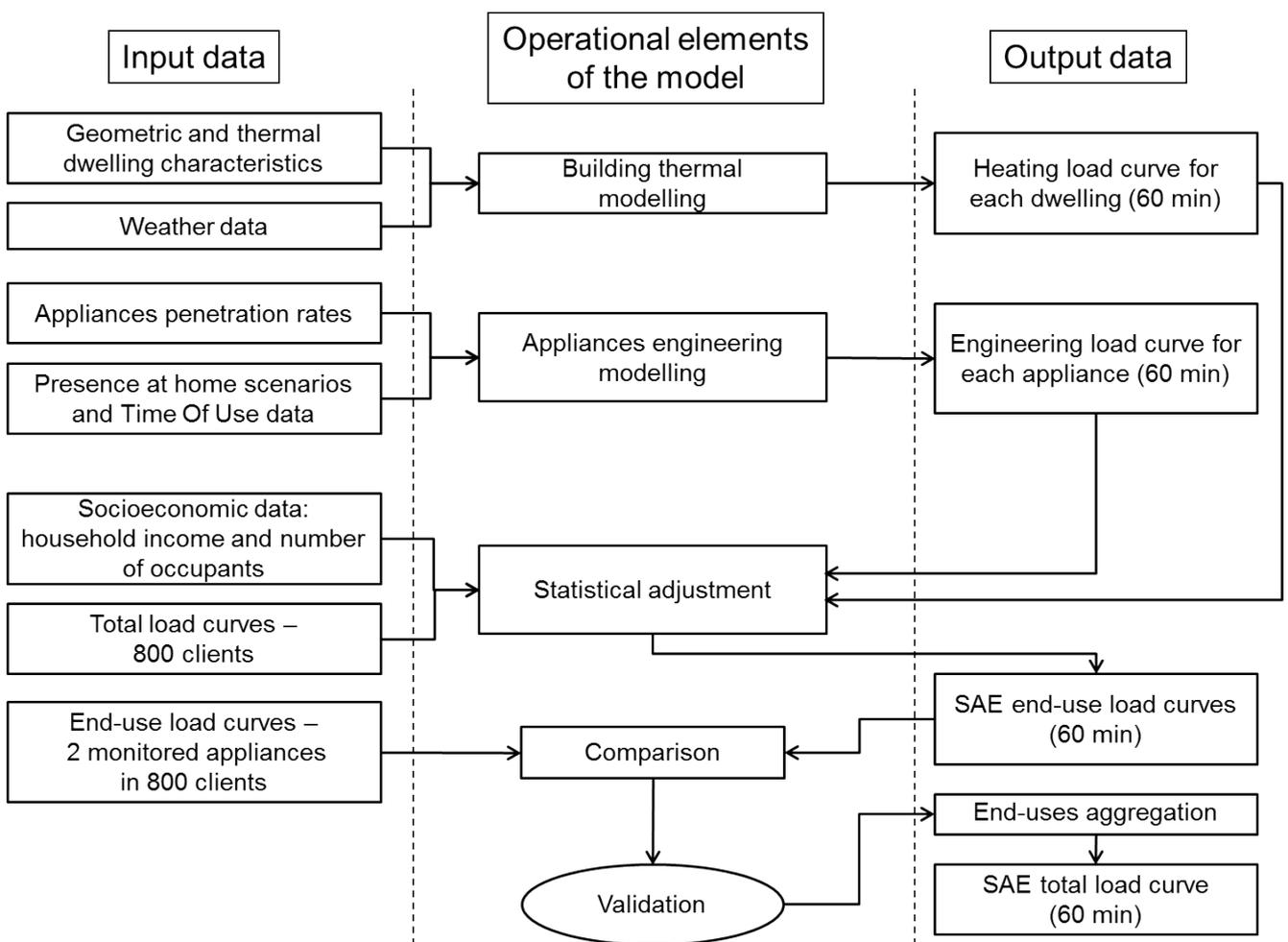


FIGURE 1.14 – Bloc diagram of Train et al.'s model [THW85]

As it is said in [TRA92], statistical adjustments on load curves can occur at each time step (1 h), for each end-use  $i$  and/or for each customer  $j$ . However it does not seem to be relevant to adjust engineering load curves on these three sources of influence because it would totally « hide » the simulated part of the load curve results.

That is the reason why the authors constructed four SAE models with progressive adjustments. These

are applied to one or several time periods  $l$  ( $l \in [1; \tilde{l}]$ ) defined within the day on the one hand and to take into account the customers' characteristics on the other hand.

In the first defined SAE method, engineering load curves are adjusted with a correction factor  $\zeta$  that is specific for each end-use  $i$  ( $\zeta = \zeta(i)$ ). Yet this coefficient is constant over the day (only one period is defined) and it does not take into account the customers' characteristics. In the second SAE model, these characteristics are taken into consideration for the adjustment of the load curves ( $\zeta = \zeta(i, j)$ ). However, once again only one period is defined for the adjustment that is to say that  $\zeta$  is not time-dependent. In the third SAE approach, four periods  $l$  are defined and the adjustments are different according to the end-use and the period but without considering the customers' specificities ( $\zeta = \zeta(i, l)$ ). Finally the last SAE model variante is the combination of model 2 and 3 : engineering load curves are corrected according to the end-use  $i$ , the customer  $j$  and the period  $l$  ( $\zeta = \zeta(i, j, l)$ ).

Whatever the adjustment level, Train et al.'s model is looking for estimating the total power demand of customer  $j$  at time  $t$ ,  $P_j(t)$ , in rectifying the simulated power demand value at  $t$  for the end-use  $i$  of the customer  $j$ ,  $R_{i,j}(t)$ <sup>15</sup>, thanks to the correction factor  $\zeta_{i,j,l}$ <sup>16</sup> and an error  $u_j(t)$  according to equation 1.12.

$$P_j(t) = \sum_{i=1}^n \sum_{l=1}^{\tilde{l}} \zeta_{i,j,l} \cdot R_{i,j}(t) \cdot D_{i,j} + u_j(t) \quad (1.12)$$

In few words, the authors observe that different adjustment periods provide a better correction of the load curves calculated with engineering models. However, introducing demographic variables (customers' characteristics) in the model do not seem to be of great interest when considering the gain in terms of results accuracy.

A wider evaluation and validation work has been conducted by Train in [TRA92]. This publication mentions an extended version of the fourth model (adjustment on nine periods, differentiation in terms of type of day, adaptation of the model according to the season summer/winter...) but conceptually it is the same method than this explained above.

In the whole, the assessment of the model has lead to general conclusions that we previously mentioned as for instance the limits of the engineering models to take into account the time of use and the regulation of thermal appliances. On top of that, Train compared the SAE, engineering and measured daily load curves with the help of the statistic  $\tilde{S}$  defined in equation 1.13.

$$\tilde{S} = \frac{\sqrt{\sum_{t=1}^{24} (P_{est}(t) - P_{mes}(t))^2}}{\sum_{t=1}^{24} \frac{P_{mes}(t)}{24}} \cdot 24 \quad (1.13)$$

where  $P_{est}(t)$  and  $P_{mes}(t)$  are the 1 h power demand value averaged on the sample of simulated customers that is estimated by the SAE (or the engineering) method and measured respectively. In fact, this is the square root of the  $NVF$  that we previously have met.

More precisely, Train focused his assessment on the comparison between SAE and engineering loads in order to quantify the impact (improvement and/or deterioration according to the considered end-use) of the statistical adjustment. The conclusion of the author is that SAE model improves the engineering results if and only if physical models do not integrate empirical knowledge (so from measures) or the structure and/or the size of the sample used for the adjustment is (are) not optimal in terms of representativeness.

To conclude, Train et al. developed an hybrid model that takes advantage of the traditional load curve modelling techniques. Measured load curves calibrate power demand profiles obtained by engineering

15.  $n$  end-uses are considered – a dummy variable  $D_{i,j}$  indicates the presence or the absence of each of them in the customers' homes

16. That depends on the end-use  $i$  and/or the consumer  $j$  and/or the adjustment period  $l$  according to the SAE model variante

models. On top of that, this method allows to get round the diversity modelling because it is contained in the adjustment coefficients.

However this approach is only relevant for short term forecasting because it reproduces past power demand profiles what is the case of some other models. Changes and evolutions in the housing stock can not be implemented and simulated in this model. On top of that, the miscellaneous category encompasses a lot of equipments. The corresponding load curve can be well estimated in the whole but what about each of the appliance in this field ?

## 1.4 Cross analysis

In this section, the authors sum up and compare the models of the literature that were described in the previous paragraphs. First of all, a general overview of them is given under a table form. This contains the in-our-opinion most important criteria that characterize a load curve reconstitution methodology. Then certain comparisons are proceeded with focusing on the advantages and the limits of each methodology. Finally, we plot the results of this cross analysis through a « pseudo 3D » representation.

### 1.4.1 Reading grid

To construct this reading grid, we began to list the to-our-mind most relevant characteristics for defining a load curve reconstitution model : these criteria correspond to the columns of the table and the models are arranged in rows.

Then we completed each cell of the table thanks to a restricted number of modes so as to facilitate the reading. This way we obtained the tables [1.2](#) and [1.3](#).

Type of model	Authors [Publication] Purpose of the model	Inputs of the model	Outputs of the model	Modelled end-uses	Nature of the generated behavioural diversity	Model validation
Statistical random bottom-up model	Aligner et al. [ASK84] To get end-uses load curves without monitoring	Assumptions concerning the non functioning of a series of equipments at certain hours within the day Measured daily load curves at the household level	Daily load curves returned at a 60 min time step for nine domestic appliances at the household level	Specific electricity appliances, heating, cooling, DHW	Deterministic	No validation
	Bartels et al. [BFG192] Simulation of the impact of various scenarios on the load curve at the regional level Decision support for the planning of the capacities of the utilities	Evolution scenarios concerning technological aspect, socio-economical aspect, appliances penetration rates, number of customers Measured daily load curves at the household level	Daily load curves per appliance returned at a 60 min time step for the regional residential class considered. The load curves are calculated according to a scenario (worked day or not worked day for a selected month)	Specific electricity appliances, heating, cooling, DHW	Deterministic	On total load curves
Statistical random bottom-up models	Yao et Steemers [YS05] Support for the design of energy systems including renewables Support for the prediction of the load curve for a selected community	Occupation scenarios Statistical daily consumptions by end-use and appliance	Daily multi end-uses load curves returned at a 1, 5, 15 or 30 min time step corresponding to, from a single household, to an entire community	Specific electricity appliances, heating, DHW	Mathematical random	On total load curves
	Walker et Pokoski [WP85] Prediction of the load curve as support for the planification of new power generation capacities  Cappaso et al. [GLP94] Support for the utilities for the evaluation of power management measures and DSM policies	Occupation, activities and end-use scenarios Unitary load cycles (from modelling and apparently from measures)  Occupation and activities scenarios Constructed unitary load cycles Consumption available at the household level (with an ambiguous temporal restitution of these data) Statistical mean annual consumption per appliance (from measures but with an ambiguous temporal restitution of these data)	Daily load curves returned at a 15 min time step for a household (or a sample of households) that is specified with taking into account of the psychological and behavioural influences  Residential daily load curve returned at a 15 min time step for a geographical area thanks to the synthesis of multi end-uses load curves at the household level	Specific electricity appliances, DHW  Specific electricity appliances	Scenario based probability  Scenario based probability	On total load curves  On total load curves
Time of use based bottom-up models	Armstrong et al. [ASR <sup>+</sup> 09] To get electricity consumption profiles for the modelling of micro-generation devices	Occupation and activities scenarios Measured and constructed unitary load cycles Mean annual consumption available at the household level Mean annual consumption of the appliances	Daily load curve for electricity specific appliances returned at a 5 min step for typical households	Specific electricity appliances	Scenario based probability	On reconstructed load curves for specific electricity appliances

TABLEAU 1.2 – Reading grid (first part) of the studied load curve reconstitution models

Type of model	Authors [Publication] Purpose of the model	Inputs of the model	Outputs of the model	Modelled end-uses	Nature of the generated behaviour /voural diversity	Model validation
Time of use bottom-up models	Widén et al. [WLV <sup>+</sup> 09] Determination of the households' energy consumption Evaluation of future changes (modifications of the behaviour, energy efficiency improvement) that impact on the electricity use	Occupation and activities scenarios Constructed unitary load cycles	Daily load curves returned at a 5 to 60 min time step and DHW draw profiles (ambiguous temporality) at the occupant level. Aggregation of previous results to get those corresponding to a household or a little community	Specific appliances, DHW electricity	Scenario based determinism	On load curves per end-use On DHW draw profiles On total load curves On daily electricity consumption per end-use On yearly consumption for DHW
	Widén et Wäckelgård [WVW09] Support for studies on electricity production at the household level	Occupation and activities scenarios Measured daily load curves per appliance Constructed unitary load cycles	Daily load curves returned at a 1 to 60 min time step for a household and possible aggregation to get results at a larger scale	Specific appliances electricity	Scenario based probability	On total load curves On the household's yearly consumption On yearly consumption per end-use
	Richardson et al. [RHC10] To anticipate the impact of « low carbon » measures on the load curve of a local distribution network	Occupation and activities scenarios Measured unitary load cycles Statistical yearly consumptions per appliance based on studies	Daily load curves returned at a 1 min time step for a household and possible aggregation for a set of households	Specific appliances, heating, DHW electricity	Scenario based probability	On total load curves
Probabilistic bottom-up models	Stokes [STO05] Evaluation of the impact of the decentralized electricity production in the case of a low voltage network	Daily load curves measured at the household level Daily load curves measured at the appliance level Unitary load cycles constructed from measures Yearly consumptions per end-use coming from surveys	Daily load curves returned at a 1 min and 30 min time step for respectively a specified and « mean » household and possible aggregation at a community scale	Specific appliances, heating, DHW electricity	Empirical probability	On total load curves
	Paatero et Lund [PL06] To get accurate domestic electricity consumption data to identify the impact of DSM measures	Daily load curves measured at the household level Unitary load cycles constructed from measures Daily consumptions measured at the household level Consumptions per appliance in proportion with the dwelling consumption (data from studies without a precise temporality)	Daily load curve at a 60 min time step for a household and possible aggregation at a larger scale (thousands of households)	Specific appliances electricity	Empirical probability	On total load curves On household's yearly consumption
Hybrid model (statistical engineering)	Train et al. [THW85] To take advantage of the statistical and the engineering models to obtain end-uses load curves	Measured daily load curves available at the household level Appliances daily load curves from modelling	Daily load curves returned at a 60 min time step per appliance and for a specified household (or a group of households)	Specific appliances, heating, cooling, DHW electricity	Statistical	On total load curves

TABLEAU 1.3 – Reading grid (second part) of the studied load curve reconstitution models

## 1.4.2 Focused comparisons

### 1.4.2.1 Modelling the diversity

As we previously said, estimating the electric power demand for the domestic sector implies the modelling of the diversity. The above descriptions of the methodologies show that various strategies have been developed by the authors to represent it. Some of the models only « catch » the diversity directly from the input data without clarifying it. Other models generate it totally or partially thanks to more or less elaborated random processes (probability distribution functions, Monte Carlo extractions, Markov chains...) or statistical approaches (correction factors).

When only considering the models that represent the diversity, Walker and Pokoski's work is the to-our-knowledge first one which separately takes into account the presence at home of the households' members on the one hand (necessary but not sufficient condition to reconstruct the domestic load curve) and the proclivity for carrying out certain domestic activities (or using appliance(s)) on the other. On top of that and like the models of Capasso et al. [CGLP94], Widén et al. [WLV<sup>+</sup>09], Widén and Wäckelgård [WW09] and Richardson et al. [RTIC10], each household's member is characterized such a way. This constitutes a real modelling refinement.

However, in order to feed models that follow such an approach, precise information have to be available. Among other data, the authors made use of time of use surveys which are only representative of a selected households sample at a certain time. In order to ensure the representativeness of the models using TOU data, these must be periodically updated which supposes permanent and so expensive studies.

In some models, a consumer mean behaviour has been defined and applied to the simulated number of dwellings. Here is open the question of the likelihood of this construction because a mean consumer only has sense with regard to one of its characteristics or some of them. As the behaviour is influenced by too many parameters, defining a realistic mean one does not seem relevant.

For energy purposes, building up a mean consumer is viable because it is possible to calculate the mean consumption of a consumer sample over a certain duration : the mean consumer is the one who uses the calculated average energy consumption.

In terms of power demand, a mean consumer has no sense ; because of the diversity in a customers sample, that among other things comes from behavioural and technical influences, there is every chance that nobody has the same load curve as the After Diversity Mean Load Curve (ADMLC). It is calculated by addition of each individual load curve and then division by the number of aggregated consumers. However what can be defined with regard to the considered sample is the marginal consumer that is to say the consumer that will be added to the original sample. Its most likely representative load curve would be the ADMLC.

Another possibility to model the diversity through the human behaviour is to calculate from input data transition probabilities from a domestic activity to another one, and this, for all activities. When integrating this within a Markov Chain, a random influence completes the variety that is already contained in the input data. However the link between the activities and the use of appliances remains to be set up what is not trivial.

To our opinion, it is possible to set a framework for the modelling of the diversity that can be less dependent on the input data. With the help of a sufficient knowledge of the simulated housing stock, logic rules and constraints can be defined and applied in order to generate diversity.

### 1.4.2.2 Allocation of appliances functioning cycles within a simulated day

Different approaches exist in order to construct scenarios that define the use of appliances at the daily level.

The first one is to calculate an end-use average daily or weekly consumption from annual consumption figures and to launch a corresponding number of appliances cycles according to the unitary consumption of each end-use. However, this method can lead to an unrealistic daily usage of each appliance as for instance a  $\frac{3}{4}$  washing-machine cycle to launch each day.

The second one is to make use of the daily end-uses (or domestic activities) probability distribution functions constructed thanks to time of use data. In order to reach the energy consumption target, the number of end-uses starts must be controlled. This way is limited in the sense that in the better case it reproduces one measured reality and is representative of it. Changes in the end-use behaviour can not be simulated.

To our mind, we think that a method based on the knowledge of appliances use frequencies at different time scales can be more realistic than the previous ones. It will rely on the definition of types of day and households. When considering the week as the modelling framework, the sequential distribution of the launching of end-uses cycles for each day will take into account these influences but also the household's availability at home duration and the number of end-uses cycles already started in the considered week. On top of that, it would be necessary to constitute some links between appliances in order to improve the realism of simulations. For instance, the use of tumble-dryer is connected to the use of washing-machine. Moreover, it will be required to construct an end-uses ranking for each simulated day so as to launch them according to a priority order.

### 1.4.3 Positioning the models on a pseudo 3D plot

As we said in the introduction, we decided to represent on a chart the positioning of the models. In order to facilitate the reading of the representation, we restricted the number of regarded criteria : thus we only considered the range of application of the model (so the simulated appliances), the way the model (re)constructs the diversity and finally its time resolution.

In figure 1.15, the models are represented by bubbles whose diameters are inversely proportional to their time resolution. That is to say that a big bubble corresponds to a model that provides load curves with a fine time accuracy.

In the north-east corner of figure 1.15, a lack of models is noticeable : to our knowledge, there is no model that simultaneously provides load curves for all domestic appliances (inclusive thermal ones), at a fine time scale and with an advanced methodology to represent the diversity.

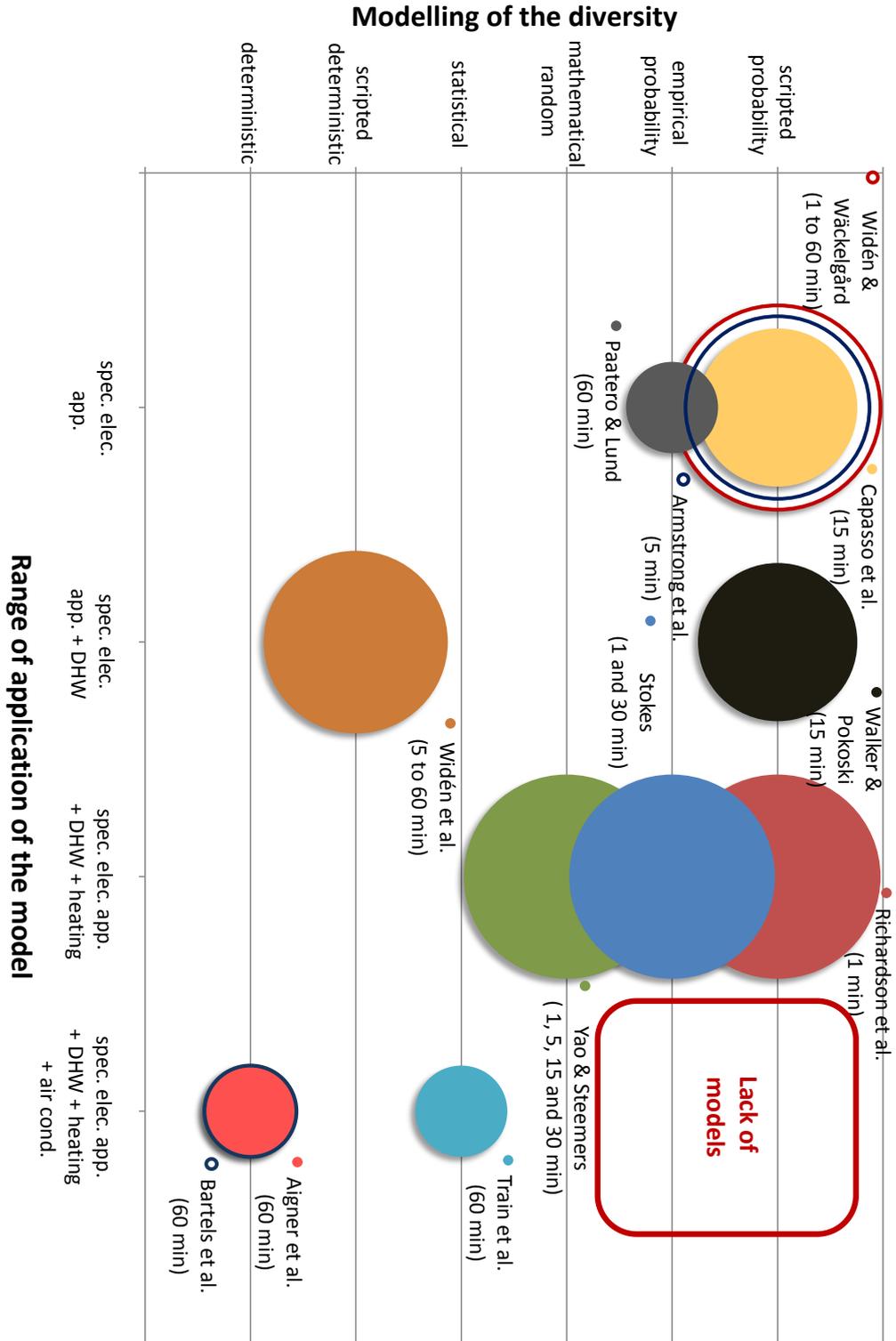


FIGURE 1.15 – Cross analysis diagram of the load curve models

## Conclusion

In this paper, we review twelve domestic load curve models. In the first part, we adopted a systematic analysis of each of them with focus on the finality of the model, its inputs/outputs, the modelled end-uses, the way diversity is taken into account and finally what kind of validation has been carried out. On top of that we built-up a synthetic bloc-diagram of each approach. In the second part of the paper, we considered them together carrying out a transverse analysis. A reading grid and a pseudo 3D plot was established at the end of the review.

This work has been voluntarily limited to a little number of models because of the willingness to analyse each of them in details. Other existing techniques for the load curve modelling, such as neural networks or agent-based method, have not been dealt with in this review. Further research has to be carried out to list, evaluate and analyse all possible methodologies for the reconstitution of the residential load curve. This purpose seems to us of great importance because of the undergoing radical transformation of the dwellings, domestic end-uses and human habits and behaviour.

## 1.5 Acknowledgment

This study could not have been carried out without the support of EDF R&D (Enerbat Department) and Mines ParisTech (Energy and Process Center).



## Chapitre 2

# Analyse fonctionnelle d'une méthode nouvelle de reconstitution de la courbe de charge électrique résidentielle

*« Je pense que le bond le plus important qui puisse être fait dans la plupart des domaines scientifiques est le passage de l'élément morphologique à l'élément fonctionnel. »*  
R. W. Gerard

### Introduction du chapitre

LE travail de réflexion que nous avons mené au sujet d'une méthode nouvelle de reconstitution de la courbe de charge a pris la forme d'une analyse fonctionnelle. Nos réflexions itératives ont mené à une série d'idées, de concepts, d'opérations de calcul et de traitements annexes de données. L'ensemble a été organisé suivant une arborescence fonctionnelle (également appelée arbre d'analyse fonctionnelle). De manière à faciliter les échanges autour de notre travail, nous avons choisi de traduire cette arborescence sous forme d'un cahier des charges fonctionnel détaillant notamment de manière individuelle chacune des fonctions de la méthode de reconstitution de courbe de charge. Pour aider le lecteur dans sa compréhension de notre méthode, nous avons choisi de maintenir dans le présent chapitre, la structure du cahier des charges que nous avons établi.

### Problèmes identifiés

La modélisation des usages électriques domestiques conduit tout d'abord à s'interroger sur la manière de représenter le cadre de l'étude. En effet, ce sont les équipements électriques qui seuls appellent de la puissance sur le réseau. Néanmoins,

- leur répartition sur une échelle géographique donnée,
- les interactions qu'ils peuvent avoir avec le bâti, les uns envers les autres mais aussi et surtout avec les utilisateurs,
- et les différents usages auxquels ils permettent de répondre

font qu'il est difficilement concevable de chercher à modéliser leur puissance électrique à chaque instant sans tenir compte des logements et de leurs occupants. Nous utiliserons le vocable de **parc habité\*** pour décrire l'ensemble formé par les bâtiments, les ménages et les équipements domestiques. Ces trois types de constituants du parc sont appelés des éléments. En outre, un bâtiment, un ménage et un jeu complet d'appareils (tous types d'usages confondus) constituent un **n-uplet\***.

La modélisation exacte du parc habité national conduirait à représenter un très grand nombre d'éléments le constituant. En fonction du niveau de raffinement de la modélisation, ce nombre peut à l'extrême, correspondre à l'effectif du parc habité puisque, du point de vue théorique, il n'y a pas deux n-uplets

identiques au sein du parc.

Pour rendre possible la modélisation et la simulation, il faut donc nécessairement réduire le nombre d'éléments constituant l'objet de l'étude. Cela revient à rechercher une simplification de la réalité en utilisant des objets en nombre limité qui sont représentatifs d'un effectif important d'éléments du parc habité c'est-à-dire des objets génériques. Dans la suite, nous introduisons et détaillons les concepts de *bâtiment-type\**, *ménage-type\** et *équipement-type\**. Indistinctement, ceux-ci sont appelés *éléments-types\**.

L'utilisation d'éléments-types peut s'avérer cependant délicate du point de vue combinatoire si les variables qui les décrivent sont nombreuses et/ou si leurs modalités sont multiples voire infinies.

Le travail d'analyse fonctionnelle a notamment eu pour objectif de sélectionner les paramètres les plus pertinents dans le cadre de la modélisation des appels de puissance électrique du secteur résidentiel.

La combinaison entre les éléments-types précédemment cités est très complexe puisqu'en réalité elle est liée à un grand nombre :

- de règles d'attribution déterministes ou probabilistes,
- d'éléments de logique,
- de contraintes,
- d'exclusions...

De surcroît, le parc habité ainsi modélisé doit évoluer au cours du temps de simulation puisque des éléments-types évoluent, intègrent le parc ou en disparaissent, de nouvelles règles entrent en jeu, d'autres deviennent caduques etc. Un deuxième objectif essentiel de l'analyse fonctionnelle a été de développer une méthode de construction du parc habité et de gestion de son évolution temporelle.

De plus, les usages thermiques soulèvent d'autres problèmes. En effet, les courbes de charge unitaires de chauffage, de climatisation et d'ECS sont très liées au climat et donc à l'échelle géographique considérée. L'utilisation de *chroniques météorologiques\** par *zone géographique\** s'avère nécessaire. À cette dernière échelle et a fortiori à celle du territoire, des influences supplémentaires affectent la courbe de charge agrégée des équipements. La densité de population, le taux d'équipement des ménages et le comportement humain, qui influe significativement sur la fréquence et l'intensité d'usage des appareils domestiques, peuvent varier significativement suivant la localisation.

Étant donné que le but du modèle est d'aboutir à des courbes de charge représentatives du secteur résidentiel, notamment à l'échelle nationale, plusieurs stratégies peuvent être adoptées du point de vue météorologique :

- définir ou choisir, à l'échelle de la France, des chroniques moyennes de température, d'ensoleillement, de nébulosité... Ces chroniques ne sont cependant pas représentatives de la réalité puisqu'artificiellement construites et elles lissent considérablement les chroniques effectives ;
- utiliser des données mesurées émanant des stations Météo-France. Cette solution pose le problème du nombre de simulations engendrées puisqu'il y a environ une station par département ;
- trouver un compromis entre les deux précédentes approches, i. e. segmenter le territoire suivant des zones climatiques qu'il faudra définir ou adopter.

Il nous semble cohérent de choisir un découpage pré-existant, en particulier si celui-ci fait office de référence dans le cadre d'études d'ordre thermique sur les bâtiments. C'est le cas par exemple des zones climatiques de la *réglementation thermique\** (RT\*) du bâtiment [MIN10]. Ces zones, actuellement – RT 2012 – au nombre de huit, sont représentatives des chroniques météorologiques qu'il est possible de rencontrer à l'échelle métropolitaine et sont valables à la fois pour l'été et l'hiver. Ce découpage semble pérenne aux prochaines évolutions du texte réglementaire.

En outre, les besoins de chauffage/climatisation sont directement dépendants des propriétés thermiques de l'enveloppe du bâti. Ces caractéristiques intrinsèques fixées à la construction des résidences peuvent évoluer dans le temps par le biais des **rénovations\***<sup>1</sup> d'une part, de la détérioration naturelle des matériaux d'autre part<sup>2</sup>. Les premières sont notamment motivées et/ou rendues nécessaires par les RT qui fixent des exigences en matière d'isolation des parois pour les habitations nouvellement construites et les logements rénovés. Là encore, afin d'être représentatif de la réalité, le modèle tiendra compte des améliorations du parc bâti par la voie des rénovations. À noter que les constructions neuves font en général appel à des matériaux présents dans la région d'implantation. Ceci laisse penser que des disparités régionales ou bien des habitudes de construction sont susceptibles d'exister lorsque le territoire français métropolitain est considéré dans son ensemble. La prise en compte de ces particularités régionales se heurte cependant à la disponibilité de données suffisantes aussi bien quantitativement que qualitativement.

À ces considérations techniques s'ajoutent l'aspect comportemental des occupants des logements. Tout d'abord, il s'agit du **confort thermique\***<sup>3</sup> perçu, c'est-à-dire des conditions d'ambiance intérieure (température, humidité, ventilation. . .) qui définissent la sensation de « bien-être » thermique dans un logement. Ces caractéristiques sont propres à chaque individu et varient notamment en fonction de l'âge et de l'état de forme à un instant considéré. Ainsi, pour une situation identique en matière de qualité de construction du bâti, de conditions météorologiques. . . deux personnes peuvent avoir un usage potentiellement très différent pour le chauffage ou la climatisation (par exemple la sensation de confort sera ressentie dès 18 °C dans un cas contre 23 °C dans l'autre).

D'autre part, les comportements d'investissement des ménages influencent grandement leurs besoins en énergie finale. C'est notamment le cas pour les usages thermiques qui d'une part contribuent largement aux consommations d'énergie dans ce secteur, comme l'indique la figure 2.1, et d'autre part, puisque ces usages peuvent facilement être modifiés suite à investissement<sup>4</sup>. L'effet de tels changements est direct : lors de travaux par exemple, le choix d'investissement conditionne la performance de l'équipement et éventuellement l'énergie utilisée dans le cas des usages thermiques ; le lien avec le besoin d'énergie est alors immédiat. C'est donc pour toutes ces raisons qu'il est important de prendre en considération l'investissement a minima pour ces équipements.

Enfin, des facteurs socio-culturels régionaux, notamment liés à des différences sensibles en matière de durée journalière d'ensoleillement et de flux solaire reçu et exprimé en kWh/(m<sup>2</sup>.an), biaisent encore les consommations théoriques pour le chauffage et la climatisation d'une part et ont une influence sur la consommation d'autres usages tel l'éclairage d'autre part. Ils doivent en conséquence être pris en compte dans le modèle.

Pour toutes ces raisons, une segmentation géographique du territoire national est requise pour mener à bien l'étude suivant les critères et les souhaits précédemment énumérés en matière de précision des simulations.

1. La rénovation est un processus d'amélioration d'une habitation (et/ou de ses éléments le constituant – systèmes énergétiques notamment) suite à la réalisation de travaux

2. Autrement dit, une habitation isolée thermiquement ne peut voir que décroître ses performances thermiques au fil du temps en l'absence de rénovations à cause de l'usure – tassement – des matériaux isolants, de l'apparition de ponts thermiques, d'une diminution de l'étanchéité de l'enveloppe. . .

3. Le confort thermique est une notion physiologique faisant référence aux perceptions sensorielles des occupants d'un logement principalement en matière d'humidité de l'air (hygrométrie), des **effets de paroi\*** et du **gradient thermique en température\***. Dans les climats tempérés, le confort thermique implique le chauffage des locaux durant l'hiver et en mi-saison ainsi que leur ventilation pour assurer un taux de renouvellement d'air suffisant. En outre, le rafraîchissement en période de forte chaleur et le conditionnement d'air sont deux opérations qui peuvent s'y ajouter selon la fonction des locaux considérés [AD04]

4. À ce sujet, un modèle de choix d'investissement des systèmes de chauffage a été proposé par Jaccard et Dennis dans [JD06] suite à l'enquête qu'ils avaient réalisée auprès d'un peu plus de 600 ménages canadiens

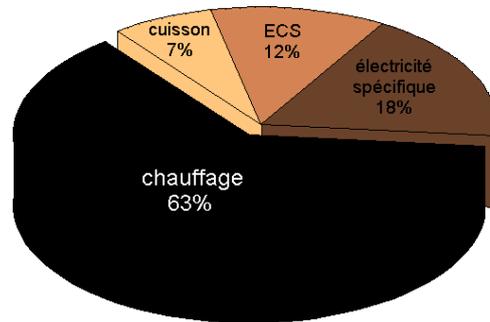


FIGURE 2.1 – Fractions par usage de la consommation énergétique finale (toutes énergies) du secteur résidentiel français (411,4 TWh) en 2010 à climat normal [CER10]

Pour assurer la cohérence du modèle et se donner les moyens d'une représentativité la plus convenable possible, il faudra régionaliser le parc habité dans sa globalité c'est-à-dire prendre en compte les éventuelles disparités locales de construction, les différences entre les ménages et les proportions entre ménages-types, les variabilités en matière de comportement et éventuellement une régionalisation des équipements-types par usage.

## Objectifs

Les objectifs de ce travail visent à :

1. disposer d'un modèle prototype (vraisemblablement industrialisé par la suite) ouvert et évolutif ;
2. valider les choix méthodologiques et les scénarios d'évolution :
  - méthode de foisonnement ;
  - scénarisation des évolutions temporelles ;
  - cohérence de ces scénarios ;
  - justesse de l'influence des paramètres par étude de sensibilité (bâti, équipement, comportement, météorologie).

Pour assurer l'évolutivité du modèle, nous pensons qu'une des meilleures façons d'y aboutir est d'opter pour une approche modulaire. Ainsi, l'intégration de nouvelles règles affectant les courbes de charge conduira à l'élaboration d'un nouveau module dans l'architecture du modèle. Celui-ci sera amené à communiquer et interagir sur d'autres parties du système. Une des nécessités de cette structure est d'anticiper et d'assurer la possibilité de communication entre des modules non encore existants. Une standardisation des trames d'informations sera très probablement indispensable.

## Les contraintes

En premier lieu, le modèle doit être compatible avec les informations disponibles et les bases de données dans lesquelles elles seront stockées. De plus, il doit être capable de s'interfacer avec les modèles thermiques disponibles (plus particulièrement l'environnement informatique avec lequel ils ont été développés) et être adapté au travail en batch (simulations itératives). De la même manière, le logiciel devra être capable de restituer les résultats sous différents formats et calculer des indicateurs caractéristiques

pertinents dans le cadre de la valorisation des courbes de charge issues des simulations.

Nous avons volontairement retiré de notre discours toute discussion autour des contraintes purement logicielles (flux de données, entrées/sorties...) qui dépassent largement le travail d'analyse fonctionnelle. Nous partons du principe que le modèle, bien que modulaire et évolutif, pourra s'interfacer convenablement avec Dymola et que l'élaboration de fichiers scripts (programmes informatiques de petite taille) est une possibilité permanente du modèle.

## Plan du chapitre

Le présent chapitre est divisé en quatre parties, chacune d'elle décrit une des fonctions principales – F 1 à F 4 – de la méthode de reconstitution de courbe de charge que nous avons élaborée. Nous précisons au passage que l'ordre de présentation choisi des fonctions n'est pas celui qui serait « naturellement » attendu (i. e. l'ordre alphanumérique : F 1 - F 2 - F 3 - F 4), ordre qui correspond davantage aux étapes de la simulation d'un scénario. En revanche, nous avons décidé de présenter les fonctions de manière à faciliter la compréhension de la méthode proposée. De plus, il s'agit également de mettre en lumière les fonctions les plus importantes (F 2 et F 3) de notre méthode de reconstitution des appels de puissance. Dans un premier temps, nous détaillons dans la partie 2.1 la fonction F 2 qui permet de fixer le cadre de travail (modélisation du parc habité à un instant donné) et qui introduit les concepts majeurs nécessaires à la reconstitution des appels de puissance électrique.

Comme nous l'avons souligné au préalable, la fonction F 2 « n'établit le parc qu'à un instant donné ». C'est pourquoi, nous avons mis en place la fonction F 1 que nous abordons dans la partie 2.2. Celle-ci permet de définir un cadre dynamique à la modélisation du parc notamment en matière de ruptures et de scénarios d'évolution entre les événements identifiés.

Une fois le parc habité construit à un instant donné avec une possible dynamique temporelle associée, tous les éléments requis pour le calcul des courbes de charge électriques correspondantes sont en place. Les calculs s'opèrent au moyen de la fonction F 3 que nous abordons au sein de la partie 2.3. Cette fonction constitue le cœur opérationnel de la méthode que nous avons définie.

Enfin, le calcul des courbes de charge ne constitue pas le point final de l'exercice de reconstitution des appels de puissance. En effet, les résultats doivent être restitués en fonction des besoins du modélisateur et des spécificités du scénario. L'obtention d'indicateurs caractéristiques peut en outre s'avérer intéressante. Ce rôle est attribué à la fonction F 4 que nous détaillons au sein de la partie 2.4. La figure 2.2 représente schématiquement l'agencement des quatre fonctions principales de notre méthode conceptuelle.

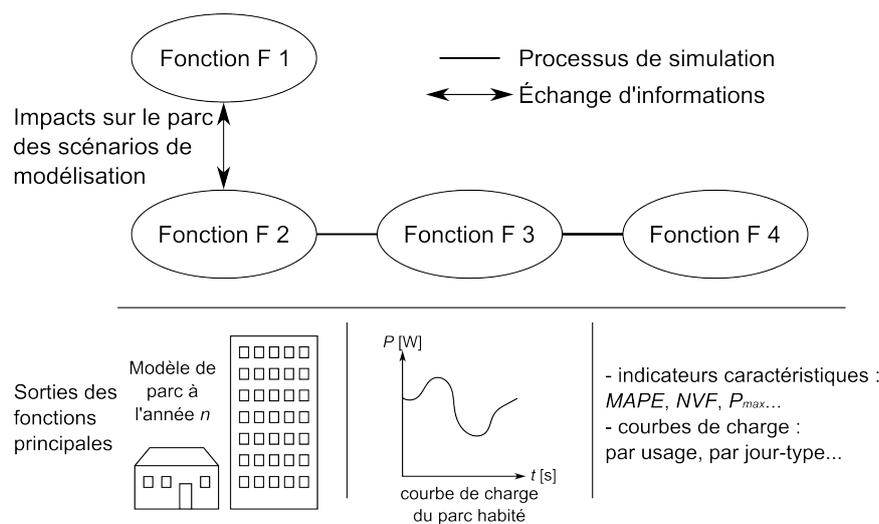


FIGURE 2.2 – Agencement des fonctions F 1 à F 4 de notre méthode de reconstitution de courbes de charge

## 2.1 Étude de la fonction « Modéliser le parc habité du territoire étudié (F 2) »

### Finalité

Le but de cette fonction est de représenter l'état d'un parc habité à un instant donné. Dans cette optique, le travail d'analyse fonctionnelle a mis en lumière trois objets principaux (définis ci-après) pour reconstituer les appels de puissance d'un parc résidentiel. Nous verrons au sein du paragraphe 2.2 que même si nous ne considérons ici que la situation d'un parc à un moment précis, celle-ci est représentative de l'état de ce parc entre deux **événements-clés**\* notés *ec*. Ce concept désigne une

*rupture dans le parc habité modélisé qui peut affecter n'importe lequel (lesquels) de ses constituants (éléments-types) que ce soit directement ou non.*

De plus, les événements-clés définissent et délimitent des **périodes**\* de simulation.

*Au sein d'une simulation de plusieurs années, une période est formée par toutes les années comprises entre deux événements-clés consécutifs.*

Les deux notions introduites ci-dessus sont schématiquement explicitées par la figure 2.3.

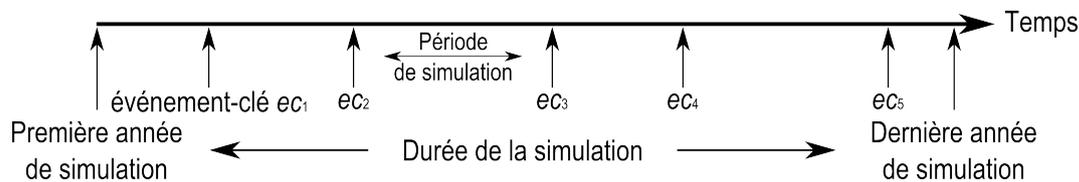


FIGURE 2.3 – Illustration des événements-clés et des périodes de simulation

Comme nous l'avons mentionné précédemment, un parc habité est constitué :

- de bâtiments-types

*habitation dont les caractéristiques typologiques, dimensionnelles, thermiques et/ou techniques sont représentatives d'un segment significatif du parc de logements considéré ;*

- de ménages-types

*foyer aux caractéristiques démographiques, économiques, sociales et comportementales typique d'un ensemble significatif de ménages de la population étudiée ;*

- et d'équipements-types

*appareil domestique dont les caractéristiques techniques et/ou de fonctionnement sont représentatives d'une famille d'équipements du marché à une période donnée.*

Dans la suite de ce document, nous désignerons par élément-type n'importe lequel des trois constituants du parc habité.

La description de la manière dont sont formés les éléments-types est effectuée dans cette partie. La figure 2.4 représente les flux d'informations entrant et sortant de F 2.

Comme le suggère la figure 2.4, la description des éléments-types n'est pas suffisante pour modéliser un ensemble de logements résidentiels en vue de leur simulation. En effet, il est ici aussi question de l'association de ces éléments dans le but de modéliser un parc habité hétérogène au sein d'un territoire sélectionné. Pour cela, la méthode de reconstitution de courbes de charge dispose d'une variété d'éléments-types stockés sous forme de bibliothèques. Grâce à des règles d'association, les bâtiments-types, les ménages-types et les équipements-types sont regroupés entre eux. Nous avons choisi de considérer les concepts suivants :

- **couple structurant**\* : l'association d'un bâtiment-type et d'un ménage-type ;
- **poids de représentativité**\* : la proportion en matière d'effectif d'un couple structurant du parc habité par rapport à l'ensemble des couples structurants ;

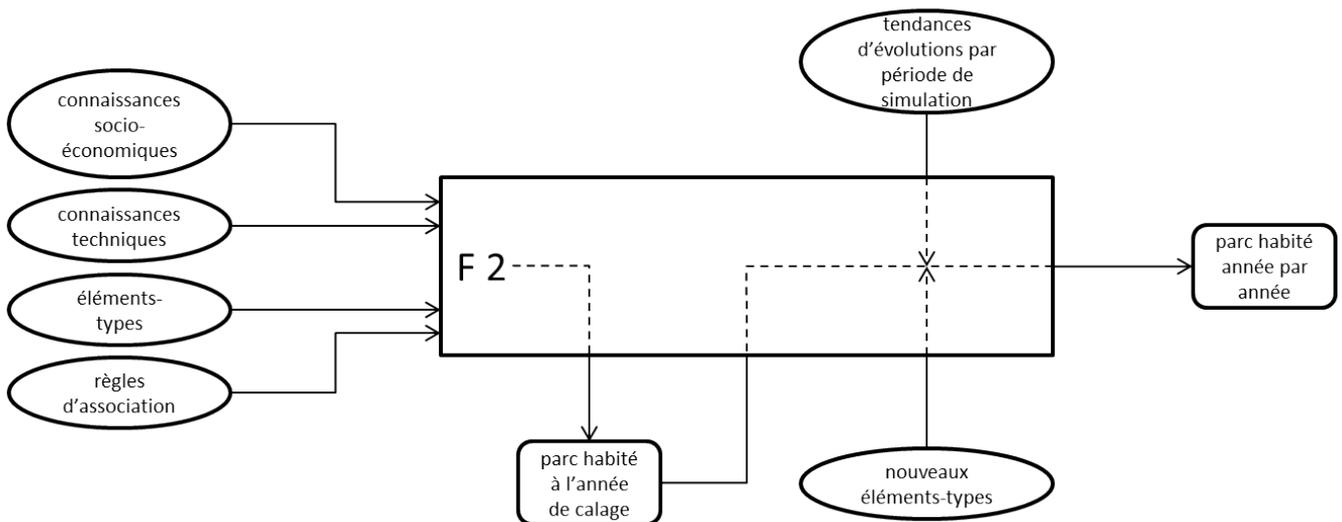


FIGURE 2.4 – Schéma d'entrées/sorties de la fonction F 2

- **triplet\*** : un couple structurant doté d'un jeu d'**équipements-types fixes\*** (cf. page 65) ;
- **n-uplet** : un triplet pourvu d'un jeu d'**équipements-types nomades\*** (cf. page 65).

La proportion numérique d'un n-uplet en matière d'individus par rapport à l'ensemble de la population est appelée **poids dans l'effectif\*** du n-uplet. Les n-uplets affectés de leur poids dans l'effectif constituent une base de données : le parc habité année par année.

*Le parc habité est constitué d'un ensemble de n-uplets affectés de leur poids dans l'effectif.*

Le mécanisme de reconstitution du parc habité est explicité au travers de la figure 2.5.

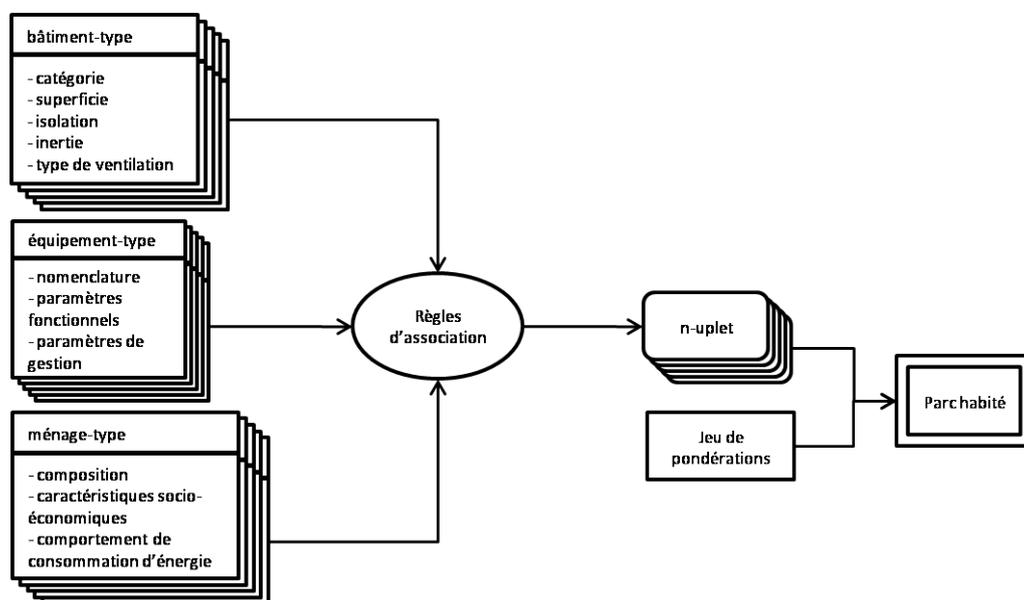


FIGURE 2.5 – Mécanisme d'obtention des n-uplets du parc habité

Dans cette section, nous détaillons également comment sont associés les bâtiments-types, les ménages-types et les équipements-types (fixes et nomades) ainsi que la façon dynamique avec laquelle l'évolution du parc est gérée au sein d'un **territoire\***. Par territoire, nous désignons la maille géographique de la modélisation. Il peut soit s'agir d'une seule zone géographique, d'un regroupement de zones géographiques ou bien encore du territoire national complet.

Par zone géographique notée  $Z$ , nous entendons un ensemble contigu de départements qui présentent des similitudes en matière de chroniques météorologiques. Cette segmentation permet en outre de caractériser chaque zone géographique en matière d'habitat régional, de densité, de nature et de modes de vie de la population.

La figure 2.6 représente un territoire constitué de quatre zones géographiques (notées Z1 à Z4).

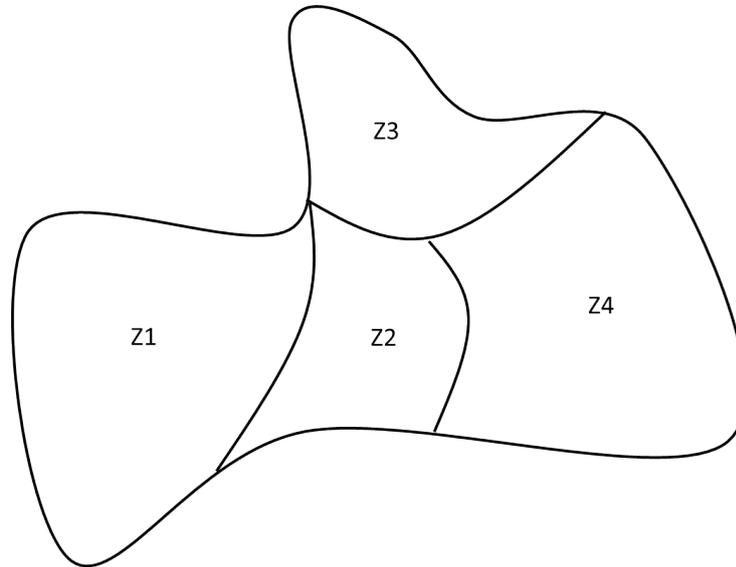


FIGURE 2.6 – Représentation du découpage d'un territoire en zones géographiques

## 2.1.1 Définir les éléments-types du parc habité (F 2.1)

### 2.1.1.1 Définir des bâtiments-types (F 2.1.1)

L'utilisation d'un certain nombre de bâtiments-types prédéfinis constitue une manière d'obtenir une représentation satisfaisante du parc observé. L'intérêt d'utiliser des bâtiments-types est de réaliser des simulations thermiques sur un nombre restreint de logements par rapport à l'ensemble des habitations du parc considéré (typiquement le parc national français de logements qui compte près de 30 millions de résidences principales). Ainsi, pour assurer une bonne représentation du parc en utilisant un nombre réduit d'éléments, il est nécessaire de sélectionner les caractéristiques des bâtiments étant à la fois les plus pertinentes du point de vue de l'appel de puissance des usages thermiques d'une part, et qui peuvent être renseignées d'autre part, au moyen de la connaissance sectorielle disponible. De ce fait, pour créer différents bâtiments-types, nous avons choisi de considérer les propriétés du bâti suivantes [CER03] :

- la catégorie (ou typologie) du logement : une habitation modélisée peut être de type **Maison Individuelle (MI)\*** ou **Logement Collectif (LC)\*** ;
- la superficie du logement : plusieurs valeurs discrètes sont choisies et affectées conformément aux données d'expertise sectorielle. A priori cette grandeur est principalement fonction de la localisation du logement (zone géographique correspondante), de son année de construction et de son énergie de chauffage ;
- le coefficient de transmission surfacique<sup>5</sup>  $U_{Bât}$  [W/(m<sup>2</sup>.K)] qui est vraisemblablement fonction de la localisation du logement et de son année de construction et/ou de son année de rénovation. Plusieurs valeurs discrètes d' $U_{Bât}$  sont retenues, valeurs qui peuvent être celles données par les RT. Chaque valeur de  $U_{Bât}$  est considérée comme caractéristique d'une classe de **performance thermique d'enveloppe\*** ;

5. La surface considérée est celle des parois du logement en contact avec l'extérieur

- l’inertie du bâtiment qui correspond à sa capacité à stocker et à restituer de l’énergie sous forme de chaleur (ou de fraîcheur). Elle est par hypothèse dépendante de la localisation du logement et de son année de construction. L’inertie diminue en général avec la rénovation de l’habitat puisque celle-ci se traduit régulièrement par la pose d’isolant à l’intérieur du logement. Comme pour le  $U_{Bât}$ , nous sélectionnons plusieurs valeurs discrètes d’inertie. Pour cela, il faut tout d’abord connaître l’étendue de variation de ce paramètre au sein du parc résidentiel étudié, corrélérer les valeurs d’inertie aux constructions<sup>6</sup> et choisir suffisamment de valeurs discrètes représentant cette propriété physique du bâti. Chaque couple ( $U_{Bât}$ , inertie) est considéré comme représentatif d’une classe de **performance énergétique du bâtiment**\* ;
- le type de ventilation variant selon toute vraisemblance en fonction de la localisation<sup>7</sup> du bâtiment-type et de son année de construction et/ou de rénovation. À ce stade, nous distinguons quatre modalités : ventilation naturelle, ventilation mécanique simple flux (extraction sans récupération de chaleur), puits canadien et ventilation double flux.

La catégorie de logement conditionne en partie la surface habitable et le type d’appareils de chauffage, de climatisation et d’eau chaude sanitaire qu’il est possible d’y installer. En outre, l’allocation des équipements-types nomades est influencée par la typologie de l’habitation considérée<sup>8</sup>.

Ces associations préférentielles sont gérées par le modèle de reconstitution de courbe de charge

De plus, un des principaux enjeux de la représentation d’un parc habité est la constitution de bâtiments-types aux performances énergétiques (caractéristiques thermiques d’enveloppe et inertie de la construction) relativement disparates<sup>9</sup>, et ce, pour représenter au mieux le parc actuel et à venir. En plus des déperditions thermiques caractérisées par le  $U_{Bât}$ , l’inertie de la construction influe directement sur la dynamique de la température intérieure d’une habitation suite à l’évolution de la température du milieu extérieur ou intérieur. Le déclenchement des systèmes de chauffage et/ou de climatisation est entre autres fonction de cette caractéristique.

Enfin le type de ventilation conditionne une partie sensible des pertes thermiques d’un bâtiment par extraction d’air intérieur chaud<sup>10</sup>.

Pour résumer, le cœur de calcul thermique que nous avons à choisir doit simuler de manière dynamique les systèmes installés dans les bâtiments-types ainsi que leur gestion. Ces habitations « génériques » représentatives de la majorité des constructions du parc modélisé, possèdent toutes les mêmes entrées et les mêmes sorties du point de vue de la nature des variables manipulées. Les bâtiments-types se différencient au niveau des modalités prises par chacune des variables les décrivant et des associations permises ou non avec les autres éléments-types. Considérer des bâtiments-types pour nos simulations rend paramétrable les aspects de la diversité des constructions du parc modélisé que nous avons retenus. En outre, cela évite de se perdre dans une approche combinatoire – voire exhaustive – dont la mise en pratique est irréaliste.

### 2.1.1.2 Définir des ménages-types (F 2.1.2)

Comme pour les bâtiments-types, il est inconcevable de définir toutes les typologies de ménages pouvant exister au niveau national. De ce fait, nous utilisons à nouveau une notion simplificatrice de la réalité, à

6. Dans le pire des cas – connaissance sectorielle insuffisante de ce paramètre notamment en matière de différenciation de cette grandeur par zone géographique – nous pouvons procéder à une affectation statistique de cette grandeur

7. Régionalisation forte du puits canadien par exemple

8. Il est par exemple peu crédible qu’un ménage en LC dispose de deux réfrigérateurs alors qu’en MI, la probabilité d’une telle situation est bien supérieure

9. De la « passoire thermique » au **Bâtiment Basse Consommation (BBC)**\* en y faisant correspondre une valeur d’inertie adéquate

10. Air à la température de consigne de l’habitation

savoir les ménages-types. Parmi toutes les caractéristiques des ménages, nous assumons l'hypothèse que les variables suivantes sont les plus pertinentes en vue de la reconstitution des appels de puissance :

- la composition du ménage : il est nécessaire de renseigner trois paramètres pour aboutir à une connaissance convenable de la composition des foyers :
    - le nombre d'occupants ;
    - l'âge respectif de chacun d'eux ;
    - leur sexe.
- Pour limiter le nombre de combinaisons possibles à partir de ces trois déterminants et pour s'assurer une conformité avec des sources de données existantes, nous pourrions reprendre tout ou partie des éléments de la catégorisation de la population établie par l'INSEE [INS07, INS10b, INS12e] :
- intervalles d'âge utilisés dans les travaux de projection de population : 0-19 ans, 20-59 ans, 60-64 ans, 65-74 ans et 75 ans et plus ;
  - types de structure familiale : ménage d'une seule personne, ménage d'une seule famille, ménage complexe ;
  - ménages par taille du foyer, sexe et âge de la personne de référence : sept catégories d'âge en regard de six modalités d'effectif du ménage ;
- les caractéristiques socio-économiques du ménage. Il faut ici renseigner des informations concernant :
    - le revenu du foyer avec au moins trois modalités (haut, moyen, bas) et/ou
    - la **Catégorie SocioProfessionnelle (CSP)\*<sup>11</sup>** du chef de famille ou des adultes du foyer (deux parents actifs) ;
    - le **statut d'activité\*** du ménage considéré à un niveau agrégé<sup>12</sup> ou de chaque membre le constituant (en activité, retraité... ) ;
  - les comportements de consommation d'énergie par usage en dissociant ceux relatifs :
    - au chauffage et à la climatisation ;
    - à l'ECS ;
    - et aux équipements d'électricité spécifique.

Une distinction en un minimum de trois classes de consommateur (énergétiquement économe, indifférent ou moyen et énergétiquement dépensier) nous semble pertinente.

Les différents éléments présentés ci-avant ont pour objectif de recréer une diversité de comportements des ménages en particulier en ce qui concerne leurs habitudes de consommation énergétique. Celles-ci se matérialisent entre autres, par une fréquence, une intensité et une temporalité d'utilisation des appareils domestiques propres à chaque foyer (a fortiori à chaque individu). En ce qui concerne la dimension temporelle des besoins en électricité au sein d'un logement, nous introduisons les deux concepts suivants : **chronogramme probabilisé d'usage\*** et **scénario Time Of Use\*<sup>13</sup>**. Nous les aborderons plus en détails et nous pourrions constater tout leur intérêt lorsque nous traiterons du calcul des courbes de charge unitaires des équipements d'électricité spécifique au sein des logements (cf. paragraphe 2.3.2.1). Soulignons toutefois à ce niveau qu'en amont de ces concepts sont pris en considération les informations relatives aux scénarios de présence infra-journalière et annuelle des ménages à leur domicile.

À l'image des bâtiments-types, il est nécessaire de réduire autant que faire se peut le nombre de paramètres qui définissent les ménages-types. Toutefois ceux-ci doivent être suffisamment bien choisis de manière à ce qu'en sortie de l'élément ménage, nous puissions obtenir des indicateurs influençant, usage par usage, le comportement général de celui-ci vis-à-vis de sa consommation d'énergie.

11. Nous pourrions par exemple reprendre la segmentation agrégée de la « Nomenclature des professions et catégories socioprofessionnelles/PCS » de l'INSEE (voir [INS03] p. 11)

12. Dans ce cas, le statut d'activité du ménage peut être identifié à celui du chef de famille par exemple

13. Également appelé scénario d'instant d'usage

### 2.1.1.3 Définir des équipements-types (F 2.1.3)

Là encore, il est impensable de tenter de modéliser tous les types d'appareils puisqu'ils sont d'une extrême diversité. De ce fait, nous avons à nouveau recours à une voie générique à savoir la modélisation d'équipements-types.

La réflexion autour des usages thermiques nous a conduit à définir deux sous-catégories parmi les équipements-types :

- les équipements-types fixes<sup>14</sup>, car « fixés » au bâtiment, correspondent aux appareils de chauffage<sup>15</sup>, de climatisation et d'ECS. Ils sont intimement liés au bâtiment dans lequel ils sont installés ;
- les équipements-types nomades<sup>16</sup> qui peuvent, par exemple, accompagner les occupants d'un foyer lors d'un déménagement. Leur possession est davantage fonction des occupants d'un ménage.

Cette catégorisation révélera toute son importance lors de l'association des jeux d'équipements-types fixes et nomades pour chacun des couples structurants<sup>17</sup>.

#### Définir les équipements par usage

Dans ce paragraphe, nous présentons les différents usages et équipements dont la modélisation nous semble pertinente par rapport à l'objectif de la thèse. De manière à clarifier les termes, nous représentons au moyen de la figure 2.7 la chaîne générale partant d'un besoin à satisfaire et allant jusqu'à l'équipement-type correspondant (a) et la même chaîne sur un exemple (b).

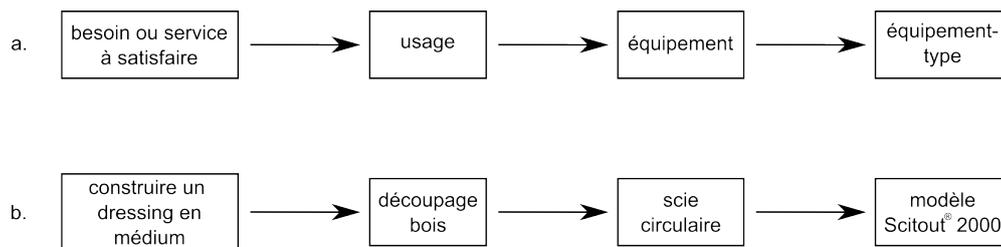


FIGURE 2.7 – Liens entre les concepts « usage », « équipement » et « équipement-type » dans le cas général (a) et sur un exemple (b)

Parmi l'ensemble des équipements domestiques du parc résidentiel français, seuls ceux qui sont pertinents vis-à-vis de la courbe de charge seront modélisés en tant qu'équipements-types.

Par pertinent, nous entendons que ces équipements présentent au moins une des caractéristiques suivantes :

- la durée de fonctionnement lors de l'utilisation de l'appareil est suffisamment grande par rapport au pas de temps de calcul des courbes de charge (a minima la durée de marche de l'équipement doit être supérieure à 10 min<sup>18</sup>) ;
- la puissance appelée en fonctionnement est importante (au-dessus d'un seuil fixé en fonction de l'étude menée) ;
- le **taux de pénétration**\* de cet équipement dans le parc est significatif à partir d'une année donnée (au-dessus d'un seuil fixé en fonction de l'étude menée).

La liste des équipements explicités dans la modélisation doit être revue tous les ans. Pour assurer la cohérence en énergie du modèle et ainsi ne pas faire fi des appareils non modélisés, nous ajoutons dans

14. Également appelés équipements immobiliers

15. Nous gardons en mémoire que cet usage peut être assuré par des équipements nomades, donc déplaçables de pièce en pièce, dans certaines situations

16. Ceux-ci sont également surnommés équipements mobiliers

17. Cf. méthode 3(d) de distribution des équipements-types décrite dans l'encadré de la page 78

18. 10 min coïncident avec le pas de temps de mesure des courbes de charge sur le réseau [ErD08]. Pour notre part, nous choisissons cette limite comme pas de temps de calcul maximal des courbes de charge d'électricité spécifique

la liste des usages considérés (voir le paragraphe <sup>19</sup> page 73) le concept de *solde fatal en énergie*\* <sup>20</sup>, noté *SFE*, qui est causé par les équipements non explicités. Comme nous le verrons lorsque nous détaillerons la fonction F 3.8 (cf. partie 2.3.8 p. 103), nous chercherons à attribuer une courbe de charge à ces appareils par l'intermédiaire d'un traitement spécifique.

Pour des raisons liées aux singularités des équipements domestiques, nous les regroupons sous forme d'usages et de familles d'usages qui sont listés ci-dessous :

- chauffage ;
- eau chaude sanitaire ;
- climatisation ;
- ventilation et auxiliaires de chauffage, de climatisation et d'ECS ;
- froid domestique ;
- préparation culinaire ;
- lavage ;
- éclairage ;
- informatique ;
- audiovisuel et télécommunications ;
- nettoyage et entretien ;
- appareils d'extérieur ;
- transport et mobilité ;
- nouvel usage ;
- équipements participant au solde fatal en énergie.

Chacun(e) de ces usages (famille d'usages) est décrit(e) dans les paragraphes qui suivent.

Le découpage par usage/famille d'usages proposé ici présente des similarités avec ceux de quelques modèles [CGLP94, YS05, STO05] que nous avons analysés au sein du chapitre 1. Toutefois dans les travaux de la littérature, la distinction de plusieurs équipements au sein de chaque famille d'usage n'est jamais aussi fine que la segmentation de notre modèle.

### *Chauffage*

Une installation de chauffage au sein d'un logement peut se présenter sous trois formes :

- *base unique*\* ;
- *double base*\* ;
- *chauffage d'appoint*\*.

La base unique indique qu'un seul dispositif de chauffage est présent dans une habitation. Ce système assure les besoins de chaleur quelles que soient les conditions extérieures. La base unique peut être électrique ou bien utiliser des ressources fossiles ou du bois. Dans le premier cas (base unique électrique), nous distinguons les équipements suivants :

- pompe à chaleur (air/air, air/eau, eau/eau) ;
- convecteur ;
- chauffage électrique à accumulation ;
- panneau rayonnant ;
- plancher chauffant électrique ;
- chaudière électrique, gaz, fioul ou bois.

Une étude plus approfondie sur les systèmes de chauffage domestique doit être conduite pour déterminer si d'autres systèmes s'avèrent employés à grande échelle.

19. « Équipements participant au solde fatal en énergie »

20. Il s'agit d'une consommation électrique non modélisée de manière explicite – pour plus de détails se reporter à la page 73

Dans le cas de la base fossile ou bois, nous devons simuler la présence d'auxiliaires de chauffage que nous détaillerons dans le paragraphe correspondant <sup>21</sup> page 68. Si cette base est sélectionnée, ces derniers éléments sont automatiquement intégrés au jeu d'équipements-types du n-uplet considéré.

Le cas de la double base correspond à une situation du logement dans laquelle un système seul ne peut répondre à l'ensemble des besoins du fait de la géométrie ou de l'agencement du logement. Nous retrouvons dans cette branche la même segmentation électrique/fossile et bois ainsi que les mêmes systèmes que nous avons présentés pour la base unique.

Enfin, un logement peut disposer d'un système de chauffage d'appoint si son système principal s'avère insuffisant à partir d'un certain seuil de température ou bien si la distribution de la chaleur au sein de l'habitation ne satisfait pas le confort thermique des occupants. Pour les solutions d'appoint, nous avons réduit le nombre des technologies possibles en ne conservant que :

- les convecteurs ;
- les panneaux rayonnants ;
- les systèmes utilisant une ressource fossile ou bois.

L'agencement des différents systèmes et des multiples configurations décrites ci-dessus suit l'arborescence issue de l'arbre d'analyse fonctionnelle présentée figure 2.8 <sup>22</sup>.

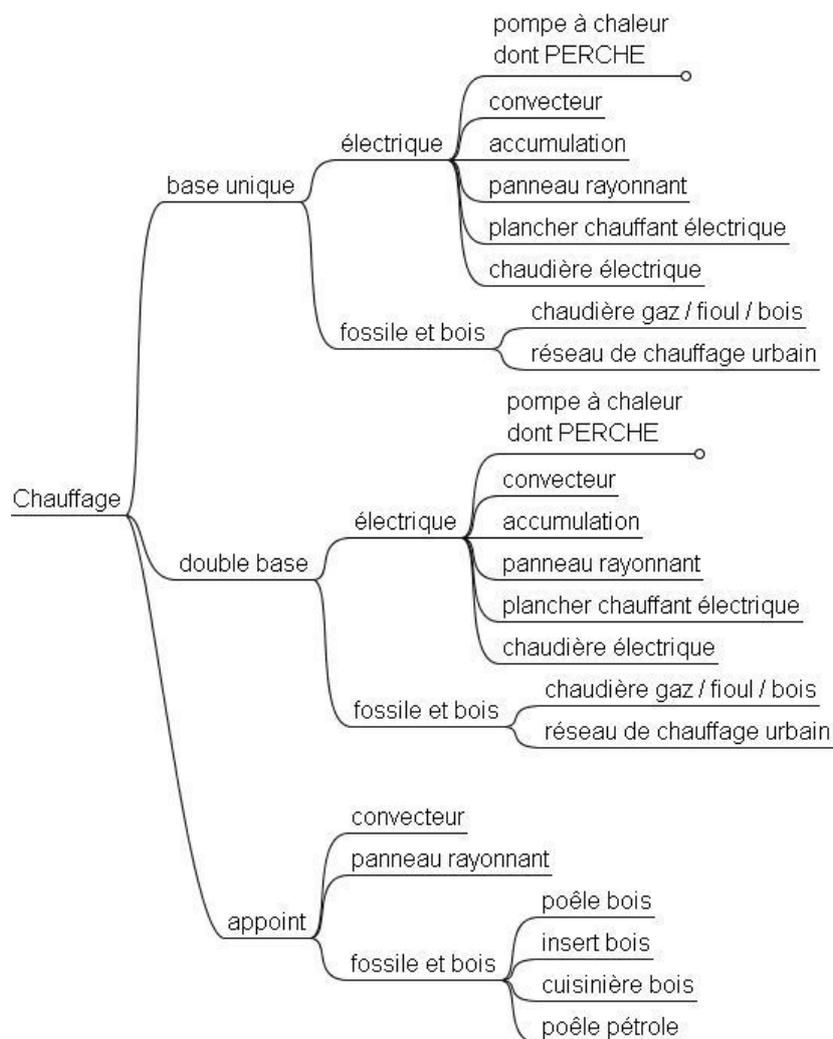


FIGURE 2.8 – Branche chauffage de l'arbre d'analyse fonctionnelle

21. « Ventilation et auxiliaires de chauffage, de climatisation et d'ECS »

22. Sur cette figure, PERCHE désigne le système de Pompe à chaleur En Relève de CHaudière Existante\*

*Eau chaude sanitaire*

Pour modéliser les appels de puissance relatifs au puisage d'ECS, nous n'avons besoin que de nous intéresser aux systèmes électriques (l'ECS issue de chaudière fossile ou bois n'implique qu'une consommation négligeable causée par leurs auxiliaires). À ce stade, nous retenons les systèmes suivants :

- ballon de stockage (avec la distinction des systèmes asservis en heures creuses de ceux qui ne le sont pas) ;
- chauffe-eau instantané ;
- chauffe-eau thermodynamique (100 % thermodynamique ou avec appoint électrique) : pompe à chaleur double service et chauffe-eau autonome asservis ou non ;
- **Chauffe-Eau Solaire Individuel (CESI)\*** à appoint électrique.

Les systèmes intégrant des énergies renouvelables comme le CESI entraînent une réduction des besoins énergétiques et de puissance des individus du n-uplet pour l'usage ECS, et ce, au pro rata de l'apport solaire disponible au sein de la zone géographique considérée. La saisonnalité est donc particulièrement remarquable sur ces systèmes. Ils sont en conséquence suppléés par un appoint qui peut être électrique ou fossile. Ce dernier est généralement dimensionné sans tenir compte du système fonctionnant à partir d'énergies renouvelables en supposant un apport nul de celui-ci au cours des périodes les plus rigoureuses.

*Climatisation*

Pour assurer les éventuels besoins de rafraîchissement d'un logement, des systèmes utilisant les énergies renouvelables (par exemple le puits canadien) peuvent être employés. En revanche, les besoins de climatisation lorsqu'ils existent, nécessitent bien souvent un système électrique. Pour cet usage, nous distinguons les trois équipements suivants :

- pompe à chaleur réversible ;
- climatiseur centralisé ;
- climatiseur d'appoint.

La majorité des pompes à chaleur air/air sont réversibles. Nous formulons l'hypothèse qu'un logement doté d'une telle machine pour son chauffage peut l'utiliser également pour assurer sa climatisation.

*Ventilation et auxiliaires de chauffage, de climatisation et d'ECS*

Pour ces usages, nous avons choisi d'explicitier les équipements suivants :

- **Ventilation Mécanique Contrôlée (VMC)\*** avec distinction entre extraction simple et extraction double flux ;
- pompe de circulation de chaudière et de système de production d'ECS<sup>23</sup> (par exemple pompe de circulation du fluide caloporteur dans les panneaux solaires thermiques) ;
- auxiliaires électroniques de chauffage/climatisation (exemples : boîtiers de commande, régulateurs).

En fonction de la ventilation sélectionnée au niveau du bâtiment-type, le modèle doit affecter à l'habitation l'appareil électrique correspondant. De même, comme nous l'avons souligné au sein du paragraphe « Chauffage » page 66, si un logement est équipé d'une installation de production thermique fonctionnant avec une ressource fossile ou du bois, la présence d'un circulateur de chaudière est configurée automatiquement par le modèle. Enfin, les systèmes de chauffage électrique sont souvent associés à des régulateurs et à des boîtiers permettant d'assurer leur gestion. Là encore, la sélection d'un système de chauffage faisant appel à ces dispositifs conduit à une affectation automatique de ces derniers au sein du n-uplet considéré.

*Froid domestique*

Pour l'usage froid domestique, nous distinguons les trois appareils suivants :

- réfrigérateur ;
- congélateur ;
- combiné réfrigérateur-congélateur.

23. Si la consommation électrique de ces circulateurs s'avère sensible, elle sera prise en compte dans les simulations

Le but de ces appareils est de maintenir, à tout instant et à une température de consigne fixée par l'utilisateur, une enceinte isolée thermiquement qui renferme des denrées alimentaires périssables. La consommation électrique de ces équipements est requise au niveau du système thermodynamique – essentiellement le(s) compresseur(s) – qui équipe ces appareils. Ce besoin énergétique varie notamment en fonction de la technologie<sup>24</sup> de l'équipement-type, des conditions d'utilisation (température de consigne, température de la pièce accueillant l'appareil, état de givrage. . .) et de l'intervention humaine (ouverture de porte, remplissage).

En matière de taux de pénétration des équipements de froid domestique, aujourd'hui [INS12f] la totalité des ménages français dispose pour ainsi dire, d'au moins un réfrigérateur comme l'indique la figure 2.9.

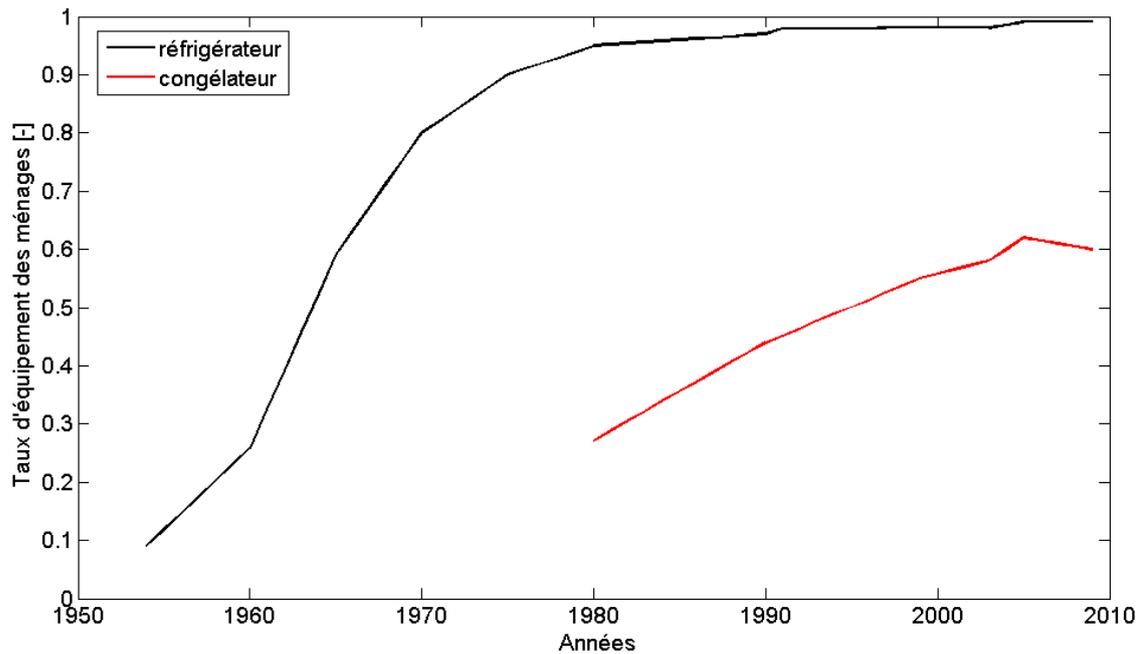


FIGURE 2.9 – Évolution des taux d'équipement des ménages français en réfrigérateur et congélateur entre 1954 et 2009 [INS10d]

Au cours de ces dernières années, l'amélioration de l'efficacité énergétique des appareils de froid domestique est notable : la campagne de mesures ADEME\*/EDF/ENERTECH 2008 [AEE08] a identifié une réduction de 40 % des consommations annuelles par logement du poste froid entre 1996 et 2008 et de 16 % entre 1999 et 2008. Cependant cette réduction est compensée par une augmentation du volume moyen des appareils qui constitue un effet rebond.

### Préparation culinaire

Dans cette famille d'usages, nous regroupons tous les équipements domestiques pouvant servir de près ou de loin à la préparation des repas. Dans une première approche, nous cherchons à modéliser les appareils suivants :

- four ;
- plaques de cuisson ;
- four à micro-ondes ;
- autres appareils de cuisson à résistance<sup>25</sup> (par exemple four à raclette, machine à pain, cuiseur-vapeur) ;
- hotte aspirante.

24. Pour les réfrigérateurs par exemple, le marché des appareils se répartit entre le froid statique, le froid brassé et le froid ventilé

25. L'augmentation tendancielle de ce type d'appareils doit être considérée avec raffinement

Ces équipements ont la particularité d'être utilisés de manière très périodique, essentiellement au moment de la préparation des repas. Conformément aux habitudes de vie des ménages, les heures de prise alimentaire dépendent du jour-type (cf. descriptif de la fonction 3.1.1 au sein de la partie 2.3.1.1) si bien qu'une périodicité hebdomadaire doit être modélisée pour une majorité des ménages-types constitués.

### Lavage

L'usage lavage est associé à trois équipements distincts :

- lave-vaisselle ;
- lave-linge ;
- sèche-linge.

Ces appareils sont utilisés plus ou moins régulièrement à l'échelle de la semaine en fonction des caractéristiques du ménage-type et en particulier suivant le nombre de personnes qui le constitue. Au niveau national, l'évolution dans le temps du taux d'équipement des ménages pour chacun de ces appareils est assez différente d'un équipement à l'autre comme le suggère la figure 2.10.

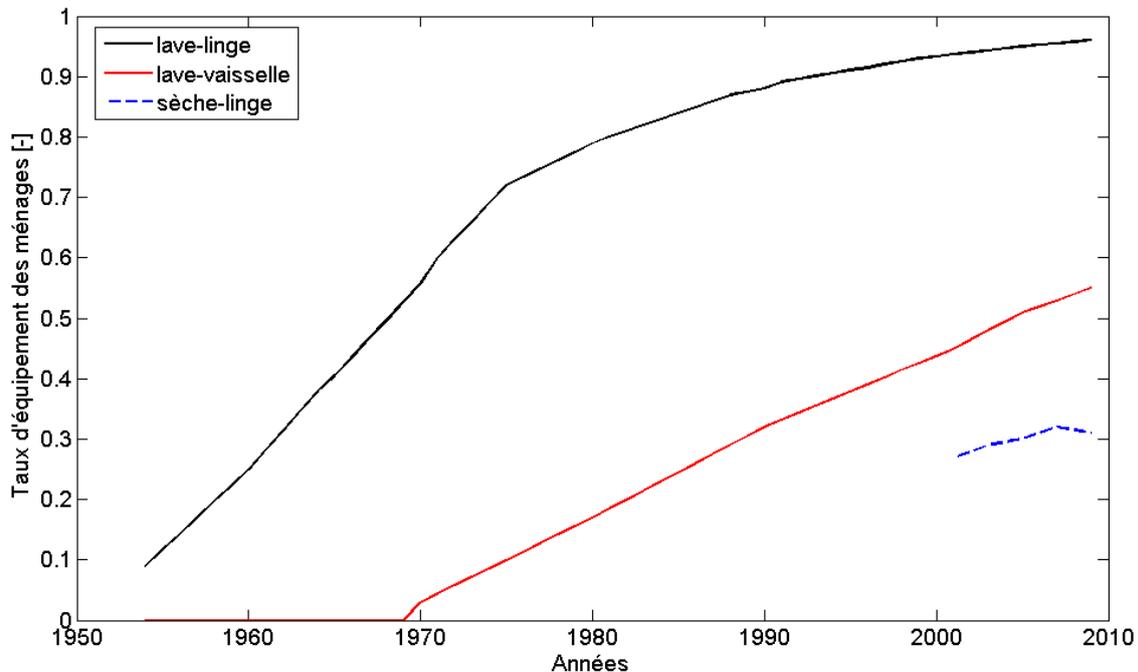


FIGURE 2.10 – Évolution des taux d'équipement des ménages français en usages blancs (lave-linge, lave-vaisselle et sèche-linge) entre 1954 et 2009 [INS10d]

Le lave-linge est un équipement qui arrive à saturation dans le parc national avec environ 95 % des ménages équipés [INS06, INS10c, INS11b, INS12f] : cette valeur indique que l'appareil est ressenti comme un bien nécessaire. Les deux autres équipements en revanche sont moins présents au sein des foyers et semblent davantage considérés comme des éléments de confort.

Soulignons également que dans notre modèle, nous n'avons pas décidé de modéliser le combiné lave-linge/sèche-linge car il ne présente pas de particularités notables de fonctionnement. La modélisation de cet appareil sera réalisée par la mise en marche séquentielle, cohérente et avec des règles logiques des équipements « traditionnels » machine à laver et sèche-linge. Concrètement, un ménage équipé d'un combiné peut être modélisé comme possédant les deux appareils aux fonctions uniques.

*Éclairage*

Les équipements que nous avons choisis de modéliser pour l'éclairage se différencient en matière de technologie :

- lampe à incandescence ;
- lampe basse consommation ;
- halogène ;
- néon ;
- **Diode Électro-Luminescente (DEL)**\*<sup>26</sup>.

L'éclairage est un usage très lié à la présence des occupants dans leur logement ainsi qu'à leur nombre. Par exemple, quatre personnes d'un même ménage se trouvant simultanément dans leur logement vont a priori solliciter davantage de points lumineux, malgré la mutualisation de l'espace (c'est-à-dire le fait que plusieurs occupants partagent la même pièce et y effectuent des usages identiques ou non), qu'un couple ou une personne seule. Deux autres influences sont particulièrement sensibles sur l'éclairage :

- la localisation du logement qui se traduit par des niveaux et des heures d'ensoleillement notablement différents ;
- la saisonnalité qui conditionne à la fois les heures de lever et de coucher du soleil et les intensités lumineuses (flux solaire) perçues par un logement.

*Informatique*

Les équipements informatiques que nous avons jugés pertinents du point de vue des appels de puissance sont les appareils suivants :

- unité centrale ;
- moniteur ;
- imprimante<sup>27</sup> ;
- ordinateur portable ;
- box internet.

Les équipements du poste informatique prennent aujourd'hui une place importante dans la vie quotidienne des ménages français qu'ils soient utilisés dans un but professionnel, pour les loisirs, pour s'équiper (achats/ventes sur internet), s'informer ou communiquer [INS11a]. La figure 2.11 témoigne de la croissance très rapide du taux d'équipement en micro-ordinateur des ménages français au cours des décennies 1990 et 2000.

Cette augmentation du taux d'équipement s'accompagne a fortiori d'une hausse de la consommation d'énergie de ces appareils à l'échelle nationale [ENE08].

*Audiovisuel et télécommunications*

Dans cette famille se retrouve une série d'équipements relatifs aux loisirs d'une part et à la communication d'autre part :

- poste de télévision avec distinction entre les téléviseurs à tubes cathodiques **CRT (Cathode Ray Tube)**\*, **LCD (Liquid Cristal Display)**\*, plasma et LED ;
- lecteur DVD/magnétoscope ;
- chaîne hi-fi ;
- appareils audio (enceintes, amplificateur, home-cinéma, charge de lecteurs audio portables) ;
- téléphone fixe/téléphone portable (charge) ;
- console de jeu.

Tout comme pour le réfrigérateur et le lave-linge, la quasi-totalité des ménages de l'Hexagone dispose d'un téléviseur [INS12f] comme nous l'indique la figure 2.12.

Au cours de ces dernières années, les technologies disponibles sur le marché ont significativement évolué. Malgré l'amélioration de l'efficacité énergétique des écrans, la consommation de ces appareils a sensible-

26. L'acronyme anglo-saxon **LED (Light-Emitting Diode)**\* est plus communément employé

27. Y compris appareils multi-fonctions imprimante-scanner-photocopieur

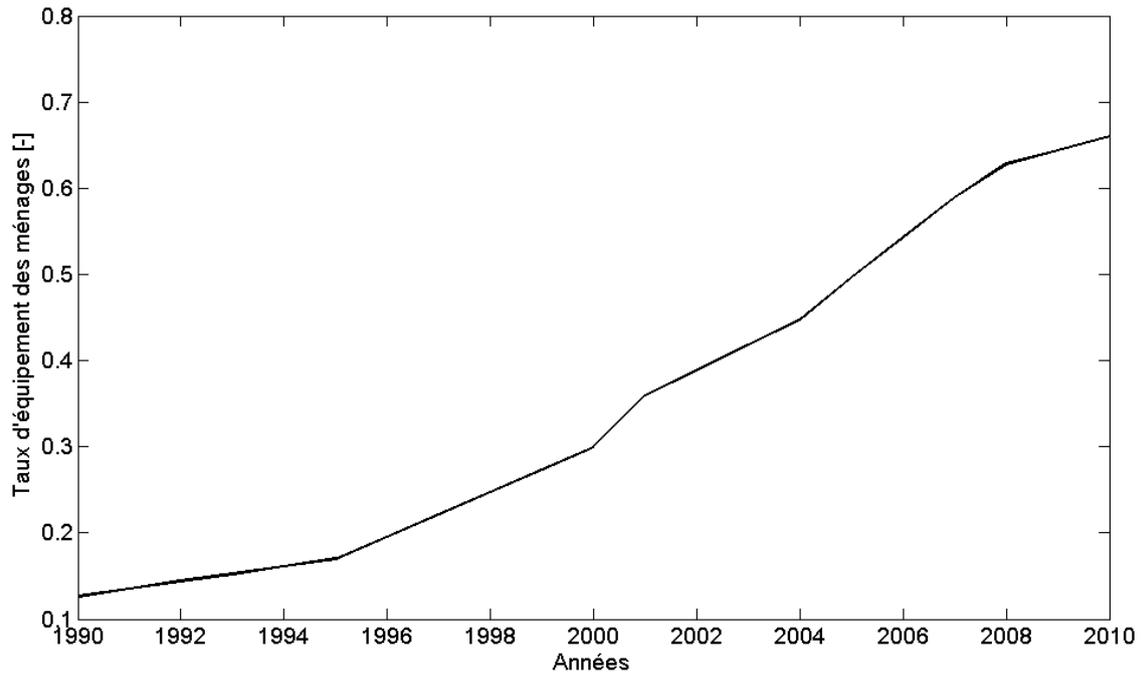


FIGURE 2.11 – Évolution du taux d'équipement des ménages français en micro-ordinateur entre 1990 et 2010 [INS10d]

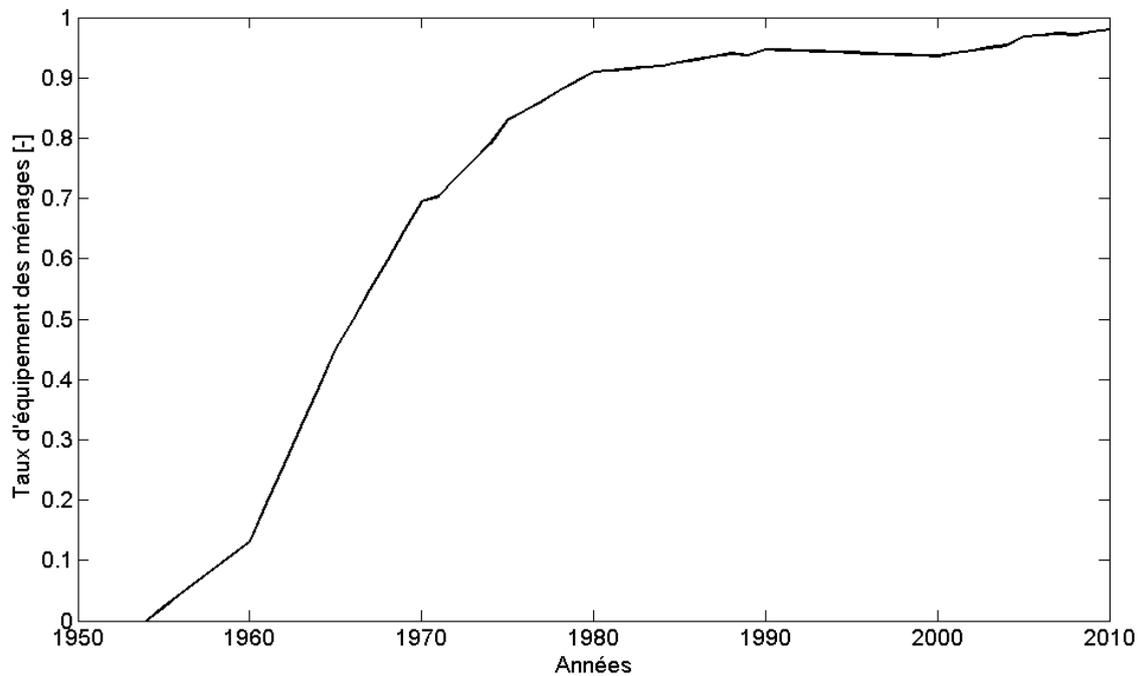


FIGURE 2.12 – Évolution du taux d'équipement des ménages français en télévision entre 1954 et 2010 [INS10d]

ment augmenté sur la période 1994-2008 [ENE08] du fait de l'agrandissement notable de la taille des diagonales des appareils vendus dans le commerce.

À l'image du réfrigérateur, l'usage de la télévision souffre d'un effet rebond qui gomme totalement l'amélioration de l'efficacité énergétique qu'a connu cet appareil au cours des dernières années [ECE10].

#### *Nettoyage et entretien*

Dans cette catégorie, nous regroupons les appareils utilisés pour l'entretien intérieur des habitations et le bricolage :

- aspirateur ;
- appareil de nettoyage à vapeur ;
- fer à repasser ;
- appareils électro-portatifs et à poste fixe (notamment ceux destinés au bricolage).

Bien qu'ils ne soient généralement utilisés qu'occasionnellement au sein d'un logement, ces appareils sont à considérer dans la mesure où ils sont susceptibles d'appeler de fortes puissances électriques lors de leur fonctionnement.

#### *Appareils d'extérieur*

Ces équipements sont exclusivement présents dans les MI puisqu'ils supposent une utilisation inutile ou impossible en intérieur :

- appareils de jardinage (tondeuse à gazon, taille-haie) ;
- appareils de nettoyage haute pression ;
- pompe de piscine.

#### *Transport et mobilité*

Cet usage est encore assez peu répandu mais il est dit promis à un bel avenir. Les équipements retenus sont les suivants :

- véhicule électrique (VE) et hybride rechargeable (VHR) ;
- deux roues électrique.

#### *Nouvel usage*

Comme nous l'avons souligné dans l'introduction de ce chapitre, nous souhaitons aboutir à un modèle évolutif. Ainsi, il est impératif de pouvoir intégrer au parc habité modélisé de nouveaux usages, équipements et équipements-types. Le fait de prévoir cette modification conduit concrètement à pouvoir faire évoluer la forme de notre base de données (l'agrandir ou la réduire en fonction de l'arrivée ou de la disparition d'un usage et/ou d'un équipement). L'avènement d'un nouvel usage dans le parc conduit à la création d'un (ou plusieurs) équipement(s) non encore existant(s).

L'introduction d'un nouvel usage correspond à un événement-clé qui impacte les équipements-types.

#### *Équipements participant au solde fatal en énergie*

Comme il l'a été expliqué dans la partie « Définir les équipements par usage » (cf. page 65), une série d'équipements consommant de l'électricité est volontairement omise par la segmentation présentée dans les pages précédentes, et ce, pour des raisons de « non pertinence » vis-à-vis des appels de puissance domestiques.

Il en résulte une **quantité d'énergie non modélisée** que nous appelons solde fatal en énergie. L'objectif est de donner une forme en puissance à cette quantité consommée sans distinguer au sein de ce solde ni les équipements ni même les usages.

Tous les individus du parc habité sont responsables d'une part du solde fatal en énergie. Nous formulons et adoptons l'hypothèse suivante : cette quantité d'énergie doit être répartie entre les individus du parc modélisé **au pro rata** de leur consommation respective d'électricité spécifique.

### Caractériser les équipements en équipements-types

Le but de cette fonction est d'attribuer des caractéristiques aux équipements précédemment listés de manière à les identifier au moyen d'une **nomenclature\***, de leur affecter des éléments relatifs à leur fonctionnement (**paramètres fonctionnels\***) et de leur transmettre des caractéristiques décrivant la façon dont ils peuvent être mis en marche, arrêtés ou programmés (**paramètres de gestion\***). Concrètement, cette opération de caractérisation des appareils nous permet de définir des équipements-types.

En matière de nomenclature nous avons retenu les caractéristiques suivantes :

- usage fourni ;
- nom de l'équipement-type ;
- nature de l'équipement-type (fixe ou nomade) ;
- âge de l'équipement-type ;
- durée de vie de l'équipement-type.

Les deux premiers paramètres permettent de sélectionner facilement un équipement-type dans la base de données associée au modèle de reconstitution de courbes de charge. Les trois suivants caractérisent l'appareil et s'avèrent particulièrement utiles dans la phase de construction du parc habité et de la gestion de sa dynamique (cf. fonction F 1 détaillée dans le paragraphe 2.2).

En ce qui concerne les paramètres fonctionnels des appareils, nous avons cherché à lister l'ensemble des caractéristiques intrinsèques pertinentes influant sur les appels de puissance des équipements. Au terme de la réflexion, cette liste est la suivante<sup>28</sup> :

- classe énergétique ;
- **puissance nominale\*** ou installée ;
- **puissance de veille\*** ;
- rendement en fonction du **niveau de charge\***<sup>29</sup> ;
- rendement en fonction des conditions extérieures<sup>30</sup> ;
- **Cycle(s) Unitaire(s) de Puissance\* (CUP\*)** :

*Un cycle unitaire de puissance est le motif singulier de courbe de charge d'un équipement-type pour un **mode de fonctionnement\*** donné*

- durée(s) du (des) cycle(s) de fonctionnement ;
- mode(s) de fonctionnement :

*Un mode de fonctionnement correspond au paramétrage d'un équipement-type programmé par l'utilisateur suivant lequel ce dernier fonctionne de manière discrète (choix du type de cycle d'une machine à laver : 30 °C, 60 °C, 90 °C...) ou continue (réglage de la température de consigne d'un réfrigérateur, de la luminosité et du niveau sonore d'une télévision...);*

- puissance de recharge ;
- durée de recharge ;
- consommation « au plus juste<sup>31</sup> »
- **capacité de l'équipement\***, niveau(x) de charge possible(s), taille.

La connaissance de données caractéristiques en puissance des appareils (cycles unitaires de puissance, puissance nominale, puissance de veille) est indispensable pour le calcul des **courbes de charge unitaires\*** (courbes de charge journalières des équipements-types d'un individu d'un n-uplet) abrégées **CDCU\***. Nous détaillons dans la partie 2.3 cette méthode de calcul opérée par la fonction F 3. Deux possibilités

28. Chacun de ces attributs s'applique spécifiquement aux équipements domestiques, i. e. la pertinence de chacun d'eux est donc fonction des propriétés des appareils

29. De manière à modéliser la perte d'efficacité à charge partielle de fonctionnement notamment

30. **Coefficient De Performance (COP)\*** d'une pompe à chaleur par exemple

31. C'est le cas par exemple des machines à laver pesant le linge

s'offrent à l'utilisateur en fonction des données dont il dispose pour chacun des équipements-types : soit il renseigne le(s) cycle(s) unitaire(s) de puissance pour le(s) mode(s) de fonctionnement disponible(s) au niveau de l'équipement-type, soit il paramètre ce dernier avec des éléments plus généraux (éléments d'une fiche constructeur par exemple) et le modèle est capable de calculer ou de faire correspondre un (ou plusieurs) cycle(s) unitaire(s) de puissance au(x) mode(s) de fonctionnement définis pour cet équipement-type.

Enfin, en ce qui concerne les paramètres de gestion des équipements-types, nous avons choisi de retenir les variables suivantes :

- asservissement à une période tarifaire ;
- éligibilité à une mesure d'effacement ;
- programmabilité ;
- pilotage (par communication avec le compteur, à distance via smart phone ou internet) ;
- contraintes de fonctionnement (« précédé de... », « prioritaire sur... », « incompatible avec... ») .

Ces attributs conditionnent de manière sensible les séquences de mise en marche et à l'arrêt des différents appareils d'une habitation. Ils interviennent au premier plan dans le calcul des courbes de charge unitaires (voir la section correspondante<sup>32</sup> page 92).

## 2.1.2 Définir le parc habité année par année pour chaque zone géographique (F 2.2)

En préambule, nous définissons le terme d'*année de calage*\* :

*L'année de calage désigne la dernière année pour laquelle des données réalisées sont disponibles aussi bien en énergie qu'en puissance.*

### 2.1.2.1 Créer le parc habité au sein de la zone géographique étudiée à l'année de calage (F 2.2.1)

Pour sa construction, le parc habité est considéré zone géographique par zone géographique. Le découpage d'un territoire en zones géographiques est opéré conformément aux chroniques météorologiques disponibles : ce sont les chroniques qui font les zones, les zones constituent ensuite le territoire. Dans la partie 2.3, nous verrons que ce découpage est essentiel pour la prise en compte de l'influence météorologique – en particulier pour les équipements-types fixes – lors des calculs des courbes de charge.

Nous avons identifié quatre méthodes potentielles pour la construction du parc habité à l'année de calage sans en privilégier une par rapport aux autres. Chacune d'elles s'effectue en trois étapes (les deux premières phases sont communes) comme l'indique le processus suivant :

1. Définir la répartition des bâtiments-types sur la zone géographique considérée ;
2. Définir la proportion de chaque ménage-type pour chaque bâtiment-type ;
3. Distribuer les équipements-types pour chaque couple structurant<sup>33</sup>...
  - (a) ... en fonction de règles statistiques ;
  - (b) ... en fonction de règles définies par l'utilisateur ;
  - (c) ... en fonction de distributions de probabilité ;
  - (d) ... en formant les triplets puis les n-uplets.

C'est donc au niveau de la troisième étape de ce processus que les méthodes se différencient. Au fur et à mesure des travaux d'analyse fonctionnelle, nous avons cherché à définir des variantes de la méthode toujours plus élaborées et cohérentes. Il s'agit des versions 3(a) à 3(d) qui apparaissent dans un ordre chronologique par rapport à nos travaux de recherche : la méthode 3(a) est la plus simple tandis que la

32. « Construire une courbe de charge journalière par équipement-type du n-uplet considéré en fonction de son comportement et pour chaque jour-type »

33. Suivant une méthode parmi les possibles

3(d) nous semble la plus aboutie.

Nous détaillons dans les encadrés ci-dessous les différentes étapes<sup>34</sup> des méthodes proposées pour la construction du parc à l'année de calage.

1. Définir la répartition des bâtiments-types sur la zone géographique considérée

Pour construire le parc habité à l'année de calage et pour la zone géographique étudiée, il faut dans un premier temps définir la répartition des bâtiments-types sur cette zone c'est-à-dire renseigner les effectifs de chacun d'eux.

2. Définir la proportion de chaque ménage-type pour chaque bâtiment-type

Il faut ensuite associer à ces bâtiments-types les ménages-types qui les occupent. Soulignons ici que notre modèle ne considère que les résidences principales : il y a donc autant de bâtiments que de ménages sur le territoire et au sein de chaque zone géographique.

Concrètement dans cette étape, il s'agit tout d'abord de constituer tous les couples bâtiments-types/ménages-types (couples structurants) qui auront été jugés possibles par les règles logiques d'association. Les couples structurants très peu voire non probables seront d'effectif nul.

En effet pour un ménage donné, le choix d'un logement est directement lié à sa composition et à son revenu (et/ou CSP). La mécanique d'attribution des ménages-types aux bâtiments-types peut être conduite par exemple au moyen de distributions de probabilité traduisant ainsi la propension différente des ménages à se loger, et ce, en fonction de leurs caractéristiques.

3(a) Distribuer les équipements-types pour chaque couple structurant... en fonction de règles statistiques

Cette solution est la voie la plus simple conceptuellement parlant et elle opère en cinq étapes :

1. définir des taux d'équipement globaux à l'échelle du parc modélisé et ce, usage par usage ;
2. définir pour chacun des usages des taux d'équipement des équipements-types : nous appellerons ces nombres les **clés de répartition\*** ;
3. construire statistiquement des jeux d'équipements-types fixes et nomades (en prenant par exemple une base de cent ménages à équiper). L'allocation des équipements-types aux couples structurants se fait aléatoirement : cent jeux d'équipements-types sont donc obtenus ;
4. corriger les combinaisons créées en assurant d'une part que les taux d'équipement des ménages équipés soient bien vérifiés et d'autre part que des associations logiques d'équipements-types au sein des ménages soient présentes ;
5. distribuer les équipements-types (fixes et nomades) sur l'ensemble du parc à partir des cent (ou moins) jeux d'équipements-types précédemment constitués.

34. De manière à éviter toute ambiguïté, nous identifions chacun des encadrés à l'aide de la numérotation de la liste précédente, numérotation qui n'a donc rien à voir avec la hiérarchie des fonctions du modèle de reconstitution de courbe de charge

3(b) Distribuer les équipements-types pour chaque couple structurant... en fonction de règles définies par l'utilisateur

Cette méthode se différencie de la précédente du fait qu'elle consiste à « ventiler » (distribuer) manuellement<sup>35</sup> les équipements-types aux ménages-types. L'utilisateur doit toutefois assurer la cohérence en matière de jeux d'équipements-types pour tous les couples bâtiments-types/ménages-types.

Pour ce faire, nous proposons deux règles directrices :

- allouer les équipements-types fixes en fonction du couple structurant ;
- allouer les équipements-types nomades en fonction des caractéristiques du ménage-type, de la typologie et de la surface du bâtiment-type.

La première règle souligne que les équipements-types de chauffage, de climatisation et d'ECS sont plus fortement dépendants du bâtiment dans lequel ils sont installés que le sont les autres appareils domestiques. Cette logique traduit également les contraintes d'installation des équipements-types fixes puisque tout système de cette catégorie ne peut effectivement pas équiper tout bâtiment-type.

Par l'intermédiaire de la seconde règle logique, nous voulons assurer une certaine cohérence pour l'allocation des équipements-types nomades notamment en vue de respecter le taux d'équipement des ménages équipés. Il est par exemple peu probable de disposer de deux réfrigérateurs en LC. De même, le nombre de points lumineux dans cette typologie de logement est généralement moindre qu'en MI.

3(c) Distribuer les équipements-types pour chaque couple structurant... en fonction de distributions de probabilité

Il s'agit ici d'une version plus élaborée que les deux précédentes pour l'affectation des équipements-types au parc de logements habités. En cela, elle se rapproche de la méthode 3(d).

La méthode 3(c) opère comme suit :

- définir les distributions de probabilité de chaque équipement-type (fixe et nomade) en fonction de chaque couple bâtiment-type/ménage-type ;
- équiper les ménages du parc conformément aux résultats des distributions définies ci-dessus ;
- vérifier la cohérence des taux d'équipement des équipements-types générés par les distributions de probabilité.

35. C'est l'utilisateur qui opère cette tâche

3(d) Distribuer les équipements-types pour chaque couple structurant... en formant les triplets puis les n-uplets

Comme nous l'avons défini page 61, l'ajout des équipements-types fixes aux couples structurants créent des triplets. Nous introduisons ici le concept d'**éligibilité d'un couple structurant**\*<sup>36</sup> par rapport à un équipement donné. Cette notion désigne la

*propension d'un couple bâtiment-type/ménage-type à être doté d'un équipement. Elle est évaluée à partir de règles logiques.*

Ces règles peuvent être implémentées dans une matrice qui comporte en ligne les couples structurants et en colonnes les modalités des règles logiques en fonction des caractéristiques du couple structurant et de la zone géographique (si cette influence est effectivement pertinente sur la distribution des équipements-types).

Ceci permet de caractériser par une série de modalités chaque couple structurant ; les combinaisons des attributs qui n'excluent pas un équipement donné pour un couple structurant le rend donc éligible à cet appareil.

Nous calculons alors le nombre de couples structurants éligibles pour chaque équipement de chauffage, climatisation et ECS en sommant les effectifs des couples structurants identifiés éligibles pour chacun des appareils.

D'un autre côté, les données sectorielles (parts de marché, chiffres de ventes...) permettent de connaître le nombre d'équipements présents dans le parc habité à l'échelle géographique considérée, équipements à distribuer sur les couples structurants éligibles.

Des hypothèses supplémentaires sont toutefois requises pour descendre au niveau des équipements-types en matière d'effectifs à répartir.

Une fois les triplets formés, il demeure seulement à les équiper d'équipements-types nomades. À partir de données sectorielles, le nombre d'équipements-types nomades à distribuer peut être calculé. Comme précédemment, des hypothèses supplémentaires sont nécessaires pour descendre au niveau des équipements-types. Des règles logiques permettent de déterminer quels sont les triplets éligibles pour chaque équipement à distribuer.

Afin de visualiser la construction du parc habité plus intuitivement que par les seules règles qui mènent à sa formation, nous avons réfléchi à une façon générique de représenter le parc. Cette représentation, que nous détaillons dans le paragraphe ci-après, est valable quelle que soit la manière avec laquelle les équipements-types sont distribués au sein des couples structurants.

### Représentation visuelle du parc

Le parc habité peut être représenté sous forme d'un tableau où les n-uplets en seraient les lignes. En colonnes seraient renseignés le bâtiment-type, le ménage-type et la présence ou l'absence de chaque équipement-type et indiquer l'éventuel multi-équipement. En plus de ces informations, les effectifs des n-uplets – les poids dans l'effectif – doivent également être renseignés. Un exemple schématique d'une telle représentation est proposé au travers du tableau 2.1.

Comme nous l'avons décrit dans cette partie, l'attribution des éléments-types entre eux est fonction de règles déterministes ou probabilistes qui se traduisent par des proportions. Ces valeurs, qui doivent être intégrées au tableau représentatif du parc modélisé à la suite de sa constitution, sont donc facilement utilisables pour effectuer des opérations de calcul.

Dans le processus de création du parc habité, nous faisons l'hypothèse que l'ensemble des ménages dispose des équipements non explicités puisqu'ils participent tous au solde fatal en énergie.

36. L'éligibilité d'un ménage-type ou d'un triplet ont également tout leur sens

n-uplet	bâtiment- type	ménage- type	équipements-types				poids dans l'effectif [%]
			$et_1$	$et_2$	...	$et_n$	
$nu_1$	$bt_4$	$mt_2$	1	0	...	2	18
$nu_2$	$bt_1$	$mt_3$	0	1	...	1	13,5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$nu_m$	$bt_{12}$	$mt_7$	2	3	...	0	0,5

TABEAU 2.1 – Principe de visualisation du parc habité sous forme d'une matrice n-uplets  $\times$  éléments-types

### 2.1.2.2 Faire évoluer le parc habité de la zone géographique étudiée année par année (F 2.2.2)

Comme nous le verrons dans la section 2.2 lorsque nous aborderons la projection des scénarios de modélisation, le parc habité est une entité qui doit évoluer de manière dynamique. Ainsi, dans notre modèle, le parc construit à l'année de calage doit être recalculé année par année suivant les scénarios d'évolution définis dans la fonction F 1.

Tenir compte de la dynamique du parc revient à reconsidérer chacun de ses constituants en suivant des étapes analogues à celles qui ont mené à sa construction à l'année de calage. Au final, il s'agit d'un processus en 5 étapes :

1. Redéfinir le parc de bâtiments-types ;
2. Redéfinir la population de ménages-types du parc habité ;
3. Redéfinir le parc d'équipements-types ;
4. Rajouter les nouveaux n-uplets ;
5. Recalculer les poids de représentativité des couples structurants.

Chacune de ces étapes est décrite dans un des encadrés ci-dessous.

## 1. Redéfinir le parc de bâtiments-types

Reconsidérer le parc de bâtiments-types consiste tout d'abord à intégrer les nouvelles constructions. Ensuite, il s'agit de suivre l'évolution dans le temps des bâtiments déjà existants. Enfin, il suffit de recalculer la part de chaque type d'habitat dans le parc immobilier nouvellement élaboré. Chacune de ces étapes est décrite dans les paragraphes ci-dessous.

### *Définir de nouveaux bâtiments-types*

Entre l'année  $n$  et  $n + 1$ , de nouveaux bâtiments-types peuvent voir le jour suite à un événement-clé. Ils doivent donc être ajoutés à la base de données pour pouvoir être intégrés dans le parc.

### *Définir des générations de bâtiments-types du parc habité d'une année sur l'autre*

Le parc de bâtiments vieillit tout au long de la durée de simulation. Pour tenir compte de l'hétérogénéité des âges de construction et pour pouvoir mieux modéliser l'évolution de l'habitat (notamment en matière de rénovation et de destruction d'éléments anciens du parc résidentiel), des générations de bâtiments doivent être modélisées.

Ce processus se déroule en quatre étapes que nous allons plus largement expliciter dans les points ci-dessous.

- Étape 1 : définir les flux pour chaque bâtiment-type du parc de référence

Commençons par définir le concept de **parc de référence**\* :

*Le parc de référence correspond au parc habité construit l'année précédente à celle présentement considérée.*

Dans un premier temps, nous cherchons à connaître l'effectif de chaque bâtiment-type dans le parc habité nouvellement modélisé. Pour ce faire, nous devons appliquer les trois règles ci-dessous :

1. définir le nombre de bâtiments disparus par bâtiment-type (par destruction ou **réaffectation**\*<sup>37</sup> du parc de référence ;
2. retirer les bâtiments disparus du parc de référence ;
3. faire perdurer les bâtiments du parc de référence qui n'ont subi aucun changement<sup>38</sup> d'une année sur l'autre.

- Étape 2 : définir le flux de bâtiments neufs pour chaque bâtiment-type

Dans un deuxième temps, nous devons renseigner l'effectif des nouveaux bâtiments entrant dans le parc. Pour cela, nous calculons les **flux nets**\*<sup>39</sup> bâtiment-type par bâtiment-type en trois étapes :

1. définir le nombre total de bâtiments neufs construits ;
2. définir la proportion parmi les constructions neuves de chaque bâtiment-type ;
3. inclure les constructions neuves dans le parc d'une année sur l'autre.

- Étape 3 : définir le flux de bâtiments rénovés pour chaque bâtiment-type

Nous devons ensuite prendre en compte l'habitat rénové. Selon l'intensité des travaux de rénovation et la qualité initiale, en matière d'isolation thermique, du bâtiment-type rénové, celui-ci passera de sa classe de performance thermique d'enveloppe d'origine à une classe meilleure.

37. La réaffectation d'un bâtiment correspond au changement de la valorisation qui est faite de lui : par exemple des locaux tertiaires transformés en habitations résidentielles

38. Rénovation, destruction ou réaffectation

39. Le flux net d'un élément-type d'un parc habité correspond à la somme des retraits du parc de cet élément-type comptés négativement et des ajouts dans le parc de cet élément-type comptés positivement

Parallèlement, la rénovation peut se traduire par une diminution de l'inertie d'un bâtiment. Dans tous les cas, la rénovation entraîne des **sauts de classe**\* de performance énergétique du bâtiment. Considérer cette modification s'effectue en trois phases :

1. définir le nombre total de bâtiments rénovés du parc de référence ;
2. définir la proportion parmi les bâtiments rénovés de chaque bâtiment-type ;
3. définir la nature et la proportion des sauts de classe de performance énergétique pour chaque bâtiment-type. Si des données sont disponibles et si cette influence est sensible, nous définirons également la part de bâtiments-types rénovés à surface constante (l'autre part correspondant aux augmentations de surface qui accompagnent parfois des travaux de rénovation).

● **Étape 4** : définir le flux de bâtiments réaffectés pour chaque bâtiment-type

Pour conclure nous devons renseigner le nombre de bâtiments réaffectés entre deux années consécutives. Ainsi nous nous intéressons particulièrement aux réaffectations de bâtiments de l'industrie (usines et entrepôts) ou du tertiaire vers le résidentiel et inversement.

D'une année sur l'autre, des bureaux peuvent par exemple être réutilisés en tant que logements et doivent être modélisés à la nouvelle année considérée. Nous incluons également dans la catégorie des bâtiments réaffectés les « nouveaux » bâtiments du parc suite à division <sup>40</sup>.

La prise en compte des bâtiments réaffectés suit une logique en trois points :

1. définir le nombre total de bâtiments réaffectés ;
2. définir la proportion parmi les bâtiments réaffectés de chaque bâtiment-type ;
3. inclure les bâtiments réaffectés dans le parc d'une année sur l'autre.

*Définir la proportion dans le parc habité de chaque bâtiment-type*

Après avoir pris en compte les flux des bâtiments neufs, rénovés et réaffectés, il faut définir la proportion dans le parc modélisé de chaque bâtiment-type résultant de ces transformations. Celles-ci doivent être renseignées « à la main » si aucun scénario d'évolution correspondant n'a été défini au sein de la fonction F 1. Cette dernière étape clôtur le processus de reconstitution d'une année sur l'autre du parc de logements.

## 2. Redéfinir la population de ménages-types du parc habité

Après avoir fixé les bâtiments-types sur le parc, l'étape suivante de reconstitution du parc habité est de définir la population de ménages qui occupe ces habitations.

Pour cela, nous devons tenir compte de trois sources de changement :

- l'arrivée de nouveaux ménages-types dans le parc ;
- l'évolution des ménages en matière d'effectifs ;
- la proportion de chaque ménage-type qui occupe chaque bâtiment-type.

Pour le premier de ces trois points, nous devons définir des ménages-types aux caractéristiques non encore existantes dans le parc à l'année simulée précédente. Pour la deuxième source d'évolution, nous devons prendre en compte les changements démographiques entre deux années consécutives en matière d'effectifs de ménages. Le **solde naturel**\* <sup>41</sup> et le **solde migratoire**\* <sup>42</sup> sont les deux causes de modification de la population des ménages du parc habité modélisé d'une année sur l'autre.

40. Typiquement une grande maison qui ne constitue qu'un logement à l'année  $n$  et qui est scindée en deux à l'année  $n + 1$

41. « Le solde naturel (ou accroissement naturel ou excédent naturel de population) est la différence entre le nombre de naissances et le nombre de décès enregistrés au cours d'une période. Les mots « excédent » ou « accroissement » sont justifiés par le fait qu'en général le nombre de naissances est supérieur à celui des décès. Mais l'inverse peut se produire, et le solde naturel est alors négatif. » [INS12b]

42. L'INSEE donne la définition suivante du solde migratoire : « Le solde migratoire est la différence entre le nombre de

Enfin, le dernier point consiste à redéfinir de nouveaux couples structurants constitués du (des) nouveau(x) bâtiment(s)-type(s) et/ou ménage(s)-type(s) à l'aide des règles logiques déjà utilisées pour la construction du parc à l'année précédente ou à partir de règles nouvellement définies par l'utilisateur.

### 3. Redéfinir le parc d'équipements-types

Pour terminer la reconstitution du parc à une année donnée, il suffit de reconsidérer les équipements-types simulés ainsi que leur effectif respectif. Des étapes analogues à celles opérées sur les bâtiments doivent être conduites : tout d'abord, il est nécessaire de prendre en compte de nouveaux équipements-types. Ensuite à l'image des habitations, les équipements domestiques ont une certaine durée de vie dont il faut tenir compte au travers de générations d'appareils. Enfin, une fois ces deux éléments abordés, des jeux d'équipements-types sont élaborés et alloués aux couples structurants. Il s'agit alors de quantifier la proportion de chacun d'eux au sein du parc nouvellement construit.

#### *Définir de nouveaux équipements-types*

D'une année sur l'autre, de nouveaux équipements-types peuvent voir le jour et intégrer le parc. Pour prendre en considération cette évolution, nous devons élargir la base de données en créant des équipements-types aux caractéristiques non encore présentes dans le parc.

#### *Définir des générations d'équipements-types du parc habité d'une année sur l'autre*

À l'image du travail réalisé sur les bâtiments-types décrit dans l'encadré<sup>43</sup> page 80, nous reprenons des étapes similaires pour modéliser le vieillissement des équipements-types. Ce processus schématiquement explicité par la figure 2.13 se déroule en trois étapes que nous décrivons dans les points ci-dessous.

- **Étape 1 : définir l'évolution numérique pour chaque équipement-type du parc de référence**

Pour prendre en compte l'évolution numérique de chaque équipement-type du parc modélisé d'une année  $n$  à une année  $n + 1$ , nous devons opérer en deux temps :

1. définir le nombre d'équipements disparus par équipement-type du parc de référence et les retirer du parc ;
2. faire perdurer les générations d'équipements qui n'ont subi aucun changement<sup>44</sup> année par année.

- **Étape 2 : définir le flux d'équipements neufs pour chaque équipement-type**

Il faut ensuite prendre en compte l'arrivée (ou le départ<sup>45</sup>) de nouveaux appareils dans le parc, c'est-à-dire un taux d'équipement de la population plus important pour un flux net positif (ou plus faible dans le cas contraire) ou bien un **suréquipement\***<sup>46</sup> (respectivement du **sous-équipement\***) des ménages déjà équipés dans le parc de référence. Pour cela, nous opérons en deux temps :

1. définir le nombre total d'équipements neufs entrant dans le parc ;
2. définir la proportion parmi les équipements neufs de chaque équipement-type.

personnes qui sont entrées sur le territoire et le nombre de personnes qui en sont sorties au cours de l'année. Ce concept est indépendant de la nationalité. » [INS12a]. Le solde migratoire représente donc le flux net entre l'immigration, l'émigration et les déménagements

43. « Définir des générations de bâtiments-types du parc habité d'une année sur l'autre »

44. Par changement nous entendons ici tout ce qui peut être considéré comme de la rénovation entraînant une modification de la manière dont l'appareil appelle de la puissance sur le réseau. La rénovation d'un système de chauffage en est un exemple.

45. Nous rappelons que nous raisonnons en flux net

46. Le taux d'équipement des ménages équipés augmente, i. e. plus d'un appareil par n-uplet équipé est présent au sein de la zone géographique considérée

• Étape 3 : définir le flux d'équipements remplacés pour chaque équipement-type  
 Un équipement-type peut atteindre sa fin de **vie** d'une année sur l'autre. Il se peut qu'il soit remplacé. Il faut donc définir le flux net de remplacement des équipements-types entre deux années consécutives. Trois étapes sont alors requises :

1. définir le nombre total d'équipements remplacés du parc de référence ;
2. définir la proportion parmi les équipements remplacés de chaque équipement-type ;
3. définir qualitativement et quantitativement le remplacement de chaque équipement-type en fonction du couple structurant. Concrètement il faut revoir la proportion de chaque équipement-type dans le parc.

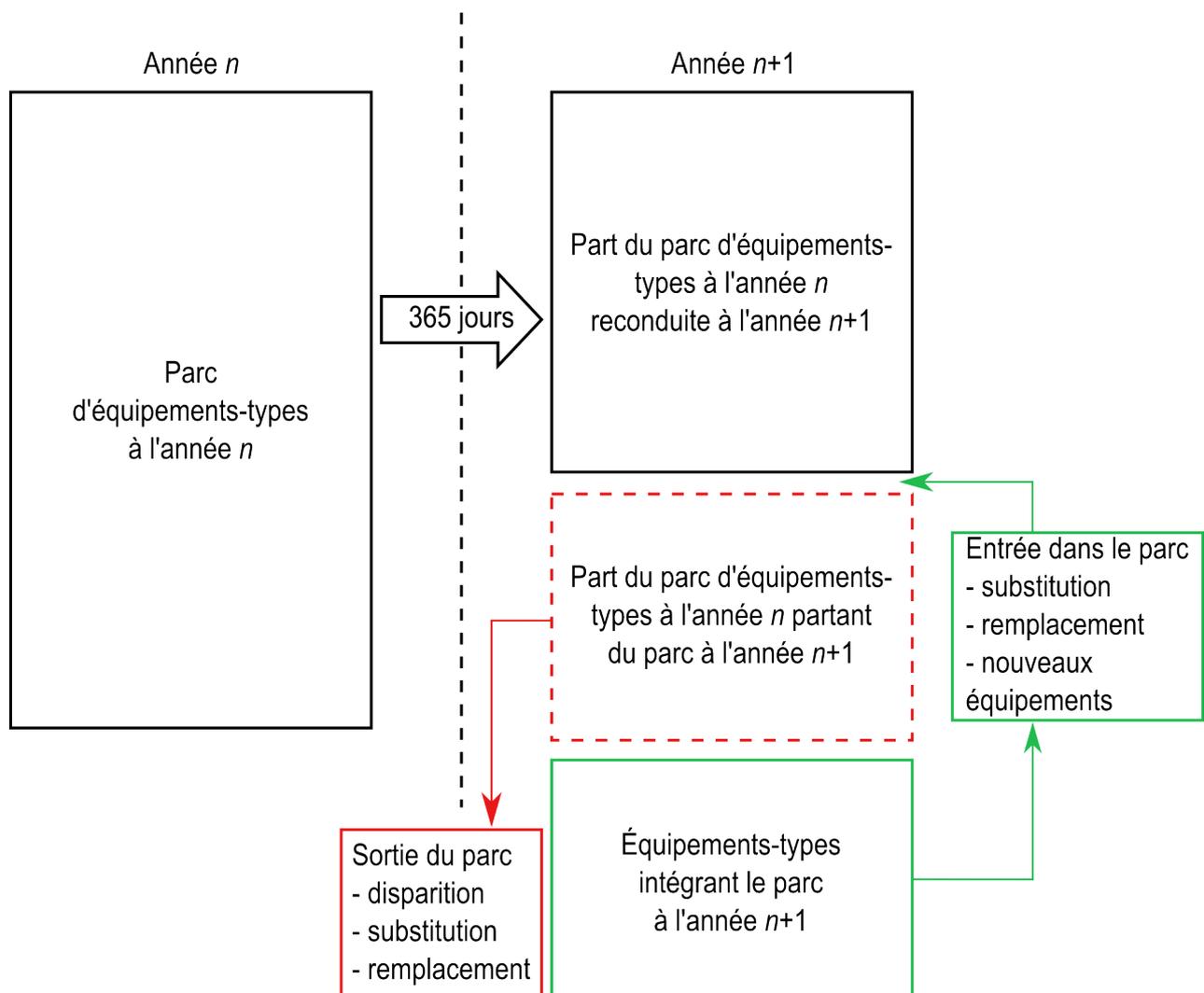


FIGURE 2.13 – Constitution de générations d'équipements-types année par année

Pour ce travail, plusieurs influences doivent être prises en compte :

- la durée de vie des appareils qui implique la vétusté ou le remplacement prématuré (panne) ;
- les évolutions et les ruptures technologiques ;
- les évolutions des caractéristiques des ménages-types ;
- l'évolution des réglementations<sup>47</sup>.

Tous ces éléments peuvent être intégrés au préalable dans un scénario d'évolution défini au niveau de la fonction F 1 ou bien ils doivent être renseignés manuellement dans le modèle au niveau de cette fonction (F 2.2.2).

*Définir la proportion de chaque jeu d'équipements-types pour chaque couple structurant*

Comme nous l'avons vu précédemment au sein de l'encadré 3(d)<sup>48</sup>, les équipements-types (fixes et nomades) peuvent être distribués en fonction des propriétés des ménages et de leur habitat respectif. Pour un couple structurant donné, différents équipements-types fixes pourront lui être associés. De même pour un certain triplet, plusieurs jeux d'équipements-types nomades peuvent lui être attribués. Pour tenir compte de ces constats, l'utilisateur doit définir la proportion de chaque combinaison d'appareils domestiques (fixes et nomades) pour tous les couples structurants modélisés. Ceci assure une répartition non uniforme des équipements-types aux couples structurants.

#### 4. Rajouter les nouveaux n-uplets

La reconstitution du parc à une année donnée peut générer des n-uplets qui n'existaient pas dans le parc de référence. Il faut donc les rajouter dans le parc nouvellement modélisé. Cela revient à rajouter des lignes à la base de données représentative du parc habité.

#### 5. Recalculer les poids de représentativité des couples structurants

La reconstitution du parc habité introduit, comme nous l'avons vu dans les pages précédentes, une série de changements qualitatifs et quantitatifs des éléments-types manipulés. Les poids de représentativité des couples structurants sont donc à recalculer.

La présente fonction (F 2.2.2) a été pensée de manière à assurer une continuité du parc habité et donc améliorer la cohérence en matière de modélisation et de répartition géographique des n-uplets.

À chaque fois que le parc est reconstitué, les différentes évolutions joueront :

- soit sur les effectifs des n-uplets ;
- ou bien sur l'existence d'un n-uplet en soi (apparition et disparition). Ceci se traduit respectivement par un ajout et un retrait de lignes à la base de données ;
- ou alors sur l'arrivée ou le départ d'un élément-type dans le parc modélisé (respectivement ajout et suppression de colonnes).

L'avantage principal de cette manière opérationnelle est de ne pas avoir à tout réaffecter, c'est-à-dire à reconstruire à partir de zéro tous les n-uplets. Pour finir, ce sont les scénarios qui feront évoluer les n-uplets en matière de composition et d'effectif respectif.

### 2.1.3 Contrôler la cohérence des poids dans l'effectif des n-uplets du parc à l'échelle du territoire (F 2.3)

Cette fonction permet d'évaluer la cohérence en matière de modélisation du parc habité à l'échelle du territoire (agrégation de toutes les zones géographiques).

47. Notamment la réglementation thermique ainsi que les réglementations de performance énergétique des équipements domestiques

48. « Distribuer les équipements-types pour chaque couple structurant... en formant les triplets puis les n-uplets »

Les poids dans l'effectif définis page 61<sup>49</sup> sont des fractions d'un effectif considéré. Dans tous les cas, ils peuvent être manipulés eux-mêmes comme des effectifs.

F 2.3 peut être tout à fait transparente vis-à-vis de l'utilisateur. Toutefois, elle doit lui indiquer les éventuelles incohérences qui peuvent apparaître par agrégation des parcs détaillés à l'échelle de chaque zone géographique. Cette fonction peut également calculer des proportions d'éléments donnés du parc habité. En effet, il peut s'avérer intéressant de connaître la part dans le parc des MI, des LC, des anciennes habitations...

La fonction F 2, que nous avons décrite dans cette section, permet de construire le parc habité à une année  $n$ . De plus, elle prend en compte les scénarios d'évolution définis au sein de la fonction F 1 pour faire évoluer dans le temps le parc modélisé. Dans la prochaine partie, nous abordons plus en détails l'élaboration de scénarios de modélisation auxquels nous avons déjà fait référence à plusieurs reprises.

---

49. Au sein du paragraphe « Finalité »

## 2.2 Analyse de la fonction « Scénariser et projeter l'évolution du contexte pour toute la durée de simulation (F 1) »

### Finalité

En entrée de simulation, il faut définir le contexte général ce qui se traduit concrètement par des événements-clés qui marquent le début et/ou la fin d'une période et des **tendances d'évolution\*** qui ont une influence d'une part sur les flux nets des éléments-types et d'autre part sur le solde fatal en énergie durant la période considérée. La figure 2.14 représente schématiquement les entrées et les sorties de la fonction F 1.

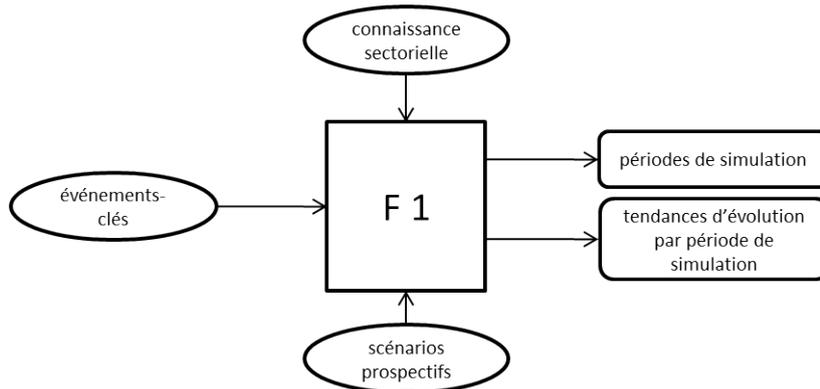


FIGURE 2.14 – Schéma d'entrées/sorties de la fonction F 1

Dans les parties relatives aux fonctions F 1.1 et F 1.2<sup>50</sup>, nous expliciterons l'intérêt de ces concepts – événement-clé et période de simulation notamment – pour la modélisation des courbes de charge.

### 2.2.1 Définir des événements-clés qui impactent les flux et le solde fatal en énergie (F 1.1)

Comme nous l'avons vu au sein de la fonction F 2, le parc habité est constitué de bâtiments-types, de ménages-types et d'équipements-types, chacun d'eux étant indistinctement désigné par le terme d'élément-type. Définis pour un instant donné, les éléments-types évoluent au cours du temps de manière dynamique d'où la nécessité du concept d'événement-clé, préalablement défini au sein du paragraphe 2.1 p. 60, qui les impacte.

Des événements-clés relatifs à un ou plusieurs éléments-types sont donc à définir en entrée de simulation. Pour un bâtiment-type, il peut s'agir par exemple de l'entrée en vigueur d'une nouvelle réglementation thermique. En ce qui concerne les ménages-types, un événement-clé peut se traduire par exemple par une tendance à la paupérisation de la population, à son vieillissement, aux changements structurels des ménages suivant leurs cycles de vie<sup>51</sup> ou encore à la tertiarisation de l'habitat (travail à domicile). Enfin, les événements-clés qui concernent les équipements-types peuvent être particulièrement variés. Un exemple trivial est celui de l'introduction d'un nouvel équipement-type dans le parc modélisé.

Ces événements-clés sont définis à l'échelle annuelle pour toute la durée de la simulation et délimitent ainsi des périodes.

Nous représentons par l'intermédiaire de la figure 2.15 comment s'organisent les concepts – événement-clé, tendance d'évolution et période de simulation – précédemment définis dans le cas où sont considérés les taux d'équipement<sup>52</sup> des ménages de quelques appareils domestiques. Sur cette figure, nous pouvons

50. Respectivement détaillées aux paragraphes 2.2.1 et 2.2.2

51. Les cycles de vie d'un ménage sont des périodes engendrées par des ruptures que sont notamment l'installation du ménage dans son logement, l'insertion dans le monde professionnel, l'arrivée d'un nouveau né, l'acquisition de la résidence principale, la séparation du couple, le départ des enfants du foyer familial, le départ à la retraite, le veuvage... [CAY11]

52. La dynamique des taux d'équipement dans la population constitue une tendance d'évolution

constater que le taux d'équipement de l'équipement-type  $et_1$  n'est pas affecté par les événements-clés définis (évolution progressive dans le temps du taux de pénétration de cet appareil) alors que le taux d'équipement en  $et_3$  est largement modifié dans le temps par trois événements-clés ( $ec_2$ ,  $ec_3$  et  $ec_4$ ).

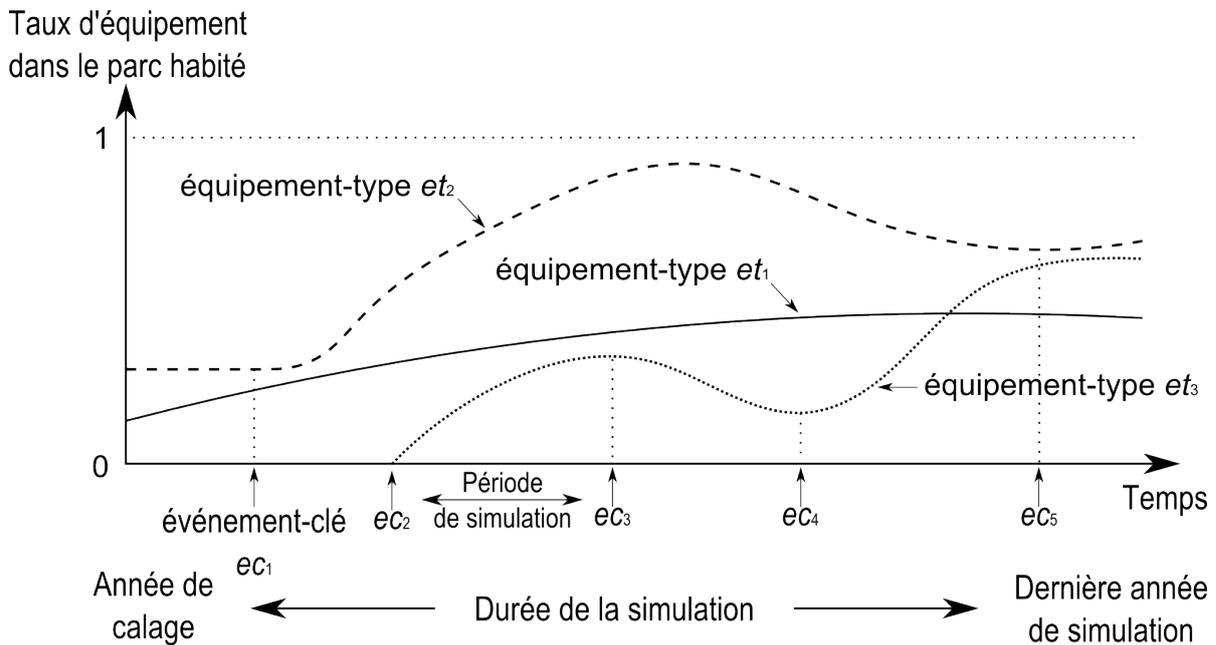


FIGURE 2.15 – Évolution du taux de pénétration de trois équipements-types dans le parc habité

De manière plus générale que le montre le graphique 2.15, les événements-clés font que les flux nets des éléments-types évoluent tout au long d'une simulation. Nous avons pu constater lors de l'étude de la fonction F 2.2<sup>53</sup>, que le renseignement des flux nets permet d'aboutir aux effectifs de chaque élément-type, et ce, année par année.

Comme nous l'avons souligné dans le préambule, nous ne pouvons pas expliciter et donc modéliser l'ensemble des équipements domestiques des ménages ce qui nous conduit à définir un solde fatal en énergie. Cette quantité est estimée par le modèle année par année selon un scénario d'évolution. Le solde fatal est par exemple directement impacté par la modélisation explicite d'un nouvel équipement ou d'un usage.

### 2.2.2 Définir les tendances d'évolution des flux et du solde fatal en énergie pour chaque période (F 1.2)

Seulement évoqué dans les pages précédentes, nous commençons ici par préciser ce que nous appelons tendance d'évolution. Il s'agit de

*la manière avec laquelle sont modifiés qualitativement et/ou quantitativement les objets (éléments-types notamment) et les paramètres de la modélisation au cours de la durée totale de la simulation.*

Définir des tendances d'évolution consiste à indiquer au modèle comment vont changer, année par année, une grande partie des données numériques manipulées par le modèle de reconstitution de courbes de charge. Ces évolutions peuvent être renseignées sous différentes formes par l'utilisateur :

- indication année par année des tendances pour toute la durée de la simulation ;
- sélection dans une base de données de tendances sur base mathématique entre le début (c'est-à-dire à l'année de calage) et la fin de la simulation ;
- choix dans une base de données des tendances entre chaque couple d'événements-clés consécutifs.

53. « Définir le parc habité année par année pour chaque zone géographique »

La connaissance sectorielle et les hypothèses générales d'évolution sont des guides pour le renseignement de ces scénarios qui doit être effectué manuellement par l'utilisateur.

Sont impactés directement par les scénarios d'évolution à définir, les éléments suivants :

- flux net de bâtiments neufs, rénovés, réaffectés et détruits dans le parc habité ;
- proportion de chaque bâtiment-type dans le parc habité ;
- flux net des ménages de la population suivant l'évolution démographique et migratoire ;
- proportion de chaque ménage-type dans le parc habité ;
- taux d'équipement des ménages ;
- proportion respective de chaque équipement-type dans le parc habité ;
- solde fatal en énergie (tendance éventuellement estimée suivant une combinaison des autres scénarios d'évolution).

## 2.3 Examen de la fonction « Calculer les courbes de charge à l'échelle du territoire année par année pour chaque période (F 3) »

### Finalité

Après avoir modélisé le parc habité par l'intermédiaire de la fonction F 2, la méthode doit ensuite reconstituer la courbe de charge électrique correspondante pour toute la durée de la simulation : c'est le rôle de la fonction F 3 comme l'indique la figure 2.16.

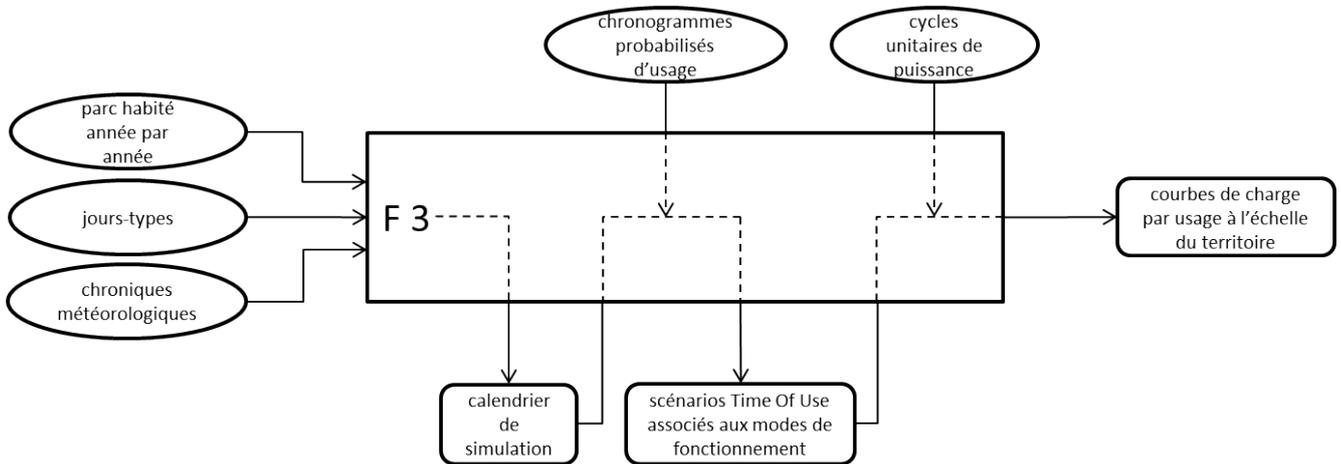


FIGURE 2.16 – Schéma d'entrées/sorties de la fonction F 3

Bien que seuls les équipements domestiques soient demandeurs d'une puissance électrique que l'on cherche à modéliser, nous avons constitué au sein de la partie 2.1 des n-uplets qui sont les entités élémentaires du parc habité. Il est donc logique que le calcul des courbes de charge se fasse n-uplet par n-uplet. D'autre part, lorsque l'on s'intéresse à la problématique des appels de puissance, nous remarquons une profonde périodicité liée aux activités humaines. Il en résulte que l'échelle de temps la plus cohérente est la journée. Enfin, nous rappelons une fois encore l'influence météorologique sur les appels de puissance des équipements présents dans l'habitat et la nécessité d'en tenir compte pour les calculs : ces derniers sont donc effectués pour le parc habité à la maille minimale de travail c'est-à-dire la zone géographique qui est caractérisée par un climat supposé homogène en tout point.

Après avoir fixé le cadre spatial de la modélisation au travers du parc habité disponible à l'année considérée, calculer les courbes de charge correspondantes nécessite tout d'abord de définir le cadre temporel de la simulation. Concrètement, il s'agit de mettre en place un calendrier que nous aborderons dans la partie 2.3.1. Une fois celui-ci élaboré, l'estimation des appels de puissance du parc est alors possible. Le calcul des courbes de charge requiert alors différentes étapes élémentaires que nous détaillerons dans les sections 2.3.2 à 2.3.4.

### 2.3.1 Construire le calendrier de simulation pour toute la durée de la simulation (F 3.1)

Le calendrier de simulation est constitué de *jours-types*\* – définis ci-dessous – caractérisant l'activité des ménages et associés à des données météorologiques.

#### 2.3.1.1 Définir les jours-types (F 3.1.1)

En préambule, nous définissons le concept de jour-type :

*Un jour-type est caractérisé par des périodes infra-journalières où sont renseignées les probabilités d'activités domestiques ou d'usages, ceci de manière transverse à l'ensemble des ménages de la population considérée.*

Ce sont les grandes habitudes de vie qui sont illustrées et prises en compte au moyen de ce concept. Un jour-type se caractérise par :

- une distribution de probabilité de l’instant de démarrage des activités domestiques pour une journée. Ce repère sera appelé  $t_{Start}$  ;
- une distribution de probabilité de l’instant de fin des activités domestiques pour une journée. Ce moment sera noté  $t_{End}$  ;
- des distributions de probabilité des usages domestiques entre  $t_{Start}$  et  $t_{End}$  ;
- les modes de gestion de chauffage/climatisation (confort, réduit, hors-gel) ;
- les **chroniques de puisage\*** d’eau chaude sanitaire (équivalentes à des distributions journalières de probabilité de puisage d’ECS).

Dans une première approche, nous pourrions définir un jour-type par jour de la semaine et par mois (soit 84 jours-types), les jours fériés étant assimilés à des dimanches. À cela doit s’ajouter un jour-type correspondant aux périodes de vacances soulignant l’absence complète des occupants de leur logement.

En considérant chaque mois de manière individuelle, c’est la saisonnalité des usages que nous souhaitons prendre en compte. Par exemple, en plus des usages thermiques dont la saisonnalité est évidente, l’éclairage est directement lié à l’ensoleillement qui varie lui-même fortement tout au long de l’année aussi bien en intensité qu’en durée à l’échelle journalière.

En outre, des attributs liés à l’activité économique indiquant la propension des occupants d’un logement à être présents chez eux ou non, viennent se superposer aux jours-types :

- jour ouvré/non ouvré (férié ou dimanche) ;
- jour travaillé/non travaillé.

Nous représentons, au moyen de la figure 2.17, la représentation incomplète d’un jour-type. En effet, sur ce graphique seules les distributions de probabilité de  $t_{Start}$ ,  $t_{End}$  ainsi que des usages audiovisuel et lavage ont été tracés.

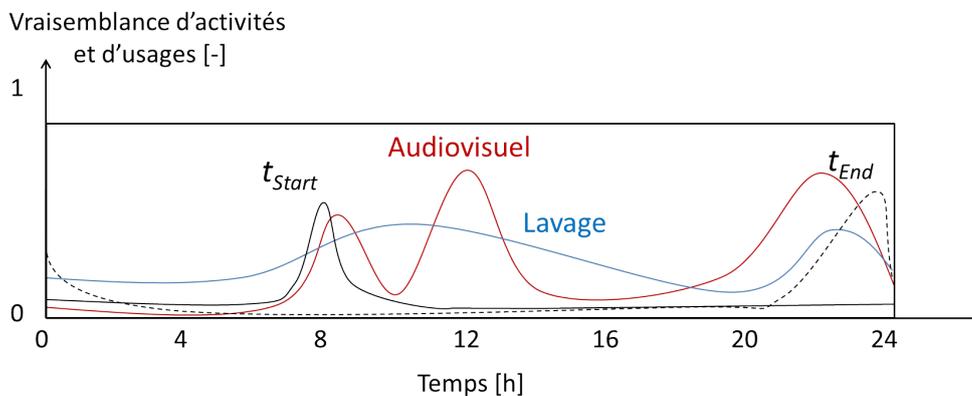


FIGURE 2.17 – Apparence mathématique (partielle) d’un jour-type

À noter sur la figure 2.17 que les distributions de probabilité des usages et des instants de début et de fin d’activités domestiques ne sont jamais nulles ce qui traduit le fait qu’à une échelle géographique simulée importante (a fortiori à la maille nationale) le modèle autorise à tout moment la réalisation de tout usage domestique.

### 2.3.1.2 Construire le calendrier de simulation en jours-types (F 3.1.2)

Une fois que chaque jour-type a été défini, il faut construire le calendrier de simulation de l’année considérée en ordonnant convenablement les jours-types de manière à recréer les diverses périodicités hebdomadaire, mensuelle, saisonnière ainsi que les vacances qui sont présentes dans le calendrier civil de l’année retenue pour la simulation.

### 2.3.1.3 Sélectionner et analyser une chronique météorologique de la zone géographique considérée (F 3.1.3)

Nous devons enfin superposer au calendrier de simulation des données météorologiques issues de chroniques mesurées. Selon la zone géographique considérée, la fonction F 3.1.3 associe les éléments de la chronique météorologique sélectionnée jour par jour au calendrier. Il s'agit a minima de :

- la température extérieure ;
- la nébulosité qui indique le niveau de couverture nuageuse du ciel ;
- le rayonnement solaire direct et diffus (ou total).

Ensuite, une seconde analyse permet de renseigner les attributs météorologiques « exceptionnels » des jours-types à savoir :

- grand froid (oui/non) ;
- canicule (oui/non) ;
- régime établi\*<sup>54</sup> (oui/non).

Enfin, la fonction F 3.1.3 renseigne également la [température d'eau froide](#)\* afin de tenir compte de la saisonnalité « fatale<sup>55</sup> » de l'ECS. Généralement cette donnée d'entrée est exprimée sous forme de moyenne mensuelle.

### 2.3.2 Calculer les courbes de charge unitaires par n-uplet pour la zone géographique étudiée et pour chaque jour de la simulation (F 3.2)

Comme nous l'avons souligné dans le préambule de cette partie, les courbes de charge sont calculées pour un individu d'un n-uplet spécifié, au cours d'un jour et au sein d'une zone géographique donnée. À noter également que nous distinguons deux méthodes pour calculer les courbes de charge des équipements d'électricité spécifique d'une part (fonction F 3.2.1), des équipements thermiques (fonction F 3.2.2) d'autre part.

#### 2.3.2.1 Calculer les courbes de charge unitaires d'électricité spécifique par n-uplet (F 3.2.1)

Les équipements-types d'électricité spécifique sont les appareils domestiques qui permettent la réalisation des usages non thermiques et ceux qui leur sont connexes. Ne font pas partie de ces équipements :

- les systèmes de chauffage et de climatisation ainsi que leurs auxiliaires ;
- les systèmes de production d'ECS et leurs auxiliaires éventuels ;
- les éléments assurant la ventilation de l'habitat ;
- certains appareils destinés à la cuisson alimentaire (plus précisément fours et plaques chauffantes).

Le calcul des courbes de charge unitaires des équipements-types est effectué en deux temps par l'intermédiaire de deux sous-fonctions présentées dans les encadrés ci-dessous. Tout d'abord, nous considérons individuellement chaque équipement-type des usages d'électricité spécifique pour un individu d'un n-uplet donné et pour chaque jour-type. Puis nous répétons le processus de calcul pour l'ensemble des équipements-types du même individu du n-uplet considéré. Ces calculs sont répétés pour les individus du n-uplet puis de même pour la totalité des autres n-uplets du parc.

54. Il est possible de représenter cet état au moyen d'une température ressentie lors de situations climatiques exceptionnelles ; celles-ci modifient le comportement des ménages face aux usages thermiques. RTE utilise par exemple une moyenne mobile sur la température des jours qui précèdent une journée spécifiée [LEF07]

55. L'eau froide puisée en hiver étant à une température sensiblement plus basse que celle prélevée en été, l'énergie requise au niveau des systèmes de production d'ECS est modifiée en conséquence tout comme le sont les courbes de charge correspondantes

Construire une courbe de charge journalière par équipement-type du n-uplet considéré en fonction de son comportement et pour chaque jour-type

La méthode de calcul des courbes de charge unitaires fait appel à plusieurs étapes que nous détaillons par la suite :

1. définir un scénario Time Of Use par équipement-type à partir des chronogrammes probabilisés d'usage issus des jours-types en assurant la compatibilité avec les autres équipements-types du n-uplet ;
2. attribuer aux scénarios Time Of Use, le(s) mode(s) de fonctionnement de chaque équipement-type pour chaque jour-type ;
3. définir un cycle unitaire de puissance pour l'équipement-type et le jour-type considérés pour chaque utilisation de l'appareil, et ce, conformément au(x) mode(s) sélectionné(s).

Avant d'aborder plus en détails le calcul des courbes de charge, nous définissons les concepts de :

- chronogramme probabilisé d'usage :

*périodes infra-journalières au cours desquelles est autorisée, n-uplet par n-uplet, une activité domestique faisant appel à l'usage d'un ou plusieurs équipement(s) électrique(s). Ces périodes sont issues du jour-type simulé et sont probabilisées conformément aux caractéristiques du n-uplet considéré. L'agrégation des chronogrammes probabilisés d'usage, pondérés par les poids dans l'effectif des n-uplets, coïncide avec les probabilités d'activités domestiques définies au niveau du jour-type ;*

- scénario Time Of use :

*instants de déclenchement – mise en marche – des équipements-types d'un n-uplet pour une journée donnée.*

Les chronogrammes probabilisés d'usage peuvent être issus d'une base de données ou bien ils peuvent être construits en deux étapes, à partir des distributions de probabilité renseignées au niveau du jour-type considéré et pour un n-uplet donné. Pour ce faire, il faut :

1. définir des plages infra-journalières où est permise l'utilisation de chacun des équipements présents chez le n-uplet considéré ;
2. attribuer à ces plages une distribution de probabilité propre au n-uplet en fonction de ses caractéristiques.

La génération de scénarios Time Of Use consiste à tirer aléatoirement dans les plages autorisées des chronogrammes d'usage de façon à définir les instants de déclenchement des équipements-types d'un n-uplet. Les différentes étapes de construction d'un scénario Time Of Use sont représentées au travers de la figure 2.18.

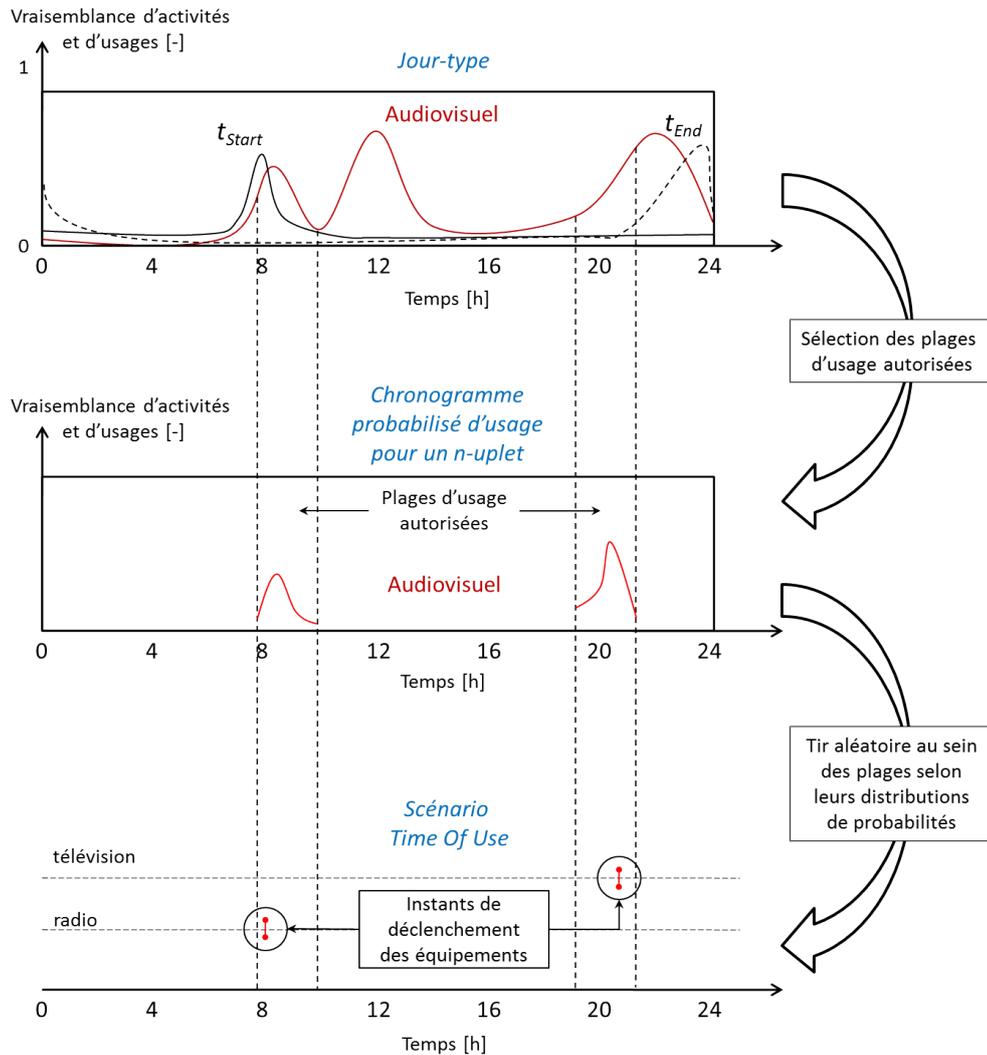


FIGURE 2.18 – Processus d'obtention d'un scénario Time Of Use pour un n-uplet donné

La construction de scénarios Time Of Use par équipement se fait tout en assurant leur compatibilité au sein d'un n-uplet. Des règles logiques fondées sur les paramètres de gestion des équipements-types<sup>56</sup> sont respectées lors de l'élaboration de ces scénarios. Ce sont notamment :

- une corrélation forte et évidente entre la mise en marche des équipements-types et la présence au domicile des occupants d'un ménage-type ;
- le respect du nombre de cycles journaliers ou hebdomadaires par équipement-type en fonction des caractéristiques du n-uplet ;
- le respect du **niveau de disponibilité\*** des membres d'un ménage-type :

*indicateur de mobilisation d'une personne à une activité domestique ou un usage donné. Il varie continûment entre 0 (la personne est occupée par une mono-tâche) et 1 (l'occupant est alors parfaitement disponible pour tout usage). De plus, la valeur -1 représente l'état de sommeil. Il est défini pour chaque membre d'un ménage-type reconnu responsable, c'est-à-dire pouvant agir sur les équipements domestiques électriques.*

Ceci se traduit par des interdictions de mise en marche simultanée d'un trop grand nombre d'équipements-types par rapport au nombre de personnes présentes au sein du logement à un instant donné<sup>57</sup> ;

- des séquences cohérentes de fonctionnement entre appareils<sup>58</sup>.

56. Voir le paragraphe « Caractériser les équipements en équipements-types » page 74

En plus de cette construction automatique des scénarios Time Of Use, l'utilisateur peut également en paramétrer de nouveaux en renseignant le nombre de cycles journaliers ou hebdomadaires et les instants de déclenchement de chaque équipement-type considéré au sein d'un n-uplet.

Une fois que l'aspect temporel (déclenchement de l'équipement-type) est fixé par le scénario Time Of Use, il faut renseigner le(s) mode(s) de fonctionnement retenu(s) à l'échelle journalière ce qui fixe la durée et l'amplitude du (des) cycle(s) unitaire(s) de puissance correspondant(s). Le lien entre mode de fonctionnement et cycle unitaire de puissance est établi au sein du modèle. Par exemple, la sélection du mode de fonctionnement lavage à 30 °C d'une machine à laver fait correspondre le *CUP* adéquat. Cependant, il doit demeurer la possibilité d'attribuer manuellement un cycle unitaire de puissance à un mode de fonctionnement donné. Celle-ci peut se faire selon deux processus :

- sélection d'un cycle unitaire dans une base de données ;
- paramétrage d'un nouveau cycle unitaire.

Pour simplifier la modélisation, nous ne tenons pas compte de la vétusté des appareils<sup>59</sup> et la modification éventuelle qui en découle sur leur(s) cycle(s) unitaire(s) de puissance. Ces derniers sont donc invariants tout au long de la durée de vie de chaque équipement-type.

Une fois que le(s) mode(s) de fonctionnement accompagné(s) du (des) cycle(s) unitaire(s) de puissance correspondant(s) ont été choisis, il reste à superposer cette information sur les scénarios Time Of Use. Une courbe de charge journalière unitaire d'un équipement-type est donc composée d'un (ou plusieurs) cycle(s) unitaire(s) de puissance « ancré(s) » à l'instant (aux instants) de mise en marche identifié(s) dans le scénario Time Of Use.

Itérer le processus pour chaque équipement-type d'électricité spécifique de chaque n-uplet du parc habité

L'opération de calcul décrite dans l'encadré précédent ne reflète que l'appel journalier de puissance d'un seul équipement-type pour un individu d'un n-uplet considéré et pour chaque jour-type. Afin de modéliser le parc habité, une phase d'itération est indispensable. Celle-ci doit être conduite sur trois plans :

- au niveau des équipements-types du même individu du n-uplet ;
- au niveau des individus du même n-uplet pour atteindre la convergence du foisonnement à l'échelle du n-uplet (voir la partie 2.3.3) ;
- au niveau de chaque n-uplet du parc habité.

### 2.3.2.2 Calculer les courbes de charge thermiques unitaires par n-uplet (F 3.2.2)

Les courbes de charge thermiques désignent celles des équipements-types des usages chauffage et climatisation (y compris leurs auxiliaires), ECS (auxiliaires éventuels inclus) et ventilation. À noter que pour leur calcul, nous distinguons deux méthodes distinctes : la première est appliquée aux appareils assurant le chauffage, la climatisation, leurs auxiliaires et la ventilation ; la seconde s'adresse exclusivement aux équipements assurant la production d'eau chaude sanitaire. En outre et quel que soit l'équipement thermique considéré, une phase d'itération est également requise pour prendre en compte l'ensemble des n-uplets du parc habité.

57. Il est en effet peu concevable d'avoir au sein d'un logement tous les appareils en fonctionnement de manière simultanée

58. Par exemple le sèche-linge ne peut être mis en marche qu'après le lave-linge

59. L'entartrage progressif d'une machine à laver en est un exemple

Dans les encadrés ci-dessous, nous détaillons les trois points – les deux méthodes de calcul de courbes de charge thermiques et la phase d’itération – qui doivent être examinés pour l’obtention des appels de puissance provenant des usages thermiques.

Calculer les courbes de charge unitaires par n-uplet pour le chauffage, la climatisation, leurs auxiliaires et la ventilation

Comme il l’a été fait pour les équipements-types répondant aux usages spécifiques de l’électricité, le calcul des courbes de charge thermiques doit être mené en deux temps :

1. construire une courbe de charge journalière par équipement-type des usages thermiques (hors ECS) d’un individu du n-uplet considéré en fonction de son comportement et pour chaque jour-type ;
2. itérer le processus pour chaque équipement-type des usages étudiés pour l’individu du n-uplet considéré.

Pour obtenir des courbes de charge des usages chauffage et climatisation incluant leurs auxiliaires ainsi que celles des équipements assurant la ventilation, nous faisons appel à un cœur de calcul thermique.

La modélisation en puissance des usages thermiques est menée en prenant en compte les **apports internes\*** [AD03] :

*il s’agit de dégagements de chaleur au sein d’un logement dont la source peut être d’origine humaine (métabolisme des occupants), technique (dissipation thermique des équipements-types, pertes de production d’ECS) ou bien liés à une activité domestique (préparation culinaire par exemple).*

D’autre part, chaque simulation est conduite :

- en fonction d’un niveau de **consigne spatio-temporelle\*** par équipement-type de chauffage/climatisation :

*il s’agit d’un paramétrage des températures de consigne qui se rapportent aux modes de gestion<sup>60</sup> des usages chauffage et climatisation et qui traduit le comportement du ménage-type considéré.*

- par **zone du bâtiment\*** :

*un bâtiment-type est constitué d’un ensemble de pièces rassemblées en plusieurs zones, a minima deux (bâtiment bi-zone). Celles-ci autorisent un réglage différencié des appareils de chauffage/climatisation entre une zone de pièces à vivre et une zone destinée au couchage par exemple.*

Les niveaux de consigne spatio-temporelle peuvent être choisis par l’utilisateur ou explicitement paramétrés.

En plus de la prise en compte des comportements « macroscopiques » des occupants au travers des paramètres renseignés au niveau des ménages-types<sup>61</sup>, l’influence de la structure tarifaire, qui a un impact potentiel sur l’utilisation individuelle du chauffage et de la climatisation, doit pouvoir être intégrée dans une version plus évoluée du présent modèle. Elle n’est toutefois pas intégrée dans le cadre du travail de thèse.

60. Définis au niveau du jour-type – voir le paragraphe 2.3.1.1 page 90

61. Voir le paragraphe 2.1.1.2 page 64

Calculer les courbes de charge unitaires d'eau chaude sanitaire par n-uplet

Le calcul des courbes de charge unitaires d'ECS requiert une méthode légèrement différente de celles qui s'appliquent aux autres usages domestiques<sup>62</sup>. Cependant comme il l'a été pour les autres usages, l'obtention de ces éléments suit un processus en deux étapes :

1. construire une courbe de charge journalière par équipement-type d'un individu du n-uplet considéré en fonction de son comportement et pour chaque jour-type ;
2. itérer le processus pour chaque équipement-type de l'usage considéré pour l'individu du n-uplet observé.

La première de ces deux opérations requiert le paramétrage de la simulation pour l'ECS ce qui revient à définir :

- un **profil de puisage\*** journalier et saisonnalisé<sup>63</sup> par n-uplet, pour chaque équipement-type et pour chaque jour-type. Le profil de puisage correspond au

*besoin en ECS exprimé en volume par unité de temps (le plus souvent en L/h) et délivré à une température de référence appelée **température de puisage\*** (par exemple 60 °C) ;*

- une température de puisage pour l'équipement-type considéré. Par convention, cette grandeur est donnée pour un point de mesure situé en sortie de ballon (ou de chaudière) pour le n-uplet étudié.

Les courbes de charge d'éventuels auxiliaires des systèmes d'ECS doivent être construites à ce niveau en fonction des profils de puisage précédemment définis.

Itérer le processus pour chaque n-uplet du parc habité

Les calculs précédemment conduits aboutissent aux courbes de charge unitaires des usages thermiques des équipements-types d'un seul individu d'un n-uplet du parc habité. Puisque ce dernier est constitué d'un ensemble de n-uplets, il faut répéter le processus pour un individu de chaque n-uplet du parc habité.

### 2.3.3 Calculer les profils de charge par n-uplet pour chaque jour de la simulation et pour la zone géographique étudiée (F 3.3)

L'obtention des courbes de charge unitaires modélise le foisonnement intra-ménage. Pour représenter le foisonnement inter-ménages, il faut tout d'abord itérer le processus mené au préalable sur les individus d'un même n-uplet aboutissant ainsi aux profils de charge des équipements-types pour ce n-uplet. Puis une seconde phase d'itération consiste à construire les profils de charge de chaque n-uplet constituant le parc habité.

#### 2.3.3.1 Construire par tirage aléatoire le profil de charge par équipement-type du n-uplet considéré pour chaque jour de la simulation (F 3.3.1)

En préambule de ce paragraphe, nous introduisons ci-dessous deux définitions supplémentaires :

- la **Courbe De Charge Moyenne Foisonnée\*** (*CDCMF\**) d'un équipement-type pour  $n$  individus correspond à

*la courbe de charge moyenne de cet appareil, tous les individus ayant été considérés. Son calcul consiste à sommer l'ensemble des courbes de charge unitaires des  $n$  individus et à diviser le résultat obtenu par  $n$  ;*

62. Équipements d'électricité spécifique d'une part et équipements des usages thermiques hors ECS d'autre part

63. De manière générale, le puisage d'ECS est plus conséquent en hiver qu'en été

– le **profil de charge\*** d'un équipement-type

*est la courbe de charge moyenne foisonnée de cet appareil, tous modes de fonctionnement confondus, pour un individu équipé d'un n-uplet lors d'un jour simulé donné.*

Un profil de charge s'obtient par un processus en cinq étapes :

1. générer et sommer  $N$  courbes de charge unitaires obtenues à partir de  $N$  scénarios Time Of Use sélectionnés et associés aux modes de fonctionnement correspondants. Une synchrone – somme de courbes de charge unitaires – est obtenue puis diviser cette dernière par  $N$  ;
2. obtenir  $N'$  courbes de charge unitaires à partir de  $N'$  scénarios Time Of Use sélectionnés et associés aux modes de fonctionnement correspondants ;
3. sommer les  $N + N'$  courbes de charge puis diviser par  $N + N'$  ;
4. comparer les deux courbes de charge moyennes obtenues au moyen d'un indicateur ou d'une distance adéquate<sup>64</sup> ;
5. itérer le processus jusqu'à satisfaction du critère de convergence  $\underline{\epsilon}$  ou l'atteinte du poids dans l'effectif du n-uplet dans le parc habité.

Ce mécanisme d'obtention d'un foisonnement satisfaisant est schématiquement représenté par la figure 2.19. Sur celle-ci,  $N$ ,  $N'$  et  $N''$  désignent des nombres génériques de courbes de charge unitaires nécessaires à une représentation satisfaisante du foisonnement. Les ordres de grandeur de ces nombres sont susceptibles de dépendre à la fois de l'usage, du jour-type et de l'équipement-type.

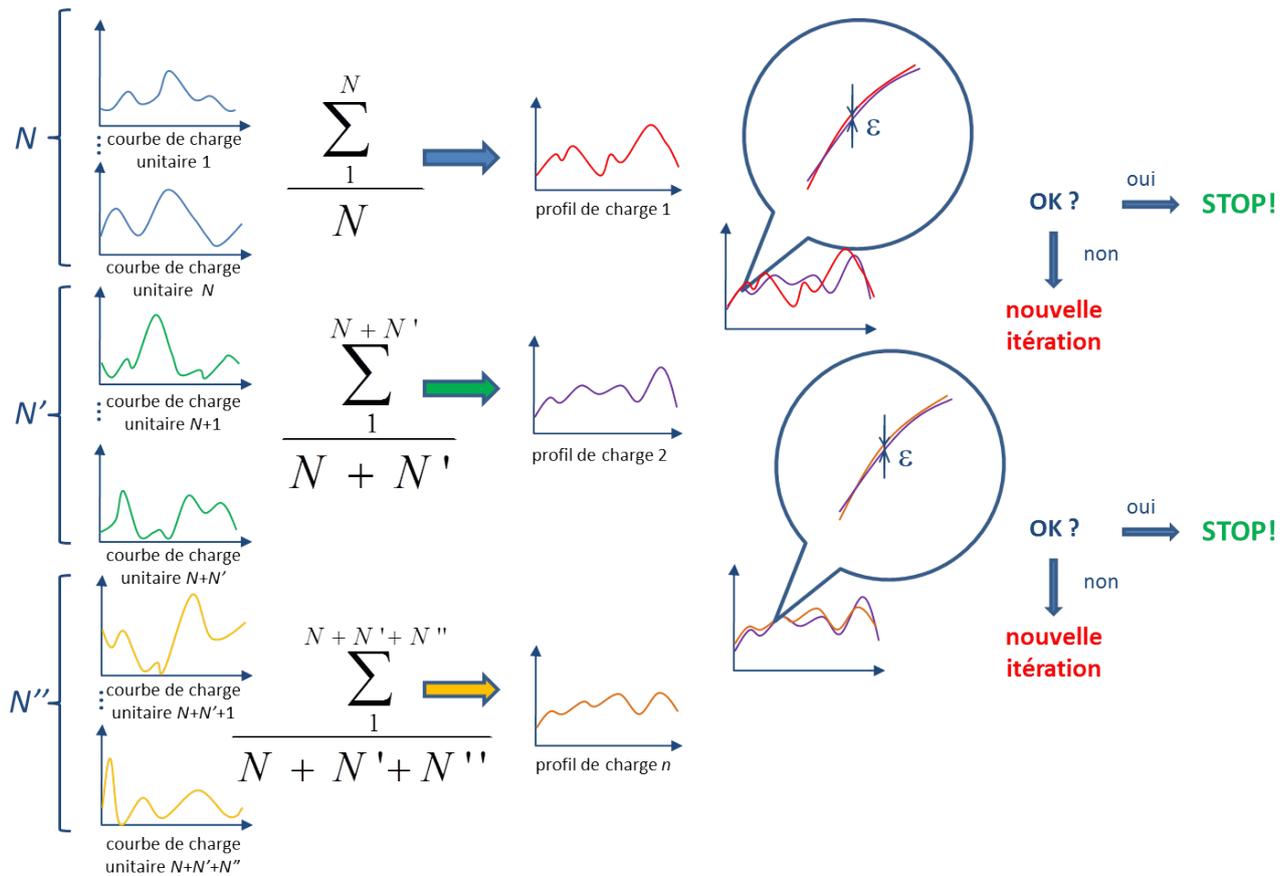


FIGURE 2.19 – Méthode d'obtention du profil de charge d'un équipement-type

64. Nous abordons plus largement ces éléments au sein de la partie 4.1 du chapitre 4

### 2.3.3.2 Itérer le processus pour chaque équipement-type d'un même n-uplet et pour l'ensemble des n-uplets du parc habité (F 3.3.2)

La fonction F 3.3.1 ne calcule que le profil de charge d'un seul équipement-type pour un n-uplet donné alors que l'ensemble des profils de charge doivent être obtenus pour modéliser le parc habité. Une fonction d'itération est donc rendue nécessaire de manière à boucler à la fois :

- sur les équipements-types d'un même n-uplet ;
- sur l'ensemble des n-uplets du parc habité.

La fonction F 3.3.2 remplit ce rôle de bouclage.

### 2.3.4 Calculer les courbes de charge à l'échelle de la zone géographique (F 3.4)

#### 2.3.4.1 Mettre à l'échelle les profils de charge de chaque n-uplet de la zone géographique étudiée en fonction de son poids dans l'effectif du parc habité pour chaque jour de la simulation (F 3.4.1)

La fonction F 3.3 fournit des profils de charge par équipement-type et par n-uplet sans toutefois tenir compte des poids dans l'effectif de chacun d'eux. De ce fait, un passage à l'échelle est nécessaire pour obtenir la courbe de charge équipement-type par équipement-type de l'ensemble des individus d'un même n-uplet du parc<sup>65</sup> de la zone géographique étudiée et pour chaque jour de la simulation.

Pour ce faire, il faut multiplier le profil de charge d'un équipement-type caractéristique d'un n-uplet par son poids dans l'effectif dans la zone géographique considérée. Par exemple, soit un parc habité de 1000 individus composé à 35 % par le n-uplet  $nu_1$ . La figure 2.20 représente la mise à l'échelle de la courbe de charge de l'équipement machine à laver pour le n-uplet  $nu_1$ .

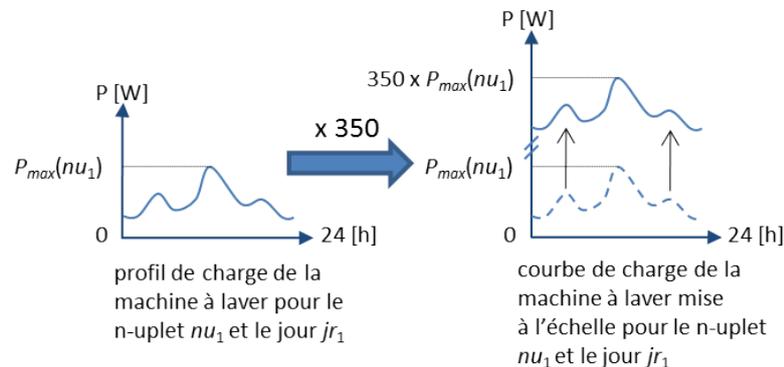


FIGURE 2.20 – Mise à l'échelle de la courbe de charge journalière de la machine à laver pour un n-uplet spécifié d'un parc habité

#### 2.3.4.2 Sommer par usage les courbes de charge d'un même n-uplet de la zone géographique étudiée et pour chaque jour de la simulation (F 3.4.2)

Pour répartir le solde fatal en énergie<sup>66</sup>, la courbe de charge d'électricité spécifique doit être disponible n-uplet par n-uplet. De manière plus générale, il faut pouvoir disposer de la courbe de charge par usage et par n-uplet pour chaque jour simulé : c'est le rôle de la fonction F 3.4.2. Celle-ci somme les courbes de charge issues de la fonction F 3.4.1 de manière à obtenir la courbe de charge par usage et par n-uplet de la zone géographique étudiée et pour chaque jour de la simulation.

Pour illustrer nos propos, nous présentons au travers de la figure 2.21 la sommation des courbes de charge d'un même n-uplet relatives à l'électricité spécifique (famille d'usage).

65. Chaque n-uplet est mis à l'échelle individuellement

66. Voir la fonction F 3.8.2 page 104

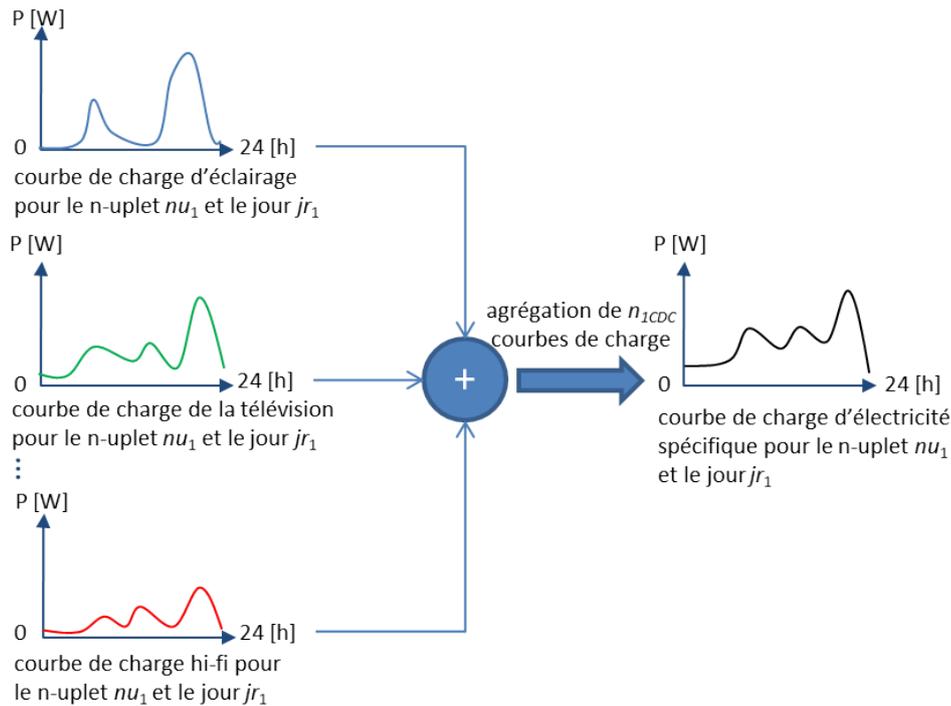


FIGURE 2.21 – Sommation des courbes de charge relatives à l'électricité spécifique pour un n-uplet donné

### 2.3.4.3 Sommer par usage les courbes de charge des n-uplets de la zone géographique étudiée et pour chaque jour de la simulation (F 3.4.3)

La fonction F 3.4.3 agrège les courbes de charge en sortie de la fonction F 3.4.2, et ce, en regroupant tous les n-uplets. L'opération effectuée par F 3.4.3 n'est qu'une simple somme de courbes de charge correspondantes à un même usage pour l'ensemble des individus de la zone géographique considérée et pour chaque jour simulé. Par exemple, dans le cas de la courbe de charge d'électricité spécifique, la fonction F 3.4.3 opère conformément au processus représenté par la figure 2.22.

### 2.3.4.4 Concaténer les courbes de charge par usage de l'ensemble des individus de la zone géographique étudiée sur toute la durée de simulation (F 3.4.4)

Les courbes de charge par usage de l'ensemble des individus du parc de la zone géographique étudiée et pour chaque jour de la simulation – obtenues au sortir de la fonction F 3.4.3 – sont ensuite concaténées sur toute la durée de la simulation. Concrètement les courbes de charge sont mises « bout à bout » dans l'ordre chronologique des jours simulés. Ce mécanisme est illustré par la figure 2.23.

### 2.3.5 Assurer la cohérence en énergie des courbes de charge par usage pour les équipements-types simulés de la zone géographique étudiée à l'année de calage (F 3.5)

Les courbes de charge calculées par le modèle doivent engendrer des consommations qui a minima sont du même ordre de grandeur que des données existantes mesurées et réalisées : la fonction F 3.5 assure ce rôle de contrôle de cohérence. Soulignons de plus que dans une première approche d'analyse fonctionnelle, nous avons pensé à définir deux fonctions – F 3.5 et F 3.6 – pour la vérification des résultats en énergie d'une part et en puissance d'autre part. Au terme de la réflexion, nous pensons que ces deux fonctions peuvent n'en former qu'une en choisissant par exemple un optimum de convergence sur ces deux critères. De surcroît, cette voie nous semble davantage cohérente du fait que des données réalisées en puissance à l'échelle du parc ne sont pour l'heure pas disponibles. Assurer la cohérence en puissance des courbes de

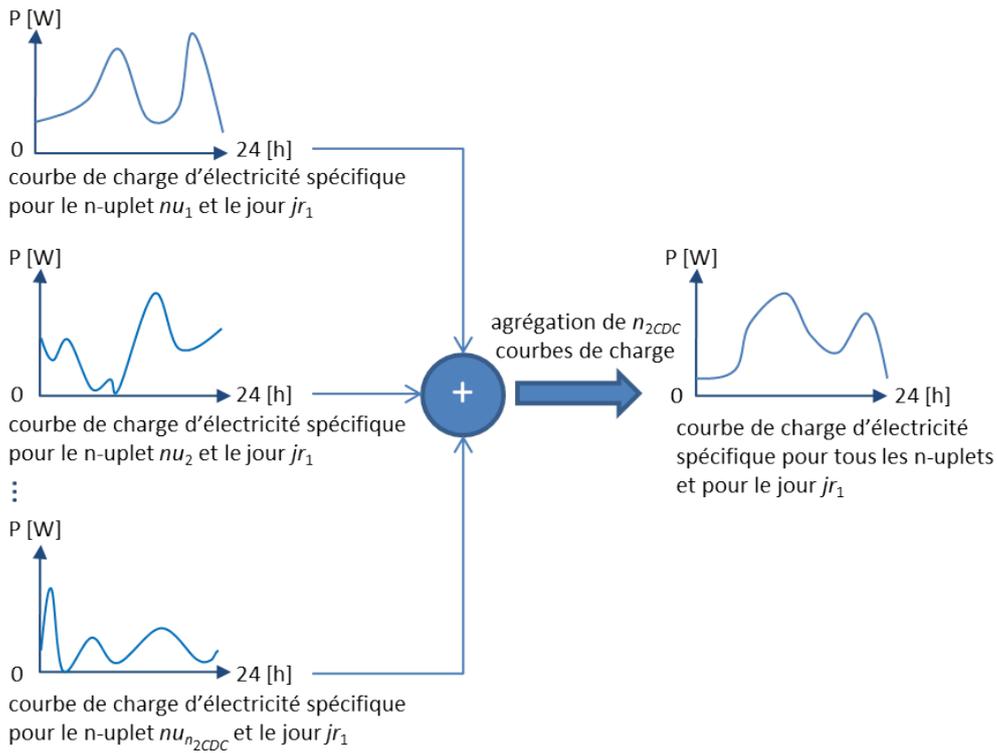


FIGURE 2.22 – Sommation des courbes de charge d’électricité spécifique tous n-uplets confondus

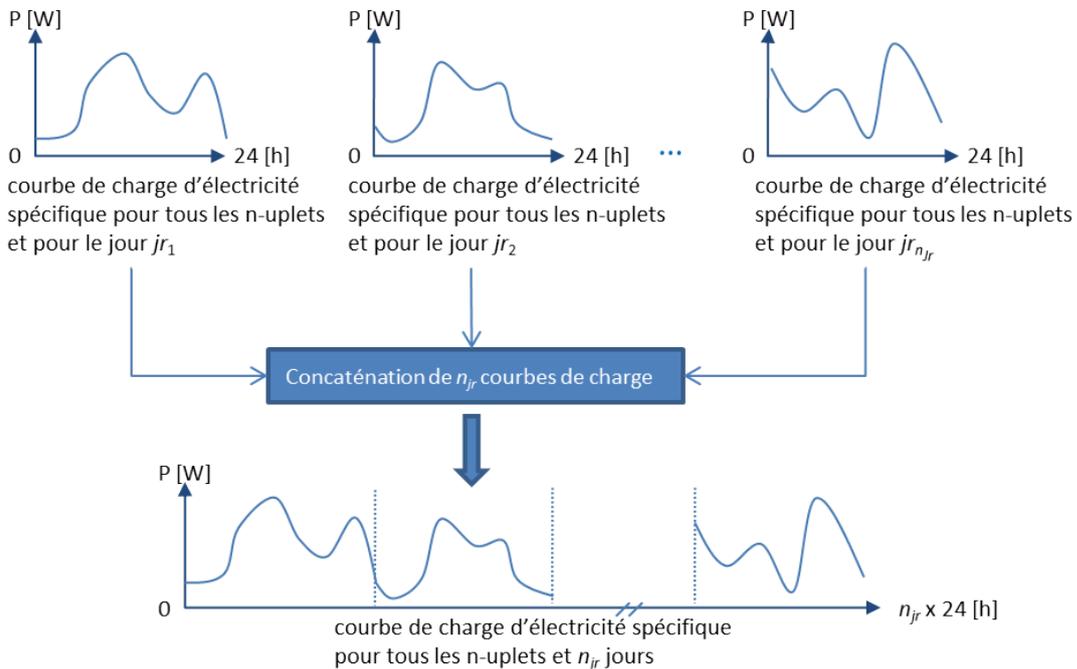


FIGURE 2.23 – Concaténation sur  $n_{jr}$  jours de la courbe de charge d’électricité spécifique tous n-uplets confondus

charge calculées n’est donc pas strictement réalisable actuellement.

Assurer la cohérence en énergie est réalisé à l’échelle du territoire et usage par usage pour plusieurs raisons :

- les valeurs mesurées qui servent de référence sont seulement disponibles à des échelles macroscopiques aussi bien du point de vue géographique que pour les usages considérés ;
- il est illusoire d’assurer la cohérence en énergie équipement-type par équipement-type puisqu’intrinsèquement nous ne modélisons pas toute la diversité du parc d’équipements domestiques existants ;
- nous considérons la consommation de l’ensemble des individus du parc habité modélisé indistinctement des n-uplets puisque la segmentation des valeurs de référence ne correspond pas à celle des n-uplets ;
- il n’est possible de parler de validation des résultats qu’avec des valeurs de référence mesurées. La validation n’est envisageable qu’à l’année de calage.

Cette correction en énergie des courbes de charge ne concerne que les équipements-types c’est-à-dire les appareils explicités.

La quantité d’énergie que nous ne modélisons pas, appelée solde fatal en énergie (cf. p. 73) et qui provient des équipements non-explicités, est quant à elle calculée et répartie en puissance respectivement par les fonctions F 3.7 et F 3.8.

Les opérations effectuées au sein de la fonction F 3.5 se décomposent en trois étapes. Chacune d’elles est prise en charge par une des sous-fonctions que nous détaillons dans les paragraphes ci-dessous.

#### 2.3.5.1 Calculer l’énergie correspondante aux courbes de charge obtenues (F 3.5.1)

Pour assurer la cohérence en énergie du modèle, il faut dans un premier temps calculer la consommation d’électricité engendrée par les courbes de charge simulées. Pour cela, il s’agit de calculer les intégrales de chaque courbe délivrée par la fonction F 3.4.4 de manière à aboutir à la quantité d’énergie consommée par usage pour l’ensemble des individus d’une même zone géographique et pour toute la durée de la simulation (année par année).

#### 2.3.5.2 Comparer à des références existantes (F 3.5.2)

Après avoir obtenu les valeurs de consommation électrique annuelles par usage issues de la modélisation du parc, il faut ensuite comparer ces quantités à des données directement disponibles ou bien déduites de consommations globales. Les sources de comparaison peuvent être nombreuses : CEREN\*, REMODECE\*, ADEME, panel de mesure... Dans tous les cas, il sera nécessaire d’obtenir des données par [poste de consommation\\*](#). Par ce terme, nous désignons un usage ou un regroupement d’usages domestiques de l’électricité. Par exemple, le CEREN distingue les postes suivants :

- chauffage ;
- froid ;
- lavage ;
- éclairage ;
- autres appareils d’électricité spécifique (appelé également solde).

En fonction des données de référence choisies et donc des postes de consommation correspondants, F 3.5.2 doit agréger différemment les consommations par usage issues de F 3.5.1.

#### 2.3.5.3 Converger vers la cible en énergie à l’année de calage selon le critère choisi (F 3.5.3)

Si les différences par poste de consommation déterminées par la fonction F 3.5.2 ne satisfont pas le critère de convergence en énergie choisi (en fonction de la simulation et/ou de la précision des valeurs de référence), il est nécessaire d’adapter les données d’entrée du modèle. Dans une première approche, nous pouvons stipuler que le nombre de leviers d’action (influence sur les paramètres du modèle par exemple) autorisés soit très restreint. Pour converger en énergie, il ne sera possible par exemple que de modifier les modes de fonctionnement des équipements-types en particulier le nombre de cycles journaliers ou hebdomadaires par mode.

### 2.3.6 Valider sur des données historiques les courbes de charge obtenues à l'année de calage (F 3.6)

Une fois la cohérence en énergie assurée grâce à des données de consommation pour l'année de calage, les courbes de charge obtenues doivent ensuite être validées en forme. Deux possibilités sont offertes a priori pour effectuer cette validation :

1. comparer les courbes de charge simulées avec des courbes de charge réalisées ;
2. calculer des indicateurs pour évaluer la validité du modèle.

Nous détaillons chacun de ces points dans les paragraphes ci-dessous.

#### 2.3.6.1 Comparer les courbes de charge simulées avec des courbes de charge réalisées (F 3.6.1)

Une première façon d'opérer pour la validation des courbes de charge simulées est de les comparer avec des données mesurées. Une telle comparaison cherche à identifier la bonne capacité du modèle à reconstituer, en plus d'une énergie cohérente avec des données de consommation, la forme des appels de puissance et leur variabilité infra- et inter-journalière. Cette fonction offre la possibilité à l'utilisateur de superposer plusieurs courbes de charge journalières émanant soit directement des résultats de la simulation soit d'une bibliothèque de courbes de charge issues de mesures. En outre, le modèle peut calculer des indicateurs appropriés.

#### 2.3.6.2 Calculer des indicateurs pour évaluer la validité du modèle (F 3.6.2)

Dans le littérature spécialisée de la courbe de charge, quatre indicateurs sont fréquemment rencontrés lorsqu'une comparaison « mathématique » de courbes de charge est requise. Il s'agit du *NVF*, du *Mean Absolute Percentage Error (MAPE)*\* [HON10], du *facteur de charge*\* de la puissance maximale [LAN85, LAN87, PAT89] et du *coefficient de foisonnement*\*. Déjà mentionnés au sein du chapitre 1, nous détaillerons davantage ces coefficients au sein de la partie 4.1.1 du chapitre 4.

D'autres coefficients non référencés ici peuvent s'avérer pertinents si bien que la liste ci-dessus est loin d'être exhaustive. D'autre part, des méthodes statistiques plus élaborées pourraient également être employées pour la validation des courbes de charge à l'année de calage (voir la partie 4.1 du chapitre 4).

### 2.3.7 Calculer le solde fatal en énergie à l'échelle du territoire (F 3.7)

Du fait de l'échelle géographique à laquelle les données de référence sont disponibles, la seule dimension où il est cohérent de calculer le solde fatal en énergie est celle de l'ensemble du territoire. En outre, cette quantité ne concerne que des équipements d'électricité spécifique pour lesquels l'influence de la saisonnalité est négligée.

Le solde fatal en énergie, notée *SFE*, est provoqué par les équipements non-explicités. Il se calcule en tant que différence entre la consommation d'électricité pour l'ensemble des usages spécifiques  $C_{elec. sp\acute{e}. tot.}$  (cette quantité d'énergie concerne donc à la fois les équipements explicités et non explicités) et la somme des consommations d'électricité des équipements-types des mêmes usages  $C_{elec. sp\acute{e}. \acute{e}q. exp.}$  (cette somme correspond à la consommation des équipements explicités) :

$$SFE = C_{elec. sp\acute{e}. tot.} - C_{elec. sp\acute{e}. \acute{e}q. exp.} \quad (2.1)$$

La manière de calculer le solde fatal diffère selon l'année simulée. Ainsi, nous différencions dans la suite de ce paragraphe l'année de calage, l'année initiale d'une période et n'importe quelle autre année. Dans tous les cas, le solde fatal en énergie est reconsidéré chaque année.

### 2.3.7.1 Calculer le solde fatal à l'année de calage (F 3.7.1)

Pour l'année de calage, la détermination du solde fatal est réalisée en deux temps :

1. calculer à l'échelle du territoire la consommation (après assurance de la cohérence en énergie) des équipements-types modélisés se rapportant aux usages d'électricité spécifique ;
2. calculer la différence entre la référence choisie et la consommation des équipements explicités précédemment calculée.

Après avoir assuré la cohérence en énergie des courbes de charge simulées à l'année de calage par l'intermédiaire de la fonction F 3.5, il est nécessaire de calculer la consommation résultante des équipements-types (donc explicités) d'électricité spécifique. Ce calcul s'effectue pour l'ensemble des zones géographiques. Concrètement, il suffit de calculer les intégrales des courbes de charge correspondantes<sup>67</sup> et d'additionner les consommations par poste.

La consommation précédemment calculée doit ensuite être comparée à la valeur de référence retenue au sein de la fonction F 3.5.2. La différence obtenue est ensuite traduite en courbe de charge par la fonction F 3.8.

### 2.3.7.2 Calculer le solde fatal à l'année initiale de chaque période (F 3.7.2)

L'année initiale d'une période est la première année simulée après un événement-clé. Ce dernier a pour effet de faire évoluer les flux au sein du parc habité et/ou le solde fatal en énergie. Si une rupture fait évoluer la liste des équipements-types explicités d'électricité spécifique, il est nécessaire de recalculer le solde fatal en énergie à l'année initiale de la période qui débute. Pour cela, il faut tout d'abord récupérer la valeur du solde fatal de l'année précédente et défalquer (ou ajouter) l'énergie correspondante à l'usage de l'équipement-type nouvellement explicité (ou retiré du parc, le cas échéant).

### 2.3.7.3 Calculer le solde fatal pour une année de la période (F 3.7.3)

Pour une année simulée qui n'est ni l'année de calage ni l'année initiale de la période considérée, le calcul du solde fatal en énergie fait appel au scénario d'évolution de cette grandeur défini par la fonction F 1.2 et de la valeur du solde à l'année initiale de la période en cours.

## 2.3.8 Répartir le solde fatal en énergie par zone géographique (F 3.8)

Le solde fatal en énergie calculé par la fonction F 3.7 doit être intégré, sous forme de courbe de charge, aux résultats validés en énergie et en puissance. Bien qu'elle soit calculée à l'échelle du territoire, cette consommation peut être considérée à la maille d'une zone géographique puisqu'elle est répartie n-uplet par n-uplet conformément à leur propre consommation d'électricité spécifique.

### 2.3.8.1 Répartir le solde fatal en énergie par n-uplet (F 3.8.1)

Pour donner une forme en puissance au solde fatal en énergie par n-uplet, nous proposons une méthode simple. Celle-ci consiste à répartir la quantité d'énergie imputable à chaque n-uplet et pour chaque jour simulé suivant :

- la courbe de charge journalière d'électricité spécifique correspondante pour une zone géographique donnée, ou bien
- un profil de charge normé exogène.

Dans le premier cas, il faut réduire les courbes de charge calculées par la fonction F 3.4.2 par la valeur de leur intégrale de manière à obtenir une [courbe de charge normée](#)\* c'est-à-dire une courbe de charge

---

<sup>67</sup>. Du fait de l'aléa autour de ce solde fatal, nous pouvons systématiquement utiliser les courbes de charge par usage délivrées par la fonction F 3.4.4 plutôt que de calculer des courbes de charge par équipement-type pour l'ensemble des individus à l'échelle du territoire et pour la durée totale de la simulation

qui engendre une consommation unitaire. Exprimer le solde fatal en énergie par n-uplet sous forme de puissance consiste alors à multiplier la courbe de charge normée par la participation du n-uplet au solde fatal en énergie au pro rata de sa consommation d'électricité spécifique, et ce, jour par jour.

Dans le second cas, il faut répartir le solde suivant le profil exogène normé défini par l'utilisateur ce qui offre notamment la possibilité de distribuer préférentiellement cette énergie à des heures spécifiées (aux heures de pointe par exemple).

### 2.3.8.2 Répartir le solde fatal de la zone géographique par jour-type (F 3.8.2)

F 3.8.1 distribue le solde fatal en énergie n-uplet par n-uplet. D'autre part, le parc habité étant modélisé à l'échelle de la zone géographique, les poids dans l'effectif des n-uplets sont connus à cette échelle. La part de solde fatal imputable à un individu d'un n-uplet, calculable à l'échelle du territoire, est une grandeur qui est également valable pour n'importe quelle zone géographique étudiée puisque la saisonnalité n'intervient pas dans le solde fatal en énergie<sup>68</sup>. La répartition du solde fatal par zone géographique découle directement de cette caractéristique.

De plus, il est possible d'envisager une répartition du solde fatal par jour-type. En effet, l'utilisation des appareils non explicités varie potentiellement selon les mêmes habitudes de consommation que les équipements-types. La manière la plus exacte de répartir cette consommation pour chaque jour simulé est de le faire au pro rata des consommations journalières d'électricité spécifique par n-uplet.

La courbe de charge finale du modèle s'obtient par sommation des courbes de charge modélisées dont la cohérence en énergie est assurée et du solde fatal en énergie exprimé en puissance par n-uplet et par jour-type.

### 2.3.9 Sommer les courbes de charge usage par usage à l'échelle du territoire (F 3.9)

L'ensemble des opérations effectuées par la fonction F 3 est réalisé à l'échelle de chaque zone géographique alors que les résultats escomptés sont ceux qui correspondent au parc habité à l'échelle du territoire. Une fonction d'agrégation des courbes de charge par usage pour l'ensemble des zones géographiques est donc nécessaire. C'est le rôle de la fonction F 3.9. Cette dernière ne fait que sommer des courbes de charge calculées au préalable par les sous-fonctions de F 3.

68. Voir le descriptif de la fonction F 3.7 page 102

## 2.4 Définition de la fonction finale « Restituer à l'utilisateur les courbes de charge journalières du parc habité (F 4) »

F 4 est une fonction basique de tout modèle délivrant des résultats. En effet, comme l'indique la figure 2.24, il s'agit ici d'offrir la possibilité à l'utilisateur de sélectionner le(s) format(s) de restitution et les dimensions spatiale(s) et temporelle(s) des courbes de charge modélisées.

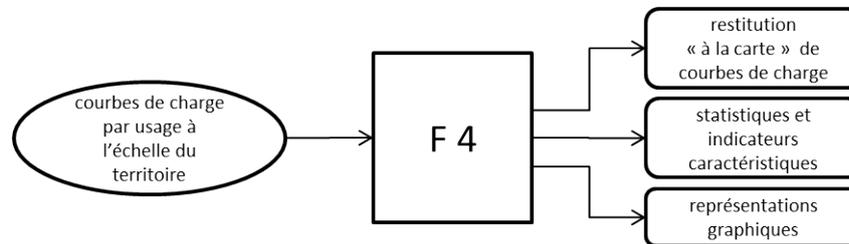


FIGURE 2.24 – Schéma d'entrées/sorties de la fonction F 4

### 2.4.1 Sélectionner le(s) format(s) de restitution (F 4.1)

En fonction des simulations réalisées, l'utilisateur peut souhaiter différents formats de restitution des résultats du modèle. A priori, le plus intéressant est d'obtenir des courbes de charge :

- par usage ;
- par poste de consommation ou famille d'usage<sup>69</sup> ;
- tous usages confondus.

F 4.1 agrège donc les résultats en sortie de la fonction F 3 selon le(s) format(s) de restitution sélectionnés.

### 2.4.2 Sélectionner les échelles spatiale(s) et temporelle(s) de restitution (F 4.2)

Les courbes de charge calculées par la fonction F 3 sont restituées à l'échelle du territoire et pour toute la durée de la simulation. Cependant, comme nous l'avons vu au travers de la fonction F 3.4, différents résultats intermédiaires sont calculés par le modèle si bien que les échelles de restitution spatiale(s) et temporelle(s) peuvent être sélectionnées par l'utilisateur.

Ainsi, il est envisageable de souhaiter des résultats pour un espace géographique :

- d'une seule zone ;
- constitué d'une sélection de zones au sein du territoire ;
- pour l'ensemble du territoire.

Les courbes de charge peuvent être également voulues par bâtiment-type et/ou par ménage-type.

De même, le choix du type de concaténation des courbes de charge journalières par usage peut présenter un intérêt certain. L'utilisateur peut notamment vouloir les regrouper et les visualiser :

- par jour-type ;
- par semaine ;
- par mois ;
- par saison ;
- par année.

### 2.4.3 Calculer des indicateurs et construire des représentations caractéristiques (F 4.3)

En fonction de l'intérêt des simulations effectuées par le modèle, différents indicateurs, paramètres et/ou représentations graphiques peuvent être intéressants pour analyser les résultats de l'étude. Ainsi, nous listons ci-dessous une série d'éléments que la fonction F 4.3 doit pouvoir fournir :

69. Par exemple chauffage, climatisation, ECS et électricité spécifique

- le **gradient thermique de consommation énergétique\***  $G$  exprimé en W/ °C qui traduit la variation de l'appel de puissance  $\Delta P$  d'un parc en fonction de la variation de la température extérieure  $\Delta T$  :

$$G = \frac{\Delta P}{\Delta T} \quad (2.2)$$

- la monotone de puissance consommée <sup>70</sup> ;
- le facteur de charge ;
- le coefficient de foisonnement ;
- des éléments de statistique descriptive :
- puissance maximale sur un intervalle  $\Delta t$ ,  $P_{max}(\Delta t)$  ;
  - puissance moyenne sur le même intervalle  $\Delta t$ ,  $\overline{P}(\Delta t)$  ;
  - variance ( $\sigma^2$ ) et écart-type ( $\sigma$ ) des courbes de charge obtenues ;
  - quantile, mode de puissance modélisée...

---

70. Rappel : nous avons détaillé cette construction dans l'introduction générale de la thèse pour le cas de la production française d'électricité

## Conclusion du chapitre

Dans ce second chapitre de thèse, nous avons présenté les résultats des travaux d'analyse fonctionnelle que nous avons menés au cours de nos recherches. Concrètement pour aboutir à la méthode nouvelle de reconstitution de courbes de charge proposée ici, nous avons adopté une démarche itérative de réflexions et de questionnements. Partant d'une feuille blanche, et ayant en toile de fond les spécifications fonctionnelles essentielles du modèle, nous avons d'abord cherché à définir avec précision le périmètre de l'étude. L'étape suivante a consisté à identifier, parmi l'ensemble des influences pouvant avoir un impact sur les appels domestiques de puissance, quelles étaient les plus prégnantes. Nous avons sélectionné un certain nombre d'entre elles avec comme critère de choix le rapport entre le gain du modèle en précision en regard de la complexification de celui-ci engendrée par la prise en compte de cette influence. En définitive, nous avons pu constater que vouloir synthétiser la courbe de charge du secteur résidentiel ne pouvait s'envisager sans penser simultanément, en plus de l'évidente dimension technique/technologique, à la fois à l'habitat, à son environnement et au très influent facteur humain.

Le modèle décrit dans ce chapitre introduit une série de concepts qui lui sont propres et qui vont au-delà de ceux que nous avons rencontrés lors de notre analyse bibliographique (cf. chapitre 1). De plus, la méthode proposée a été voulue générique c'est-à-dire qu'elle demeure parfaitement applicable quel que soit le parc domestique considéré. Les spécificités de ce dernier doivent en revanche être introduites dans la mise en œuvre opérationnelle de la méthode. Le chapitre 3 développe justement nos choix en matière de mise en pratique des concepts définis dans cette partie. Les concepts ne sont pas utilisables directement dans une simulation c'est pourquoi une représentation mathématique de ceux-ci doit être proposée. De la sorte, ils deviennent manipulables et peuvent être employés au sein d'algorithmes comme nous le verrons dans le prochain chapitre.



# Chapitre 3

## Modélisation des concepts et mise en œuvre par voie algorithmique

*« Nous ne raisonnons que sur des modèles. »*

Paul Valéry

### Introduction du chapitre

À la suite des développements conceptuels que nous avons présentés au sein du chapitre 2, nous avons cherché à mettre en pratique notre méthode de reconstitution de courbe de charge. Pour cela, il nous a fallu dans un premier temps dériver des concepts précédemment définis, des objets adaptés à une implémentation informatique : concrètement il s'est agi de les modéliser. Nous présentons dans la partie 3.1 la manière avec laquelle nous avons décidé de représenter les principaux concepts de notre méthode que sont les jours-types, les ménages-types et les équipements-types. De nouveaux concepts se sont révélés nécessaires et ont été introduits dans le modèle.

Une fois adaptés les éléments-types de la méthode, nous avons ensuite dû mettre en place les procédés et les techniques nécessaires pour les faire interagir et ainsi aboutir au résultat escompté : la courbe de charge électrique. Étant donné l'ampleur des travaux à mener, nous avons poursuivi la modélisation par ordre de priorité. Reconnu comme étant le point dur principal de la thèse, la génération explicite du foisonnement d'origine comportementale a occupé une très large place dans notre travail. Ce foisonnement, facilement observable à une échelle agrégée sur des mesures réelles de puissance électrique, est le fruit de nombreuses influences et d'interactions complexes qui le rendent difficilement modélisable de manière explicite, c'est-à-dire sans appliquer de façon systématique sur des courbes de charge simulées des traitements mathématiques (par exemple de type déformation statistique) dont l'origine réside dans la connaissance sectorielle (celle-ci trouvant notamment sa source dans les campagnes de mesure).

Générer de la diversité comportementale en matière de consommation électrique consiste en priorité à faire en sorte que l'utilisation des équipements au sein d'un parc simulé de logements soit semblable à la réalité. Cette utilisation est notamment caractérisée par un ensemble de fluctuations à différentes échelles temporelles : annuelle, saisonnière, mensuelle, hebdomadaire, journalière, horaire et infra-horaire.

Le travail de modélisation a abouti à un algorithme stochastique qui cherche à assurer le déclenchement réaliste des équipements domestiques d'électricité spécifique au sein d'une journée donnée. Nous décrivons ce processus stochastique séquentiel dans la partie 3.2.

Enfin dans la section 3.3, nous proposons une mise en perspective de notre algorithme stochastique. Il s'agit à la fois de faire la synthèse de notre travail et ensuite de le comparer aux autres méthodes de génération de diversité d'origine comportementale que nous avons trouvées dans la littérature.

### 3.1 Modélisation des éléments conceptuels de la méthode de reconstitution de courbe de charge

Dans cette partie, nous exposons et expliquons nos choix en matière de modélisation pour la reconstitution des appels de puissance électrique du secteur résidentiel français. Parmi l'ensemble des concepts introduits au chapitre précédent, les plus importants d'entre eux sont reproduits au travers de la figure 3.1. Cette illustration représente le schéma-blocs de notre méthode de reconstitution de courbe de charge. Sur ce graphique, nous pouvons visualiser que le calcul des appels de puissance<sup>1</sup> ne constitue que l'étape finale d'un processus plus complexe. Les différentes phases<sup>2</sup> en amont de cette reconstitution de la puissance ont pour principale vocation de reproduire les multiples sources de diversité qui affectent largement les besoins instantanés en électricité des logements.

Avant d'entrer dans les détails, nous précisons que **nous ne traiterons pas dans ce chapitre (et dans le reste du manuscrit) de la modélisation thermique du bâtiment et de ses systèmes énergétiques qui lui assurent chauffage, climatisation, ventilation et eau chaude sanitaire.**

En effet, comme nous l'avons souligné dans l'introduction générale de ce manuscrit, nous avons fait le choix de nous concentrer sur les équipements d'électricité spécifique. D'une part, ceux-ci engendrent une consommation électrique supérieure à celle des usages thermiques et d'autre part ils synthétisent les principales difficultés en vue de l'élaboration d'un modèle de génération de foisonnement.

En outre, nous ne décrivons dans cette partie qu'une des versions possibles de la modélisation des concepts de notre méthode générique de reconstitution de courbe de charge. Cette modélisation ne représente tout d'abord qu'une partie des concepts abordés au chapitre 2. En outre, elle ne correspond qu'à un point d'étape intermédiaire : un certain nombre de concepts de la méthode n'ont pas été modélisés. D'autre part, les choix de modélisation que nous présentons au sein de ce chapitre n'ont aucun caractère définitif. Les modèles peuvent donc être enrichis par la suite.

Dans la partie 3.1.1, nous détaillerons la modélisation des jours-types qui s'appuie largement sur les *intervalles infra-journaliers\**, concept que nous définirons. La modélisation des ménages-types sera abordée au sein du paragraphe 3.1.2. Enfin, nous détaillerons dans la partie 3.1.3 la modélisation des équipements d'électricité spécifique vers laquelle nous avons concentré nos efforts.

#### 3.1.1 Modélisation des jours-types

Jusqu'à présent, le fonctionnement de la plupart des appareils consommateurs d'électricité qui se trouvent dans les habitations, est directement dépendant de la présence et/ou de l'intervention humaine. Modéliser le besoin en puissance électrique d'un ménage implique nécessairement de s'intéresser tout particulièrement aux périodes de la journée où le foyer est occupé par des personnes éveillées, c'est-à-dire disposées à faire l'usage d'équipements domestiques (la nuit et les périodes de vacance<sup>3</sup> du logement font appel à un traitement différencié). Partant de ce constat, nous avons cherché à identifier les moments ou les périodes d'une journée qui conditionnent en partie la présence des occupants au sein de leur logement.

Dans la modélisation que nous avons choisie, une journée est segmentée de façon générale en 6 intervalles infra-journaliers qui sont les suivants<sup>4</sup> : *Réveil\**, *Matin\**, *Midi\**, *Après-midi\**, *Soir\** et *Nuit\**. Ces périodes sont définies, dans un premier temps, à l'échelle du *ménage agrégé\**, c'est-à-dire au niveau d'un foyer considéré comme une unité indivisible (non prise en considération de manière individuelle des membres qui le composent). Dans la suite, elles pourront être repérées par leur indice noté  $\iota$  (*Réveil*  $\equiv$  1, ..., *Nuit*

1. Pour mémoire, ces calculs sont assurés par la fonction F 3 de notre modèle conceptuel

2. Dans la méthode conceptuelle, ces étapes préliminaires sont prises en charge, rappelons-le, par les fonctions F 1 et F 2

3. Le logement peut être inoccupé à cause :

- des absences journalières occasionnelles liées aux activités ménagères, associatives, sociales, administratives, de loisirs. . .
- des absences journalières régulières causées par l'activité professionnelle, scolaire ou étudiante ;
- des absences prolongées lors des périodes de congés.

4. L'utilisation de majuscules nous permettra de différencier ces intervalles infra-journaliers de la signification habituelle de ces noms communs

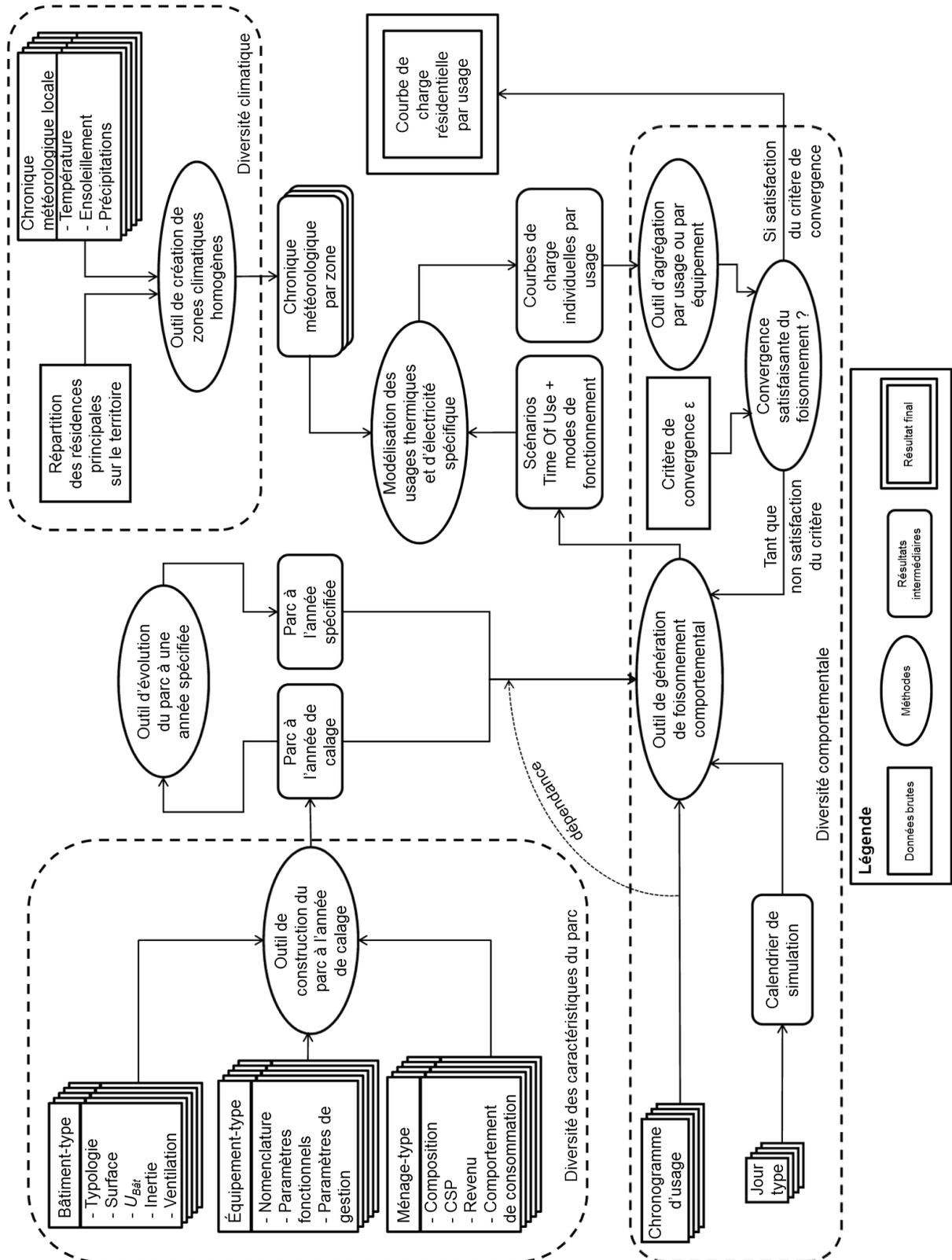


FIGURE 3.1 – Schéma-blocs du modèle de courbe de charge développé

≡ 6). Elles sont délimitées par des **instants caractéristiques**\* notés  $t_\alpha$  de manière générique. Ceux-ci sont détaillés<sup>5</sup> ci-dessous :

- $t_{Start}$  : début des activités domestiques au sein d'un foyer spécifié au cours d'un jour  $n$  simulé. Cet instant coïncide avec l'heure du lever des membres du foyer c'est-à-dire au début de l'intervalle Réveil ;
- $t_1$  : fin de l'intervalle Réveil (début de l'intervalle Matin). Cet instant correspond au départ au travail (ou à l'école) des occupants du foyer. À partir de  $t_{Start}$ , il survient après un intervalle de temps noté  $\Delta t_{Départ}$  ;
- $t_2$  : début de l'intervalle Soir (fin de l'intervalle Après-midi). Ce moment se calcule en ajoutant à  $t_1$  la durée journalière d'absence notée  $\Delta t_{Abs.Jour.}$  des individus du ménage (coïncide avec l'instant de retour du premier occupant) ;
- $t_3$  : début de l'intervalle Midi (fin de l'intervalle Matin). Ce moment correspond à l'heure méridienne de retour au domicile, le cas échéant, des occupants du ménage ;
- $t_4$  : fin de l'intervalle Midi (début de l'intervalle Après-midi). Cet instant est calculé en ajoutant à  $t_3$  la durée de présence méridienne au domicile notée  $\Delta t_{Prés.Mér.}$  ;
- $t_{End}$  : fin des activités domestiques au sein du ménage correspondant à l'heure du coucher des membres du foyer (fin de l'intervalle Soir, début de l'intervalle Nuit) ;
- $t_{StartBis}$  : début des activités domestiques au sein du foyer simulé pour le jour  $n + 1$  (fin de l'intervalle Nuit du jour  $n$ ).

Deux précisions importantes sont à souligner : tout d'abord et **dans une première approche, nous nous plaçons au niveau du ménage agrégé** c'est-à-dire que les  $t_\alpha$  sont ceux qui correspondent à l'emploi du temps du premier occupant levé/rentre et du dernier membre du foyer couché/parti. À l'heure de la rédaction de la thèse, la modélisation au niveau du ménage agrégé a toutefois laissé la place à celle au niveau de l'individu (voir la partie 3.2.3) de manière à améliorer le réalisme de la représentation des ménages dans notre modèle. Concrètement **dans cette seconde approche, l'ensemble des  $t_\alpha$  doit être défini pour chaque individu du ménage considéré susceptible de faire l'usage des équipements domestiques**. Dans le cas particulier du ménage composé d'un seul individu, l'approche ménage agrégé est bien entendu confondue à celle de l'approche individu par individu.

D'autre part, **la simulation d'un jour  $n$  ne correspond pas nécessairement à une durée de 24 h**. En effet, les jours calendaires (de 0 h 00 à 23 h 59) présentent des difficultés potentielles de modélisation :

- la nuit est constituée d'un seul ou de deux intervalles disjoints en fonction des habitudes de vie des ménages (couchage après minuit par exemple) ;
- les équipements déclenchés en fin de journée se reportent potentiellement sur la journée suivante. Il faut donc être capable d'assurer la continuité en puissance du fonctionnement des équipements entre deux jours calendaires consécutifs.

**Pour ces raisons, nous avons choisi de considérer la nuit d'un seul bloc et donc de simuler, i. e. d'opérer le placement\* des équipements, entre deux  $t_{Start}$  consécutifs.**

Les 6 intervalles infra-journaliers construits à partir des instants caractéristiques sont consécutifs les uns aux autres et de durées respectives non nulles. De ce fait, un contrôle de cohérence est nécessaire. Pour chaque jour-type, la relation 3.1 :

$$t_{Start} < t_1 < t_3 < t_4 < t_2 < t_{End} < t_{StartBis} \quad (3.1)$$

doit être vérifiée. Ceci constitue une première spécification en matière de modélisation des instants caractéristiques.

La reconstitution de la courbe de charge électrique d'un parc résidentiel impose la prise en compte de diversité qu'elle soit d'origine technique, comportementale, socio-économique, climatique ou autre.

5. Les explications proposées partent de l'hypothèse que le jour considéré est un jour travaillé

Comme nous l'avons vu au travers de la fonction F 3.2.1 <sup>6</sup>, le calcul des courbes de charge d'électricité spécifique – ce qui in fine nous intéresse le plus particulièrement – passe par la simulation répétée d'un même scénario qui renferme une variabilité potentiellement forte.

Les instants caractéristiques  $t_\alpha$  reproduisent la première source de diversité d'origine comportementale au travers de la représentation de l'emploi du temps journalier des membres d'un foyer. La variabilité de ces éléments constitue donc une seconde spécification essentielle. Ainsi, afin de favoriser un maximum de diversité, la sélection des  $t_\alpha$  doit varier d'une part selon différents scénarios mais aussi d'autre part au sein d'un même scénario. Concrètement, pour un couple [jour-type, ménage-type] donné, les valeurs des  $t_\alpha$  évoluent d'une simulation à une autre. Pour cela, nous avons choisi de les définir au moyen de tirages aléatoires.

Les deux spécifications préalablement décrites nous ont donc orientés vers la modélisation suivante des instants caractéristiques. Chaque  $t_\alpha$  est tiré aléatoirement entre deux bornes temporelles de variation (borne inférieure notée  $\underline{inf} t_\alpha$  et borne supérieure notée  $\underline{sup} t_\alpha$ ) paramétrées par le modélisateur <sup>7</sup>. Entre ces limites, il est possible de sélectionner la distribution <sup>8</sup> suivant laquelle les tirages aléatoires sont effectués. La figure 3.2 illustre nos propos des paragraphes précédents et plus particulièrement elle résume le découpage d'un jour-type.

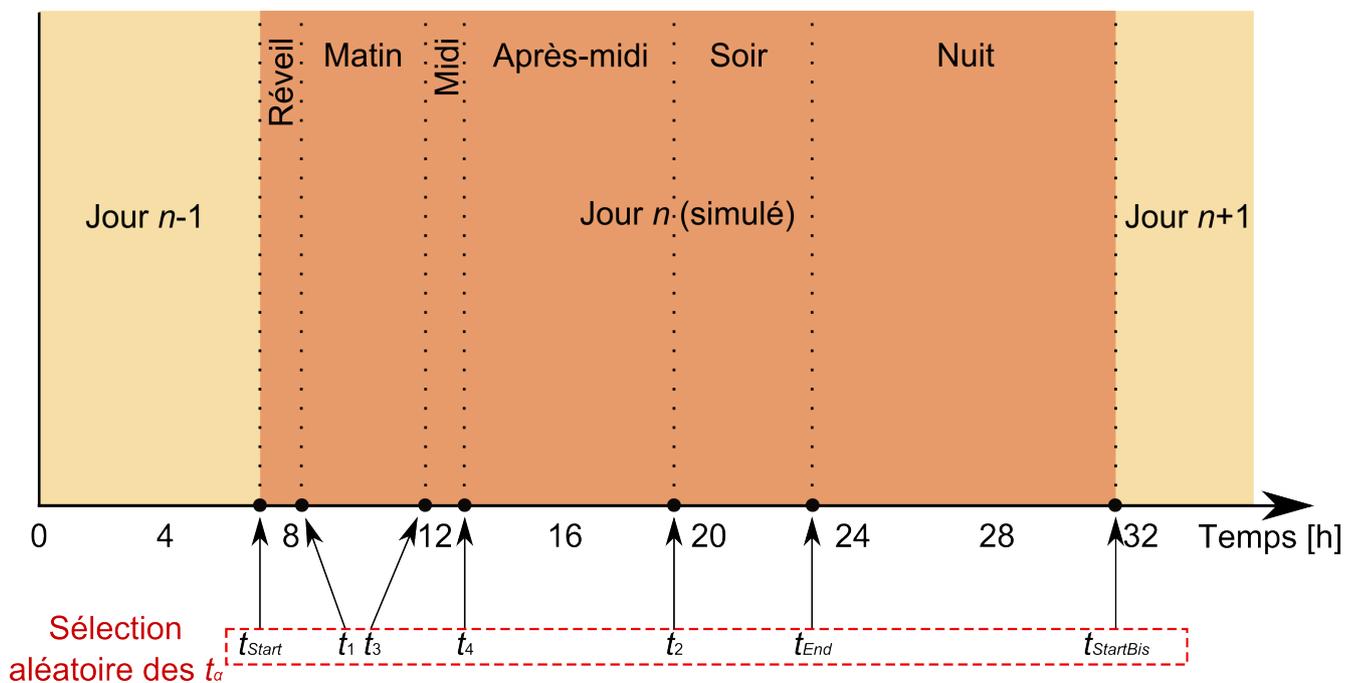


FIGURE 3.2 – Segmentation d'un jour-type en intervalles infra-journaliers

Lorsque nous étudions les comportements humains, des périodicités peuvent être remarquées. En effet, les échelles temporelles du jour, de la semaine, du mois... structurent le travail, les loisirs, les temps physiologiques... et forgent ainsi nos habitudes de vie. Ces périodicités sont également retrouvées lorsque l'on observe des courbes de charge électrique, et ce, quel que soit le secteur (résidentiel, tertiaire, industriel...) considéré. Plus précisément pour le secteur domestique, la différenciation à minima entre jours de semaine (du lundi au vendredi) et week-end (samedi et dimanche) est pertinente <sup>9</sup> du fait des activités

6. Voir la section 2.3.2.1 du chapitre 2

7. Des valeurs par défaut sont toutefois proposées

8. Par défaut, nous avons choisi des distributions uniformes de probabilité entre les couples des bornes temporelles de variation des  $t_\alpha$

9. L'étude dans le détail de courbes de charge du secteur résidentiel montre l'existence de jours de semaine « similaires » : les mardis, jeudis et vendredis sont comparables entre eux alors que les lundis et les mercredis présentent leurs propres

professionnelles et des rythmes scolaires qui impactent largement la forme et l'amplitude des courbes de charge électriques.

Dans le modèle générique du jour-type, nous introduisons donc une première variable intitulée `TypeJour` dont les deux premières modalités sont `Semaine` et `Weekend`. De plus, nous définissons une modalité supplémentaire à cette variable et plus précisément le mode `Absence` qui traduit la vacance du logement de ses occupants.

En ce qui concerne la différenciation entre jour de semaine et week-end, nous pensons que cette influence se traduit par des valeurs différentes des bornes de variation des  $t_\alpha$  et plus spécifiquement les valeurs hautes, et ce, de manière à augmenter l'étendue des intervalles dans lesquels les instants caractéristiques sont sélectionnés. La modification des bornes supérieures de variation des  $t_\alpha$  se fait automatiquement dans le modèle en fonction de la modalité de la variable `TypeJour` (`Semaine/Weekend`). En ce qui concerne la modalité `Absence`, celle-ci n'a pas d'impact sur les valeurs des bornes des  $t_\alpha$ . En revanche, elle impose un certain **profil de disponibilité\*** comme nous le verrons dans le paragraphe 3.2.2.

Pour finir, soulignons que c'est au niveau du jour-type qu'est opéré le choix du paramètre qui renseigne la résolution temporelle de la simulation c'est-à-dire le pas de temps auquel les appareils électriques pourront être placés dans la journée et les courbes de charge calculées. Nous appelons ce paramètre `ResolutionCDC` (résolution courbe de charge). Définir un pas de temps de calcul revient à chercher une solution optimale entre le réalisme de la simulation qui est directement lié au(x) **temps caractéristique(s)\*** du (des) système(s) simulé(s), c'est-à-dire l'ordre de grandeur de leur durée de fonctionnement, et le temps de simulation. De ce fait, nous avons autorisé de mener des simulations jusqu'au pas 1 min (le choix 1 min est celui par défaut) puisque cette durée nous semble caractéristique de la majorité des équipements domestiques que nous modélisons. De plus, dans le cadre d'une simulation journalière, nous rappelons que le placement des équipements s'effectue entre deux  $t_{Start}$  consécutifs et la durée de simulation, notée  $d_{Simulation}$  (`DureeSimulation`), est une grandeur qui demeure paramétrée par le modélisateur. Celle-ci doit donc être en accord avec  $inf t_{StartBis}$  et  $sup t_{StartBis}$ .

Suite au choix des instants caractéristiques et à la sélection de la résolution temporelle de la simulation, la construction des intervalles infra-journaliers dans lesquels seront déclenchés les appareils domestiques est relativement immédiate. En effet, la modélisation des intervalles infra-journaliers s'appuie sur les éléments suivants :

- $t_{Initial}$  : début de l'intervalle infra-journalier ;
- $t_{Final}$  : fin de l'intervalle infra-journalier ;
- $l_{Intervalle}$  : longueur de l'intervalle ;
- une **forme développée d'intervalle\*** qui correspond à un tableau unidimensionnel de taille  $l_{Intervalle}$  dont chaque élément représente le niveau de disponibilité du ménage considéré. Construit en début de simulation, ce tableau est ensuite mis à jour tout au long de la durée de celle-ci ;
- un **résumé d'intervalle\*** comprenant les **événements\***, couples [instants de déclenchement, **signature d'équipement\***], qui ont lieu au sein de l'intervalle et qui sont classés chronologiquement. Tout comme la forme développée de l'intervalle, cet élément évolue dynamiquement au cours des calculs de courbes de charge. Par ailleurs, le résumé d'intervalle est directement utilisable pour la construction des appels de puissance des équipements domestiques.

Nous proposons au travers des tableaux 3.1 et 3.2 respectivement la représentation de la forme développée d'un intervalle et de son résumé.

Nous détaillerons à la section 3.2.2 l'étape qui relie les intervalles infra-journaliers aux profils de disponibilité, c'est-à-dire l'affectation du niveau de disponibilité du ménage dans chacun des intervalles infra-journaliers.

---

spécificités [AEE08]

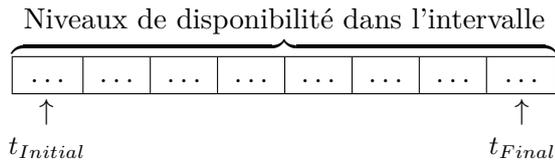


TABLEAU 3.1 – Exemple d'intervalle développé

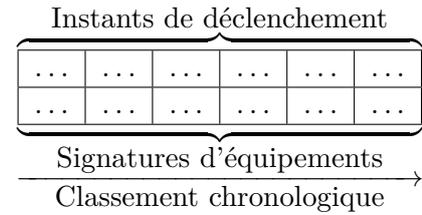


TABLEAU 3.2 – Exemple de résumé d'intervalle

### 3.1.2 Modélisation des ménages-types

Comme nous l'avons indiqué en préambule de cette partie, la mise en œuvre des concepts que nous présentons ici n'est ni exhaustive ni complète : il s'agissait en effet dans un premier temps de construire une modélisation même partielle afin de tester notre modèle de reconstitution de courbe de charge. Ainsi, initialement lorsque nous considérons le ménage-type à un niveau agrégé la seule variable implémentée dans le modèle du ménage-type fut `TypeMenage`. Cette dernière traduit le statut d'activité du ménage et elle présente trois modalités décrites ci-après :

- **Actif** : les occupants ont une activité professionnelle et/ou scolaire dans la journée. Ils sont donc absents de leur domicile entre  $t_1$  et  $t_2$  ;
- **ActifDomicile** : les membres du ménage ont également une occupation en journée mais ils reviennent (au moins l'un d'entre eux) chez eux pour déjeuner au cours de la pause méridienne ;
- **Inactif** : cette catégorie correspond aux personnes sans emploi ou retraités qui demeurent au domicile entre  $t_{Start}$  et  $t_{End}$ .

Considérer le ménage à un niveau agrégé est une approche dont les limitations sont substantielles. En effet, cette représentation du ménage-type n'opère pas de distinction en matière de potentiel de disponibilité entre des ménages de taille différente. Concrètement, l'utilisation simultanée d'un certain nombre d'appareils domestiques au sein d'un logement, qui est une situation quotidienne chez les ménages (notamment constitués de plusieurs personnes), est très mal modélisée par cette méthode. En outre, les absences occasionnelles (non liées à des activités professionnelles et scolaires) ne sont pas considérées dans cette représentation : ainsi, il n'est pas prévu par exemple que les occupants des ménages de type **Inactif** puissent quitter leur domicile.

Conscients de ces limites, nous avons recherché une manière plus évoluée de représenter les ménages-types. Dans une seconde approche, nous avons donc décidé de modéliser les ménages en tenant compte de chacun de leurs membres. Nous introduisons donc dans le modèle du ménage-type, une première variable supplémentaire intitulée `TailleMenage` qui correspond au nombre de personnes constituant le foyer modélisé. Sachant que nous gardons à l'esprit notre objectif de modéliser la puissance électrique requise par les appareils domestiques lors de leur utilisation, nous devons définir qui, au sein du ménage, peut mettre en marche ces appareils. De manière plus pragmatique, il faut définir au sein des foyers le nombre de personnes ayant un **niveau de responsabilité**\* suffisant pour déclencher des équipements. Dans le modèle, la variable correspondante est un entier nommé `NbPersonnesAResponsabilite`. Dans un premier temps, le niveau de responsabilité est représenté par un codage binaire (niveau de responsabilité = 0  $\equiv$  impossibilité de déclencher des appareils domestiques, niveau de responsabilité = 1  $\equiv$  autorisation d'utiliser l'ensemble des appareils domestiques sans restriction). Un codage plus fin avec des niveaux intermédiaires<sup>10</sup> peut être envisagé mais nous n'avons pas approfondi nos recherches en la matière.

Pour fixer le niveau de responsabilité des individus du ménage, des règles simples peuvent être suivies comme par exemple n'attribuer un niveau de responsabilité compatible avec l'utilisation de tous les appareils ménagers qu'aux individus du ménage âgés de plus de 15 ans. Dans notre modèle, c'est l'utilisateur qui définit le nombre d'éléments à responsabilité du ménage.

10. Par exemple, ceux-ci pourraient représenter le fait qu'un individu d'un foyer ne soit autorisé à faire l'usage que d'une sélection d'appareils domestiques

Ceci étant fait, la modélisation du ménage au niveau de l'individu impose de définir le statut d'activité de chaque membre du foyer à savoir *Actif*, *ActifDomicile* ou *Inactif*. Ces modalités ont la même signification que lorsqu'elles s'appliquaient au ménage considéré de manière agrégée. En outre, pour améliorer le réalisme de la modélisation, nous définissons une série de variables pour représenter les absences infra-journalières autres que celles liées aux activités professionnelles, étudiantes ou scolaires. Soulignons que dans notre modélisation, ces absences journalières occasionnelles ne concernent que les personnes du ménage à responsabilité et de type *Inactif* (nous aborderons cette question plus en détails au sein du paragraphe 3.2.3).

Pour prendre en considération ces absences occasionnelles, nous introduisons tout d'abord, un *seuil d'absence\** – *SeuilAbsence* – qui doit être renseigné. Celui-ci correspond à la probabilité journalière qu'au moins une absence ait lieu pour une personne du foyer à qui ce type d'absence peut être attribué. Ensuite, la durée totale de cette catégorie d'absence journalière est tirée aléatoirement par une loi uniforme<sup>11</sup> entre deux bornes inférieure et supérieure choisies par l'utilisateur<sup>12</sup>. Une fois la durée totale d'absence sélectionnée, celle-ci est répartie entre un certain nombre de périodes infra-journalières durant lesquelles la personne considérée n'est pas chez elle. Le nombre d'absences journalières de ce type est également tiré aléatoirement par une loi uniforme entre deux bornes haute et basse sélectionnées par le modélisateur. Malgré le fait que cette seconde approche soit plus élaborée que la première, il subsiste toujours des limites à la modélisation. Par exemple, nous pouvons noter le fait que cette représentation du ménage-type ne prévoit pas qu'un ménage puisse s'absenter en soirée : la modélisation impose d'être chez soi au moment du Réveil et du Soir. De plus, celle-ci n'assure pas la répétitivité du comportement des membres d'un ménage lors d'une succession de plusieurs jours simulés de même type.

### 3.1.3 Modélisation des équipements domestiques d'électricité spécifique

Dans cette partie, nous traitons en particulier de la manière avec laquelle nous avons représenté les équipements domestiques d'électricité spécifique sous forme de modèles. Deux constatations immédiates s'imposent pour mieux juger de la difficulté de cet exercice. Tout d'abord, il est bien perceptible qu'au sein d'un logement donné, les appareils électriques présents sont caractérisés par leur extrême diversité aussi bien en matière d'usage (ou de service rendu), de *type de fonctionnement\**, d'appels de puissance que de caractéristiques techniques (forme, taille, volume, capacité...) ou encore de paramètres de gestion<sup>13</sup>. D'autre part, lorsque l'on considère tout ou partie des logements présents sur un territoire, une hétérogénéité supplémentaire affecte les équipements disponibles sur le parc observé du fait de la diversité présente sur le marché. Il est ainsi statistiquement peu probable que deux foyers disposent exactement des mêmes appareils électriques à un instant donné (à moins de le vouloir sciemment) et cette probabilité devient nulle dès lors que l'on considère ce jeu d'équipements dans la durée.

De surcroît, la manière avec laquelle sont utilisés les équipements domestiques influe largement sur leurs appels de puissance. Celle-ci est éminemment complexe puisqu'elle englobe à la fois les choix des utilisateurs, la temporalité de l'usage des appareils, les habitudes et les modes de vie des foyers...

**L'enjeu de la modélisation des équipements domestiques est donc double : reproduire l'hétérogénéité d'ordre technique/technologique mais aussi et surtout la diversité d'origine comportementale.** Sur ce dernier point, nous avons développé un algorithme reproduisant le déclenchement réaliste, autant que faire se peut, des usages domestiques d'électricité spécifique que nous décrirons dans la partie 3.2 de ce chapitre. Nous détaillons ci-après les éléments techniques que nous avons pris en considération pour constituer une certaine hétérogénéité des appareils modélisés et leurs impacts en termes d'appels de puissance.

Malgré leur grande diversité, les équipements d'électricité spécifique peuvent être regroupés en différentes familles selon les critères observés que nous nommons *filtres\**. De cette manière, cela permet d'éviter la modélisation exhaustive de chacun d'eux qui constitue une tâche sans fin et donc de se concentrer sur

11. Ce choix peut bien sûr être modifié

12. Des valeurs par défaut sont toutefois proposées

13. Se reporter au paragraphe « Caractériser les équipements en équipements-types » page 74

une **modélisation générique\*** par famille d'équipements, modélisation qui s'applique par héritage aux appareils respectifs de chacun des groupes constitués.

Comme nous l'avons indiqué plus haut, un grand nombre de paramètres peuvent différencier les équipements domestiques entre eux. Cependant du point de vue de la modélisation des appareils ménagers et en vue de la simulation de leurs appels de puissance, tous ces critères n'ont pas le même intérêt. Ainsi, les filtres qui à notre avis demeurent pertinents pour la modélisation sont :

- le type d'usage fourni : le lavage, la préparation culinaire, l'audiovisuel, le froid domestique en sont des exemples ;
- le type de fonctionnement : à cycle de fonctionnement, thermostaté, à états ;
- l'**opportunité technique de pilotage\*** : interruptible, reportable, adaptable en puissance ou pas de pilotage possible ;
- l'**acceptabilité humaine du pilotage\*** : forte, moyenne, faible, inenvisageable.

Selon nous, le choix le plus logique dans ce travail de modélisation est le filtre du type de fonctionnement puisque c'est celui qui influence les appels de puissance des appareils de manière prépondérante vis-à-vis des autres. Dans la partie 3.1.3.1, nous détaillons la modélisation générique que nous avons mise en place et qui est appliquée aux équipements à cycle de fonctionnement et aux appareils à états. En outre, nous ferons référence dans les paragraphes 3.1.3.2 et 3.1.3.3 aux adaptations apportées à la modélisation générique conformément aux spécificités de chacun de ces types d'appareils.

Les particularités des équipements thermostatés (exemple du froid domestique) nous ont imposé de créer une modélisation distincte de la précédente. Nous présentons celle-ci dans la partie 3.1.3.4.

### 3.1.3.1 Modélisation générique des équipements d'électricité spécifique

À l'image de la modélisation des intervalles infra-journaliers, nous avons conçu une représentation des équipements domestiques qui lui est compatible. Pour y aboutir, nous avons d'abord dressé la liste des caractéristiques nécessaires à formaliser. En effet, un équipement est défini par :

- des caractéristiques générales telles que sa signature, définie ci-après et notée  $\underline{s}_x$ , ses éventuels **équipements liés\*** (y compris lui-même) et le **report nécessaire\*** ou non de l'utilisation de l'équipement en fonction du scénario modélisé ;
- des caractéristiques intrinsèques relatives à la **disponibilité requise\*** pour utiliser l'équipement. Nous définissons pour cela cinq périodes notées  $\underline{d}_\beta$  qui encadrent l'utilisation de l'équipement dans sa temporalité ;
- des caractéristiques extrinsèques relatives à la **compatibilité de l'équipement\*** face aux intervalles infra-journaliers. Nous verrons dans le paragraphe 3.2 comment sont établies ces caractéristiques.

La signature de l'équipement, **SignatureEquipment**, est une variable numérique qui rassemble les informations nécessaires pour reconnaître l'équipement assurant un certain usage et la durée de ce dernier. La **durée d'usage\*** d'un équipement  $i$ , notée  $\underline{d}_{Usage,i}$ <sup>14</sup>, est codée par un paramètre entier **DureeUsage**. Ce dernier s'exprime en nombre de pas de temps de calcul selon la résolution de courbe de charge choisie. Parmi les autres caractéristiques de base de la modélisation générique des équipements, nous retrouvons les paramètres fonctionnels « classiques » suivants : puissance nominale, puissance de veille, capacité, niveau de charge. À cela s'ajoutent des éléments de la nomenclature : nom de l'équipement, nom de l'équipement-type, mode de fonctionnement désiré. Enfin, de manière à tenir compte au plus près des apports internes occasionnés par le fonctionnement de l'équipement, la pièce d'accueil de l'appareil doit être renseignée dans le modèle.

14. Dans la suite, lorsqu'il ne sera pas nécessaire de désigner un équipement particulier (considération d'un équipement de façon générale), nous ne ferons pas systématiquement l'usage de l'indice  $i$ . Ainsi, la durée d'usage d'un équipement quelconque sera notée  $\underline{d}_{Usage}$ . Nous ferons de même pour certaines autres variables intervenant dans la modélisation des équipements ou dans celle des intervalles

Le déclenchement d'un équipement assurant un usage donné peut être conditionné par le fonctionnement :

- du même équipement : nous parlons d'**équipement auto-lié**\* ;
- d'un ou plusieurs équipement(s) domestique(s) distinct(s) de lui : dans ce cas, il s'agit d'équipements liés.

Les modalités « lié » et « auto-lié » constituent ce que nous appelons les **liaisons**\* possibles entre équipements.

Pour être dit auto-lié, un équipement doit présenter les caractéristiques suivantes : d'une part l'appareil concerné n'est présent qu'en un seul exemplaire dans le logement simulé et d'autre part l'utilisation multiple de cet équipement ne peut se faire que séquentiellement. Dans une majorité de ménages, la machine à laver est un exemple d'équipement auto-lié.

Parmi les nombreux exemples d'équipements liés, le sèche-linge fonctionnant généralement une fois que le cycle de la machine à laver est terminé ou le lecteur DVD qui n'est le plus souvent utilisé que lorsque la télévision est en marche, sont des plus représentatifs et perceptibles au quotidien. Ainsi, la temporalité des liaisons possibles entre équipements peut prendre deux modalités : **lien simultané**\* ou **lien séquentiel**\*. Le dernier cas suggère deux « sous-modalités » évidentes : avant/après.

De manière à simplifier la modélisation, nous définissons le concept d'**équipement de tête**\*. Ce sont eux qui déterminent la temporalité de fonctionnement de leurs équipements liés. Par exemple, pour la liaison machine à laver/sèche-linge l'équipement de tête est le premier de ces deux appareils. Après recensement des différentes liaisons entre équipements et identification des équipements de tête, seule la liaison séquentielle « après » demeure à considérer. D'autre part dans la modélisation, nous limitons à 1<sup>15</sup> le nombre d'équipements séquentiellement liés après un autre<sup>16</sup> mais nous ne fixons pas de limites aux équipements simultanément liés.

Sur un échantillon de population donné et du fait de nos habitudes de vie, la probabilité de mise en marche d'un équipement n'est pas constante au fil de la journée. Plus précisément, le fonctionnement de la majorité des appareils est davantage favorisé ou exclu<sup>17</sup> au sein de certaines périodes infra-journalières. Des stratégies particulières d'usage de ces équipements existent effectivement<sup>18</sup> et elles doivent donc être prises en compte dans notre modélisation. Pour cela, nous avons introduit six<sup>19</sup> variables binaires, notées  $AD_{\iota}$  ( $\iota \in [1; 6]$ ), dans le modèle générique d'équipement (une par intervalle infra-journalier). Chacune d'elles renseigne l'**autorisation de dépassement**\* de l'équipement par rapport à un intervalle infra-journalier spécifié, autrement dit la permission pour l'équipement de voir son utilisation s'étendre de l'intervalle infra-journalier accueillant son instant de déclenchement à l'intervalle infra-journalier (aux intervalles infra-journaliers) qui lui est (sont) postérieur(s).

La nécessité de report de l'usage d'un appareil domestique tient compte à la fois des autorisations de report d'un intervalle à un autre mais aussi des éventuelles liaisons entre l'équipement considéré et les autres appareils.

Pour réaliser un usage, la présence et l'action de l'Homme sont bien souvent requises. Cette intervention potentielle avant, pendant et/ou après l'activation d'un équipement se traduit dans notre modélisation par cinq périodes  $d_{\beta}$  caractérisées par une durée  $d_{\beta, Durée}$  et un niveau de disponibilité  $d_{\beta, Disp}$ . Elles sont détaillées ci-dessous :

- $d_1$  : période d'un besoin éventuel de disponibilité avant le déclenchement de l'équipement ;
- $d_2$  : attente avant le déclenchement de l'appareil ;
- $d_3$  : période de fonctionnement de l'équipement assurant l'usage ;

15. Cette limite peut bien évidemment être paramétrée et modifiée

16. Il s'agit d'une simplification de la modélisation mais pas d'une réduction de la réalité puisqu'un enchaînement de plusieurs équipements liés séquentiellement « après » peut être généré par notre modélisation

17. Par exemple, l'utilisation de l'aspirateur durant la nuit est une activité qui nous semble marginale

18. Déclenchement de différents appareils avant la nuit par exemple pour bénéficier d'un tarif d'heures creuses

19. `AutorisationDepassementReveil`, `AutorisationDepassementMatin`, `AutorisationDepassementMidi`, `AutorisationDepassementApresMidi`, `AutorisationDepassementSoir`, `AutorisationDepassementNuit`

- $d_4$  : attente après le fonctionnement de l'appareil ;
- $d_5$  : période d'un besoin éventuel de disponibilité après le fonctionnement de l'équipement.

La durée d'usage est obtenue en sommant les  $d_{\beta, Durée}$  ( $\beta \in [1; 5]$ ).

Ces périodes doivent être renseignées en durée et en niveau de disponibilité requise : pour ce dernier paramètre, la modélisation adoptée est une représentation binaire : 0 équivaut à un besoin nul de disponibilité alors que 1 signifie le besoin de l'intervention d'un individu à « 100 % » ce qui correspond dans ce cas à la mono-tâche. Notons toutefois que des règles logiques ont été implémentées pour assurer une cohérence dans la modélisation des équipements :

1. la variable  $d_{1, Durée}$  (respectivement  $d_{5, Durée}$ ) est conditionnée par  $d_{1, Disp}$  (respectivement  $d_{5, Disp}$ ) de sorte que la durée de cette période ne peut être positive qu'à la condition que le niveau de disponibilité correspondant le soit également ( $\forall \beta \in \{1; 5\}, d_{\beta, Disp} = 0 \Rightarrow d_{\beta, Durée} = 0$ ) ;
2. la variable  $d_{2, Durée}$  (respectivement  $d_{4, Durée}$ ) est conditionnée par  $d_{1, Durée}$  (respectivement  $d_{5, Durée}$ ) de sorte que la durée de cette période ne peut être positive qu'à la condition que la durée de la période qui la précède (respectivement qui la suit) soit non nulle ( $d_{1, Durée} = 0 \Rightarrow d_{2, Durée} = 0$  et  $d_{5, Durée} = 0 \Rightarrow d_{4, Durée} = 0$ ) ;
3. les périodes d'attente sont caractérisées par un besoin nul de disponibilité ( $d_{2, Disp} = 0$  et  $d_{4, Disp} = 0$  quel que soit l'équipement considéré).

L'utilisation d'un appareil domestique peut être modélisée sous deux formes : la [forme condensée de l'usage d'un équipement](#)\* qui correspond à une matrice  $2 \times 5$  comme l'indique l'exemple présenté dans le tableau 3.3 (détaillé ci-après) et la [forme développée de l'usage d'un équipement](#)\* qui par défaut se construit directement de la précédente représentation en « dilatant » chaque période  $d_{\beta}$  comme l'indique le tableau 3.4.

	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$
$d_{\beta, Durée}$ [min]	3	0	55	0	8
$d_{\beta, Disp}$ [-]	1	0	0	0	1

TABLEAU 3.3 – Exemple de forme condensée de l'usage d'un équipement

Le tableau 3.3 présente la forme condensée de l'usage d'un équipement ; supposons qu'il s'agisse d'une machine à laver et que la résolution de courbe de charge choisie soit égale à 1 min. Dans cet exemple, l'usage du lave-linge requiert 3 min de disponibilité ( $d_{1, Durée} = 3$  avec  $d_{1, Disp} = 1$ ) avant le déclenchement de l'appareil (ici le remplissage du tambour de la machine à laver). Le délai d'attente avant fonctionnement  $d_{2, Durée}$  est nul (lancement de l'équipement une fois le remplissage fini). Le cycle dure 55 min ( $d_{3, Durée} = 55$ ). Pendant cette période, l'intervention humaine n'est pas nécessaire :  $d_{3, Disp} = 0$ . Le retrait des vêtements lavés est conduit sitôt la fin du cycle ( $d_{4, Durée} = 0$ ). Cette dernière opération, qui requiert la disponibilité à 1 d'une personne, dure 8 min ( $d_{5, Durée} = 8$  avec  $d_{5, Disp} = 1$ ).

1	1	1	0	...	<sup>20</sup>	0	1	1	1	1	1	1	1	1	1	1	1	1	1
---	---	---	---	-----	---------------	---	---	---	---	---	---	---	---	---	---	---	---	---	---

TABLEAU 3.4 – Exemple de forme développée de l'usage d'un équipement

Pour uniformiser la représentation des équipements et notamment la forme développée de leur usage, nous adoptons les règles suivantes :

- l'usage d'un équipement dont la durée est inférieure au pas de temps de la simulation (`ResolutionCDC`) sera représenté par une unique case contenant un 0 puisqu'il est « invisible » d'un point de vue disponibilité requise ;

20. Dans ce cas suite de 55 zéros au total dans la forme développée de l'usage du lave-linge

- la forme développée de l’usage de tout équipement tel que  $d_{Usage} > \text{ResolutionCDC}$  débutera par un 1<sup>21</sup>. Pour cela, des tests sont effectués au sein de la fonction du modèle qui génère la forme développée à partir de la forme condensée. Plus précisément, ces tests sont conduits sur la valeur de  $d_{3,Disp}$  :
  - si  $d_{3,Disp} = 1$ , la construction de la forme développée de l’usage s’opère par simple développement de sa forme condensée (construction par défaut) puisque dans ce cas, il est certain que la forme développée de l’usage de l’équipement commence par au moins un 1 ;
  - si  $d_{3,Disp} = 0$ , la valeur de  $d_{1,Durée}$  est testée : si  $d_{1,Durée} = 0$ , on la fixe égale à 1 tout comme  $d_{1,Disp}$  (de sorte que la forme développée de l’usage commence par au moins un 1) et la suite de la construction de la forme développée se poursuit identiquement à la manière par défaut. De la même façon si  $d_{1,Durée} > 0$ , rien n’est modifié dans la construction par défaut de la forme développée de l’usage de l’équipement.

Enfin un équipement se caractérise par sa compatibilité vis-à-vis des intervalles infra-journaliers. Dans une première approche, nous choisissons une représentation binaire pour la modéliser : 0  $\equiv$  non compatible, 1  $\equiv$  compatible. Cette information, qui se présente sous forme d’un tableau de taille 6 intitulé [tableau de compatibilité équipement/intervalles\\*](#) est construite<sup>22</sup> à l’instant initial de la simulation et mise à jour tout au long de la phase de calcul en fonction des autres équipements déjà déclenchés le jour considéré. En outre, au sein de la modélisation de l’équipement, nous introduisons le concept d’[intervalle limite\\*](#) – `IntervalleLimite` – qui correspond au dernier intervalle infra-journalier dans lequel un équipement peut être utilisé lors d’un jour donné. Pratiquement il s’agit de l’indice du premier élément valant 1 dans le tableau de compatibilité équipement/intervalles parcouru dans l’ordre inverse des indices des intervalles infra-journaliers.

Pour affiner la notion de compatibilité face aux intervalles infra-journaliers, chaque équipement reçoit trois informations supplémentaires venant des autres éléments de la modélisation :

- l’indice de l’[intervalle à considérer\\*](#) noté  $\iota_{IAC}$  (`IndiceIntervalleAConsiderer`) ;
- le [repère de l’équipement à considérer\\*](#) noté  $r_{EAC}$  (`RepereEquipementAConsiderer`) ;
- la [durée à considérer\\*](#) notée  $d_{AC}$  (`DureeAConsiderer`).

Nous verrons dans la partie 3.2 comment sont construites ces informations.

Pour finir, en ce qui concerne la modélisation du pilotage des équipements domestiques, seuls quelques éléments de réflexion seront abordés dans cette partie puisque le temps imparti nous a contraint à limiter nos recherches dans ce domaine.

Actuellement dans le parc résidentiel français, un nombre assez limité d’équipements domestiques sont pilotables et parmi eux, tous ne sont pas effectivement pilotés. Traditionnellement, le déclenchement automatique d’un appareil domestique est commandé par un signal horo-journalier émanant du réseau électrique. Dans l’Hexagone, ce système a été massivement mis en place pour l’asservissement des chauffe-eau électriques à accumulation aux heures creuses de la journée (fin de soirée et nuit) dans un but de lissage de la courbe de charge [BRA96].

La modélisation du pilotage nécessite donc de représenter les éléments assurant la commande d’un appareil. Ainsi, nous introduisons le concept de [profil exogène\\*](#). Il s’agit de la connaissance de la fluctuation au cours du temps de la valeur d’une variable d’intérêt. Cette évolution est restituée à une certaine résolution temporelle (par exemple l’heure) et sur une période donnée (un jour typiquement). La modélisation du pilotage d’un équipement fait donc appel à un profil exogène de commande. Celui-ci traduit la possibilité pour un appareil de fonctionner durant certaines périodes de la journée et au contraire il exclut la marche de l’appareil pour d’autres moments du jour considéré.

21. Ceci témoigne de l’interaction presque systématique, aussi courte soit elle, qu’il existe entre l’Homme et un appareil juste avant le fonctionnement de celui-ci

22. cf. description du processus de construction dans la partie 3.2.4

En matière de modélisation, si la commande est réduite à du tout ou rien, une représentation binaire nous semble la plus pertinente. Cette information doit être connue à chaque instant de la simulation (conformément au choix du paramètre `ResolutionCDC`) et, à l'image de la forme développée de l'usage d'un équipement ou d'un intervalle infra-journalier, le profil exogène peut être représenté sous forme d'un tableau.

Ceci conclut la partie relative aux éléments de la modélisation générique des équipements d'électricité spécifique. Dans les paragraphes qui suivent, nous abordons plus en détails les adaptations de cette modélisation conformément aux spécificités liées à chaque type de fonctionnement.

### 3.1.3.2 Les équipements à cycle de fonctionnement

#### Spécificités de ces équipements

Les équipements de cette catégorie sont caractérisés par un **fonctionnement déterministe** suivant la sélection, par l'utilisateur domestique, des paramètres de marche de l'appareil. Plus précisément en ce qui nous concerne, connaissant l'instant de déclenchement et le mode de fonctionnement de l'équipement, ce sont la temporalité<sup>23</sup> et l'amplitude (l'ordre de grandeur) des appels de puissance électrique requis par l'appareil qui sont largement conditionnés, mais non totalement figés, par le choix du type de cycle de fonctionnement. En effet, en plus de ces déterminants principaux sur les besoins énergétiques de ce type d'équipements, il subsiste des influences « secondaires » et exogènes qui impactent la consommation de l'appareil domestique lors de son fonctionnement. En outre le plus souvent, celles-ci ne peuvent pas être contrôlées par l'utilisateur de l'appareil.

De façon générale, les instants de déclenchement de cette catégorie d'équipements font directement suite à l'intervention humaine. Toutefois, pour certains d'entre eux, la mise en route peut être programmée à l'avance par l'utilisateur. Dans ce cas, l'appareil obéit à un signal interne issu de son horloge. En revanche, comme nous l'avons déjà suggéré, l'arrêt de ce type d'équipement se fait de manière automatique. Enfin dans tous les cas, l'utilisateur peut interrompre un cycle de fonctionnement d'un appareil de ce genre et il a de plus la possibilité, pour un certain nombre d'appareils, de le reprendre à l'état d'avancement du cycle atteint avant l'interruption.

Nous illustrons nos précédents propos sur l'exemple de la machine à laver qui d'une part est bien représentative de cette catégorie d'appareils et qui, d'autre part, a été le premier équipement considéré dans notre modélisation et pour lequel nous avons obtenu les résultats les plus fournis (voir chapitre 4). Si l'on considère un cycle de fonctionnement à 30 °C avec une vitesse d'essorage à 1400 tours/min, le cycle unitaire de puissance correspondant du lave-linge aura l'allure de celui présenté par la figure 3.3.

Le plus communément, un cycle unitaire de puissance de machine à laver est composé de trois phases principales – identifiées sur la figure 3.3 – qui se déroulent dans l'ordre chronologique suivant :

1. chauffage : la température de l'eau froide injectée dans le tambour en début de cycle est élevée à la consigne programmée par l'utilisateur. Ceci occasionne un appel de puissance d'amplitude maximale et d'une durée proportionnelle à la différence de température entre l'eau froide soutirée et la température du programme de lavage ;
2. lavage/rinçage : le linge est brassé un certain nombre de fois relativement au mode de fonctionnement retenu (en particulier du type de linge nettoyé) conduisant à des appels de puissance d'amplitude faible ;
3. essorage : les habits lavés sont ensuite essorés à haute vitesse provoquant un appel de puissance d'amplitude faible à moyenne conformément à la vitesse d'essorage sélectionnée par l'utilisateur.

---

23. Celle-ci témoigne tout d'abord des éventuelles phases de natures différentes qui s'enchaînent, dans un ordre déterminé, lors d'un unique cycle de marche d'un appareil. Ensuite, c'est la durée totale de marche de l'équipement qui est connue au travers de cette temporalité

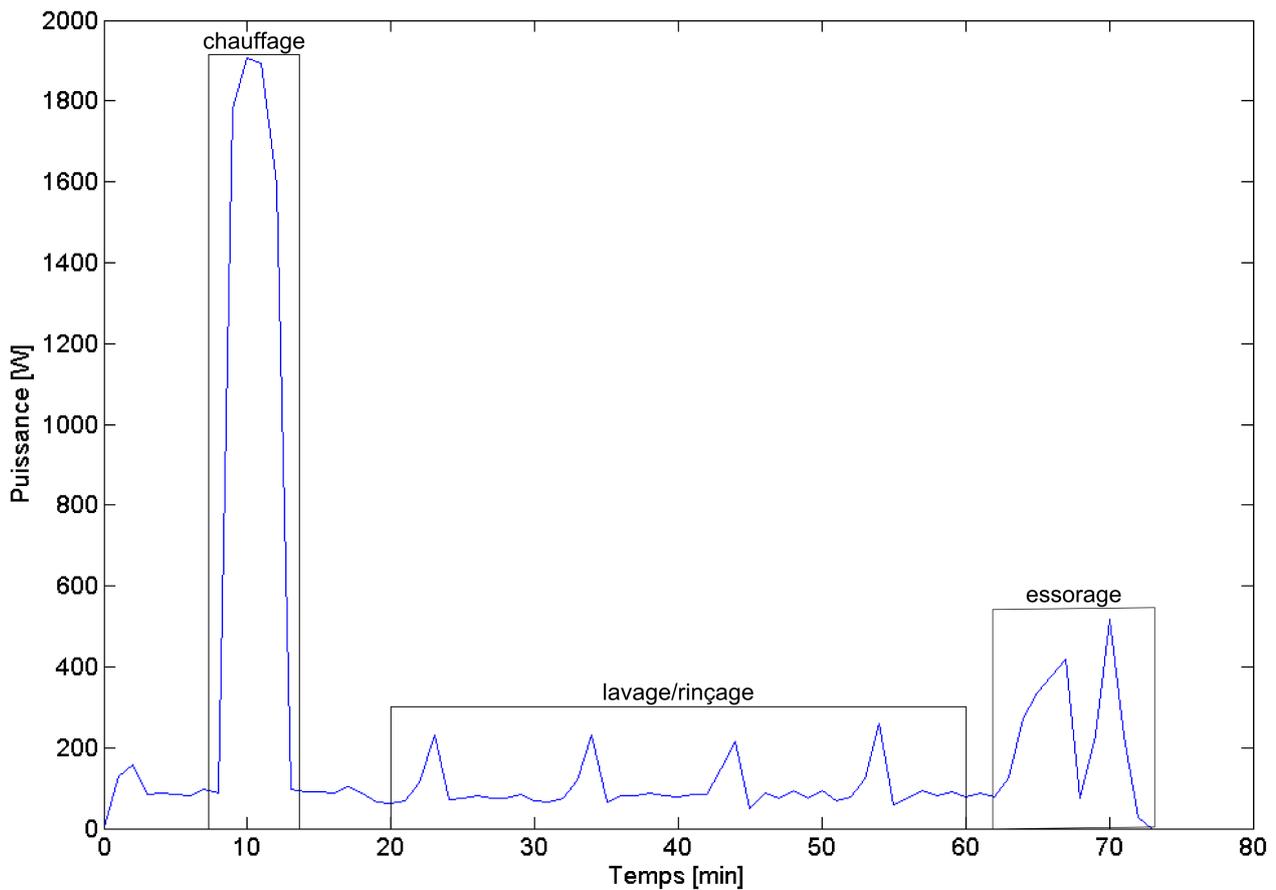


FIGURE 3.3 – Puissance électrique moyenne au pas 1 min d'un cycle à 30 °C et 1400 tours/min d'une machine à laver

De manière plus générique, Stamminger dans [STA09] p. 25 propose une représentation du *CUP* d'une machine à laver (« general pattern »). En plus des phases listées ci-dessus, un cycle de fonctionnement de lave-linge comporte des étapes de trempage et de vidange. En outre et en fonction de la machine utilisée, une étape de pré-lavage et/ou une phase de défroissage peuvent être programmées.

Pour la machine à laver, la température d'eau froide alimentant l'appareil est une influence qui affecte l'appel de puissance de chauffage de l'eau de manière secondaire. En effet, les écarts entre la valeur moyenne annuelle de la température d'eau froide et ses valeurs extrêmes sont moins importants (d'un facteur 10 environ, il ne s'agit que de quelques degrés Celsius) que la différence de température entre celle de l'eau froide puisée à un instant donné et la température d'un programme de lavage conduit au même moment (plusieurs dizaines de degrés Celsius). Habituellement<sup>24</sup>, l'eau froide puisée par le lave-linge est à une température fluctuante en fonction de la saison et cette température est un paramètre exogène au fonctionnement de l'appareil.

### Adaptations apportées à l'approche générique pour la modélisation de ces équipements

Les éléments génériques de modélisation que nous avons présentés dans le paragraphe 3.1.3.1 s'appliquent directement aux équipements à cycle de fonctionnement. En effet, aucune de leurs spécificités n'implique l'introduction de nouveaux concepts et/ou de nouvelles variables par rapport à ceux/celles de la modé-

24. i. e. en dehors des machines à laver avec deux sources d'alimentation en eau dont une chaude produite par exemple par des panneaux solaires thermiques

lisation générique des appareils domestiques. Pour ce type d'appareils, nous avons simplement introduit les modalités adéquates pour quelques paramètres du modèle générique d'équipement domestique, essentiellement pour la nomenclature des appareils et la désignation des types de cycle de fonctionnement.

### 3.1.3.3 Les équipements à états

#### Spécificités de ces équipements

Les équipements dits « à états » désignent les appareils domestiques fonctionnant :

- de manière discontinue dans le temps : bien qu'un niveau de puissance de veille puisse être appelé de façon constante, le fonctionnement « normal » d'un équipement de ce type se fait de manière intermittente et conduit à des appels de puissance d'un niveau supérieur à celui de sa veille éventuelle ;
- pendant des durées non déterministes<sup>25</sup> à partir de leur instant de déclenchement fixé par l'utilisateur (contrairement aux équipements à cycles, l'instant de fin de la période de marche est le plus souvent choisi par l'Homme également) ;
- selon un ou plusieurs mode(s) de fonctionnement caractérisé(s) par un niveau « pseudo-constant » d'appels de puissance. Une fois le mode de fonctionnement choisi, l'ordre de grandeur du niveau de puissance appelée par l'appareil est conditionné. Cette puissance peut cependant varier en fonction de l'environnement d'utilisation de l'appareil et des possibles [perturbations\\*](#) qu'il peut subir, c'est-à-dire des modifications d'origine humaine ou environnementale de ses conditions d'utilisation lors de son fonctionnement, modifications qui ont un impact en matière d'appels de puissance de l'équipement considéré.

Sont inclus dans cette catégorie les équipements audiovisuels, ceux relatifs au poste informatique, certains appareils de cuisson, l'éclairage... Bien souvent pour ce type d'appareils, les trois modes de fonctionnement sont marche, arrêt et veille. Les figures 3.4 et 3.5 représentent respectivement la courbe de charge journalière d'une télévision mesurée au pas 10 min et les niveaux de puissance de plaques à induction en fonction de l'intensité d'utilisation.

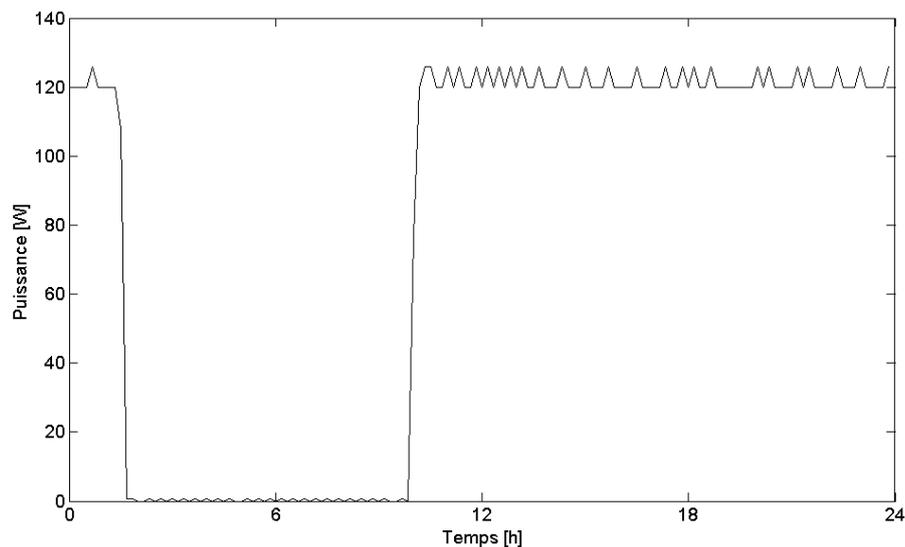


FIGURE 3.4 – Courbe de charge journalière d'une télévision mesurée au pas 10 min issue de REMODECE

Les courbes de charge présentées au travers des figures 3.4 et 3.5 témoignent des fluctuations en matière d'appels de puissance des équipements à états pour un mode de fonctionnement donné de ces appareils. En outre, la figure 3.5 illustre l'impact en puissance de l'intensité d'utilisation des plaques à induction

25. Toutefois, pour certains appareils, des durées par défaut sont programmées par le constructeur

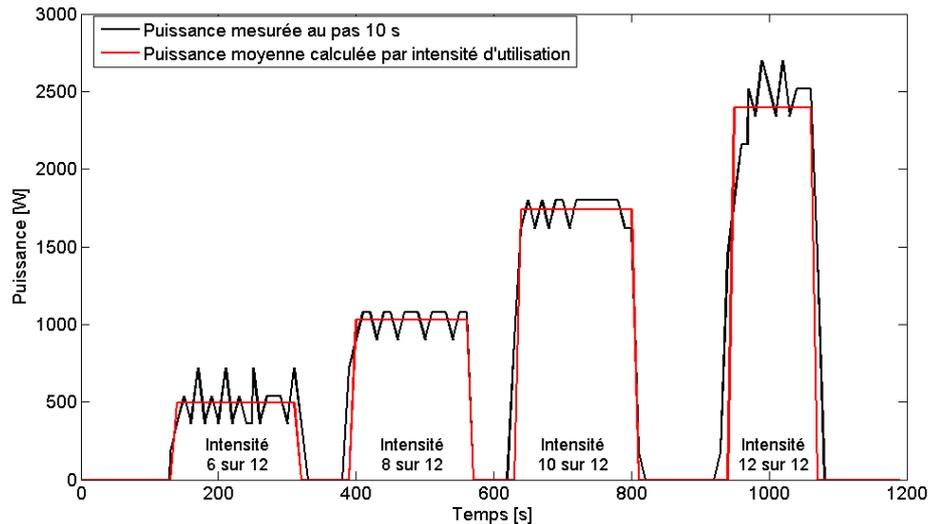


FIGURE 3.5 – Appels de puissance de plaques à induction mesurés au pas 10 s pour différentes intensités d'utilisation (identifiées de 1 à 12)

(différents modes de fonctionnement), phénomène qui peut être généralisé à l'ensemble des appareils de cette catégorie (i. e. plus l'intensité d'utilisation d'un équipement de ce type est élevée, plus les appels de puissance engendrés sont importants).

### Adaptations apportées à l'approche générique pour la modélisation de ces équipements

Les équipements à états se rapprochent largement des équipements à cycles de fonctionnement. Seule la temporalité d'utilisation de ces appareils doit être revue par rapport à la modélisation générique que nous avons développée dans le paragraphe 3.1.3.1. Concrètement, une variable supplémentaire est introduite de manière à renseigner la durée journalière d'utilisation de ce type d'équipement. Celle-ci peut ensuite être répartie en une ou plusieurs périodes, chacune étant d'une durée spécifique. En fonction du nombre de périodes de fonctionnement identifiées ci-avant, autant d'instant de déclenchement de l'appareil doivent être sélectionnés à partir des éléments de la modélisation de l'équipement précédemment décrits. Durant les phases de fonctionnement normal des appareils à états et sans qu'ils soient soumis à une perturbation, nous avons fait le choix de représenter leurs appels de puissance par un niveau constant qui correspond à la puissance nominale de fonctionnement paramétrée par l'utilisateur.

#### 3.1.3.4 Les équipements thermostatés de froid domestique

##### Spécificités de ces équipements

Comme nous l'avons mentionné au sein du paragraphe 2.1.1.3 du chapitre 2 page 68, le froid domestique désigne les équipements réfrigérateur, congélateur et combiné réfrigérateur-congélateur.

Leur rôle est de maintenir à tout instant (dès lors que l'appareil a été branché au secteur et mis en mode marche) une enceinte fermée à une certaine température de consigne programmée par l'utilisateur. Pour la majorité des applications domestiques, l'action de refroidissement est effectuée au moyen d'un système<sup>26</sup> composé d'un évaporateur, d'un compresseur, d'un condenseur et d'une valve de détente. Le compresseur est le seul élément qui nécessite un apport électrique. Selon Stamminger [STA09] p. 100, lors du fonctionnement en régime établi d'un appareil de froid domestique, le compresseur fournit du travail durant 20 à 35 % du temps. Ceci suppose que l'appareil ne subisse pas de perturbations. A contrario,

26. Nous ne considérons pas d'autres systèmes de génération de froid – opérant par absorption ou par effet Peltier – dans notre modélisation car ils nous semblent marginaux pour les applications domestiques

le remplissage de l'appareil avec des éléments à refroidir peut conduire au fonctionnement continu du compresseur (100 % du temps) jusqu'à ce que la consigne de température soit atteinte.

Autrement dit et pour résumer, les appels de puissance de ces équipements permettent d'entretenir un équilibre thermique qui est perturbé par l'intervention humaine ou par le milieu extérieur :

- ouverture et fermeture de la (des) porte(s) principalement au moment des repas pour prélever des aliments ;
- remplissage de l'appareil après que les occupants aient effectué des achats ;
- fluctuation de la température de la pièce dans laquelle est disposé l'appareil.

Ce type de fonctionnement induit un aspect cyclique des courbes de charge de ces équipements comme le montre la figure 3.6.

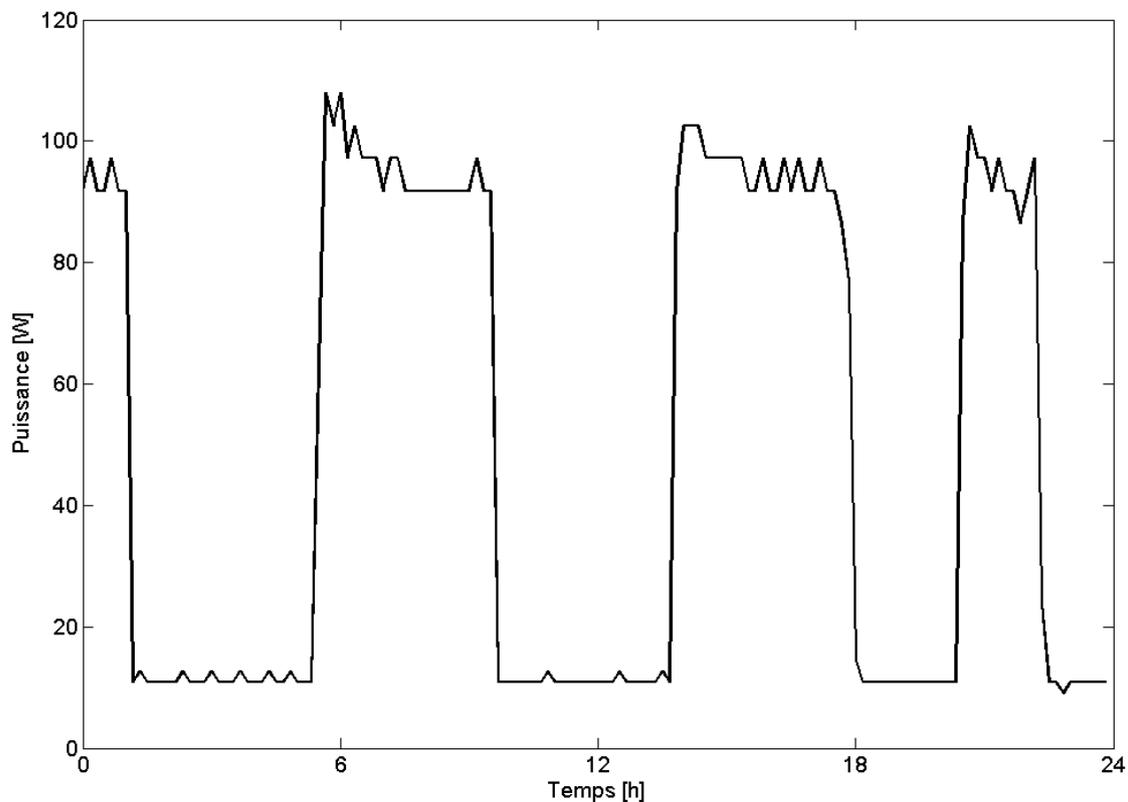


FIGURE 3.6 – Courbe de charge journalière d'un réfrigérateur mesurée au pas 10 min issue de REMODECE

Néanmoins, pour un appareil donné, les courbes de charge journalières correspondantes ne sont pas strictement identiques (présence d'un déphasage entre elles) puisque l'enclenchement d'un cycle de compression peut survenir à tout instant. En effet, le compresseur cherche à contrebalancer les déséquilibres qui affectent l'enceinte réfrigérée de manière stochastique. Bien entendu, la probabilité de ceux-ci est d'autant plus grande autour des heures des repas par exemple.

En plus du mode de marche « classique » de refroidissement, certains appareils sont capables d'assurer leur dégivrage. Celui-ci peut être conduit de manière automatique ou manuelle. Dans le dernier cas, c'est à l'utilisateur de débrancher son appareil et d'attendre la fonte du givre présent dans l'enceinte thermostatée. Les dégivrages dits automatiques (automatique « naturel », automatique « no frost » – ou « free frost » – et automatique par inversion du cycle) sont déclenchés par l'appareil après  $n$  cycles de

compression (après  $n$  heures de fonctionnement pour certains modèles) ou alors ce mode est commandé par l'utilisateur de manière ponctuelle.

### Modélisation spécifique des équipements thermostatés de froid domestique

De par leur type de fonctionnement et leur caractère autonome, la modélisation des équipements de froid domestique est tout autre que celle des équipements à cycle de fonctionnement ou bien de celle des équipements à états. Concrètement, les éléments de la modélisation générique relatifs au déclenchement de l'équipement deviennent caduques. En outre, la notion de liaison entre appareils n'a plus de sens lorsqu'elle s'applique aux équipements de froid domestique. De plus, la prise en compte du besoin de disponibilité pour faire l'usage d'un tel équipement ne constitue pas le cœur de la modélisation de ce type d'appareil. Néanmoins, cela ne signifie pas pour autant que le besoin de disponibilité est inexistant : en effet, il s'exprime différemment dans la mesure où il concerne la modélisation des perturbations qui peuvent affecter le fonctionnement de l'appareil. Pour les appareils de froid domestique, la principale perturbation réside dans l'ouverture de porte et le remplissage/vidage du volume thermostaté.

Tenir compte de ces spécificités de fonctionnement de manière précise requiert, de façon théorique, une modélisation physique assez fine de ces appareils (le chapitre 2 de [ZOU03] présente des éléments d'une telle démarche). Toutefois, celle-ci dépasse largement le cadre des recherches menées dans cette thèse si bien que nous avons opté pour une modélisation simplifiée de ces équipements.

Pratiquement, nous avons cherché à reproduire le fonctionnement cyclique des appareils de cette catégorie en introduisant une variable renseignant la durée caractéristique entre deux cycles de fonctionnement pour chacune<sup>27</sup> des parties réfrigération et/ou congélation. Pour générer une certaine diversité entre les appareils, cette durée caractéristique est sélectionnée aléatoirement au moyen d'une distribution uniforme entre les bornes haute et basse paramétrées par le modélisateur<sup>28</sup>. Dans notre modélisation, les déclenchements des cycles de compression correspondent aux moments de la simulation qui sont multiples de la (des) durée(s) caractéristique(s) de l'appareil. Sans un petit raffinement supplémentaire, ce choix entraîne le synchronisme à l'instant initial des simulations de l'ensemble des appareils de froid domestique : ce phénomène apparaît clairement sur la figure 3.7. Celle-ci représente les courbes de charge moyennes foisonnées de réfrigérateurs obtenues à partir de 100 simulations unitaires, respectivement avec puis sans ajout d'un décalage temporel à l'instant initial de simulation. La courbe de charge calculée sans ajouter de décalage présente un pic de puissance prévisible à  $t = 0$  s.

En effet, 0 est multiple de la durée caractéristique de chacun des appareils quelle que soit la valeur de cette dernière. Pour pallier ce phénomène, un décalage temporel aléatoire est calculé et appliqué à la fraction du nombre des équipements qui sont susceptibles d'être en marche à l'instant initial de la simulation. Ce décalage désynchronise ainsi le déclenchement du premier cycle de compression des appareils au cours d'une simulation. L'ajout de ce décalage temporel a un impact notable visuellement comme nous l'avons constaté au travers de la figure 3.7, mais il modifie également les indicateurs usuels (cf. partie 4.1.1 du chapitre 4) qui caractérisent des courbes de charge comme l'indique le tableau 3.5.

Indicateur [unité]	Symbole	Valeurs des indicateurs pour les courbes de charge calculées	
		avec décalage temporel	sans décalage temporel
Puissance moyenne [W]	$\bar{P}$	20,3	21,1
Puissance maximale [W]	$P_{max}$	32,3	95,9
Facteur de charge [-]	$F_{P_{max}}$	0,63	0,22
Coefficient de foisonnement [-]	$K_{f,n}$	0,304	0,903

TABLEAU 3.5 – Impact du décalage temporel sur la valeur des indicateurs usuels décrivant les courbes de charge moyennes foisonnées des réfrigérateurs issues de simulations

27. L'affectation de deux durées caractéristiques n'intervient que dans le cas du combiné réfrigérateur/congélateur disposant de deux compresseurs

28. Des valeurs par défaut issues de notre connaissance sectorielle sont proposées

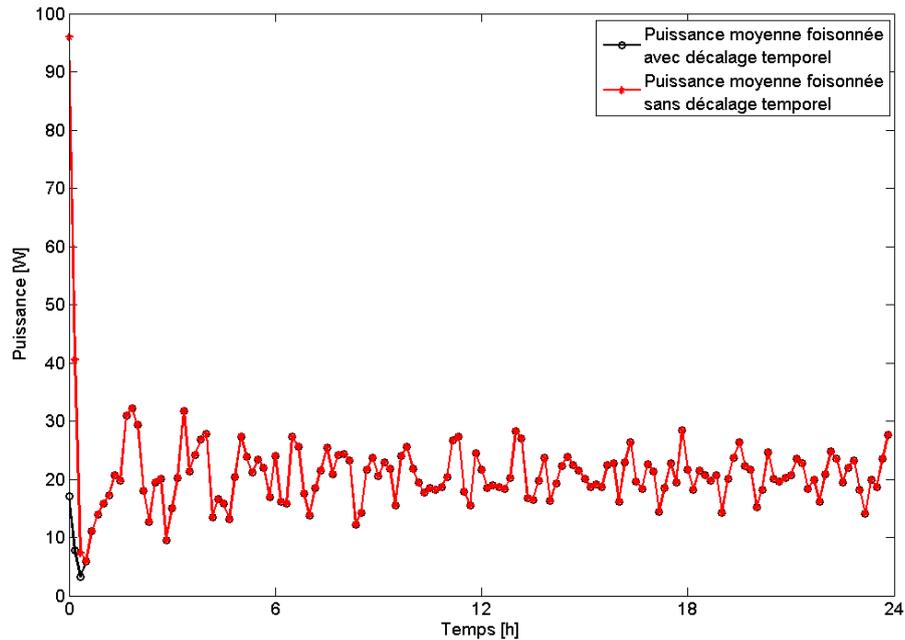


FIGURE 3.7 – Courbes de charge moyennes foisonnées de réfrigérateurs, avec et sans décalage temporel à  $t = 0$  s, obtenues à partir de 100 simulations

Les valeurs très distinctes de la puissance maximale de chacune des courbes de charge moyennes foisonnées de réfrigérateurs impactent directement les autres indicateurs que sont la puissance moyenne, le facteur de charge et le coefficient de foisonnement.

À partir de la (des) durée(s) caractéristique(s) mentionnée(s) ci-dessus, nous calculons la durée du cycle de fonctionnement (i. e. durée de compression) pour la zone de réfrigération et/ou de congélation en la multipliant par un coefficient spécifique à chaque partie de l'appareil (le cas échéant <sup>29</sup>) qui est paramétré par le modélisateur (les fractions de temps avancées par Stamminger [STA09] p. 100 peuvent faire office de référence). Durant les phases de compression, la puissance appelée par chaque partie de l'appareil (le cas échéant) correspond à la puissance nominale paramétrée par le modélisateur. Malgré l'utilisation sur le marché des appareils de froid domestique de compresseurs à vitesse variable [ECE08], nous n'avons pas intégré cette caractéristique dans notre modélisation si bien que les appels de puissance en fonctionnement forment des créneaux. En dehors des phases de marche du (des) compresseur(s), l'appareil soutire une puissance de veille dont la valeur est choisie par le modélisateur.

De plus, celui-ci doit renseigner les limites haute et basse du nombre d'ouvertures journalières de porte de la partie réfrigération et/ou congélation. Un entier choisi aléatoirement entre ces bornes définit ce nombre d'ouvertures. Celles-ci sont également renseignées en durée : l'auteur du scénario doit en effet fixer des limites inférieure et supérieure à la durée des ouvertures de porte pour chacune des parties (réfrigération et/ou congélation) de l'appareil (le cas échéant). La durée de chaque ouverture de porte survenant lors de la simulation du scénario est sélectionnée au moyen d'une distribution uniforme entre les bornes haute et basse précédemment évoquées.

À l'heure de la rédaction de la thèse, la modélisation des équipements de froid domestique se limite aux éléments présentés ci-dessus. Le positionnement au cours d'une journée simulée des ouvertures de porte n'est pas pris en compte. Néanmoins, nous avons défini les prémices du processus permettant de fixer les moments où surviennent ces perturbations. Celles-ci doivent en effet être en rapport direct avec les

29. cf. note 27

heures de prise des repas. D'autre part, il manque également dans notre modélisation, l'implémentation de l'impact en puissance de l'ensemble des perturbations. L'étude de données expérimentales serait un moyen d'estimer, dans une première approche, l' (les) effet(s) en puissance des perturbations sur un régime de marche établi de ce type d'appareils.

Dans cette partie 3.1, nous avons exposé et expliqué nos choix en matière de modélisation des concepts de la méthode de reconstitution de courbe de charge. Les représentations évoquées précédemment constituent le socle minimal permettant par la suite de mener des simulations. En effet, les modèles présentés ci-avant sont capables d'interagir entre eux de façon dynamique dans le but de générer de la diversité d'origine comportementale en matière de consommation d'électricité dans le secteur domestique. Cette dernière thématique est abordée plus en détails dans la section 3.2.

## 3.2 Algorithme pour la génération de diversité d'origine comportementale

Comme nous l'avons souligné dans l'introduction de ce chapitre, des fluctuations de la consommation d'électricité du secteur résidentiel peuvent être repérées à différentes échelles temporelles. La reproduction de ces particularités essentielles de la courbe de charge ne peut être envisagée que si elles sont considérées individuellement. Ainsi, nous nous sommes concentrés avant tout sur la modélisation de la diversité à l'échelle journalière incluant par là même les variations horaires et infra-horaires. Pour cela, nous avons défini un processus en trois étapes :

1. distribution des tâches journalières pour un jour-type et un ménage-type ;
2. création d'un profil de disponibilité par ménage-type en fonction du jour-type ;
3. placement des équipements domestiques conformément au profil de disponibilité du ménage-type lors du jour-type simulé.

Nous détaillons chacune de ces étapes dans les prochains paragraphes.

### 3.2.1 Distribution des tâches journalières pour un jour-type et un ménage-type

La première partie du processus de génération de diversité d'origine comportementale consiste à allouer un certain nombre de tâches domestiques, sous forme d'équipements dont il faut faire l'usage, conformément aux caractéristiques du jour-type et du ménage-type simulés. L'étude du comportement des ménages, qui est notamment réalisée au moyen de l'Enquête Emploi du Temps de l'INSEE [INS10a], dévoile également des périodicités : le jour, la semaine, le mois et l'année constituent les principales échelles de temps autour desquelles s'organisent les activités au sein des foyers. Distribuer les tâches domestiques au jour le jour impliquent donc de prendre en considération ces dynamiques de comportement.

Cette étape nécessite des informations macroscopiques (ou agrégées) concernant les habitudes des ménages. Par exemple, la fréquence hebdomadaire de cycles de machine à laver ou la durée quotidienne de visionnement de la télévision sont des éléments à connaître lorsque l'on s'intéresse à ces deux appareils. Les équipements domestiques permettent d'assurer des usages qui peuvent être regroupés en un nombre restreint de classes<sup>30</sup> que nous appelons **macro-tâches\*** domestiques. Ce terme s'applique à l'échelle macroscopique c'est-à-dire qu'à ce niveau, nous considérons l'ensemble des ménages du territoire observé. La segmentation suivante a été établie :

- Préparation Culinaire et Repas (PCR)\*<sup>31</sup> ;
- Hygiène Corporelle (HC)\* ;
- Sommeil (SO)\* ;
- L'Avage (LA)\* ;

30. À l'image de ce que nous avons déjà présenté au sein du paragraphe 2.1.1.3 du chapitre 2

31. Comme pour les intervalles infra-journaliers, nous avons décidé d'utiliser des majuscules pour différencier les éléments de la modélisation des notions éponymes qui font partie de notre quotidien

- Entretien Domestique (ED)\* ;
- Travail À Domicile (TAD)\* ;
- Loisirs, Repos, Communications et Mobilité (LRCM)\*.

La définition des tâches journalières pour un ménage-type lors d'un jour-type donné revient à distribuer et à lister l'ensemble des équipements qui seront utilisés au sein du logement le jour considéré. Soulignons à ce stade que dans une logique d'uniformisation de la modélisation, nous assimilons les activités domestiques ne faisant pas nécessairement appel à l'électricité, le sommeil par exemple, à des équipements aux appels de puissance nuls. Malgré le fait qu'elles ne soient pas systématiquement sources d'appels de puissance, ces activités doivent être prises en compte du fait de leur impact sur le comportement des ménages. Celles-ci participent notamment à la constitution des rythmes chronobiologiques et/ou rendent improbable l'utilisation de certains appareils tant qu'elles sont en cours (c'est en particulier le cas de l'activité SOMmeil).

Plus précisément, la définition des tâches journalières consiste en la spécification des valeurs des paramètres fonctionnels (mode de fonctionnement ou puissance nominale et durée de marche) pour chacun des équipements utilisés au cours d'un jour-type donné. De plus, pour faciliter le placement des équipements au cours d'un jour simulé, les tâches du jour doivent être ordonnées de telle sorte que les équipements de tête arrivent en premier sur la liste par rapport à leur(s) équipement(s) lié(s) (avant eux donc). Le tableau 3.6 est un exemple d'une telle liste de tâches domestiques au sein d'un jour-type donné.

Macro-tâche	Équipements	Paramètres fonctionnels		
		Mode de fonctionnement	Puissance nominale [W]	Durée de marche [min]
PCR	réfrigérateur	2 °C	–	1440
	four à micro-ondes	–	360	15
	bouilloire	–	2000	3
HC	sèche-cheveux	–	2000	10
	rasoir	–	10	15
SO	sommeil	–	0	420
LA	lave-linge	30 °C	–	70
	lave-vaisselle	50 °C	–	80
ED	aspirateur	–	1400	30
	fer à repasser	–	2400	35
TAD	ordinateur fixe	–	300	120
	imprimante	–	100	20
LRCM	télévision	–	150	180
	lecteur DVD	–	200	120
	chaîne Hi-Fi	–	300	45
	halogène	–	200	30
	véhicule électrique	–	500	360

TABLEAU 3.6 – Liste des équipements à placer avec leurs paramètres fonctionnels pour un jour-type donné

Comme nous l'avons souligné ci-avant, l'obtention d'une telle distribution des tâches journalières nécessite une fine connaissance des habitudes des ménages. En plus de ces éléments sectoriels, des règles doivent être définies pour recréer une cohérence en matière d'utilisation journalière des appareils. D'autre part, il faut également introduire dans cette phase du processus (de génération de foisonnement d'origine comportementale) des éléments conduisant à reproduire de la diversité affectant la fréquence d'utilisation des appareils, le choix des modes de fonctionnement et des modes de gestion des équipements. En outre, ce processus doit aussi faire en sorte que la distribution des équipements domestiques tienne compte des singularités du jour-type, de celles du ménage-type (y compris de son comportement général) et enfin, des tâches réalisées les jours antérieurs notamment. L'influence météorologique sur nos activités journalières

n'est pas nulle et doit ainsi être prise en considération dans ce processus.

Les quelques éléments précédents suggèrent la complexité de cette phase de listage. Pour être obtenues de manière automatique en fonction du contexte modélisé, ces distributions de tâches journalières doivent être le fruit d'un algorithme tenant compte de toutes les influences mentionnées au préalable. Étant donné que de telles distributions de tâches journalières peuvent être facilement établies par l'auteur d'un scénario de modélisation, nous avons décidé de ne pas nous concentrer sur la construction d'un tel algorithme.

Pour notre part, les scénarios modélisés que nous détaillerons dans le chapitre 4, s'appuient sur des distributions **déterministes** de tâches journalières puisqu'elles sont élaborées par le modélisateur et elles ne mettent en jeu que quelques équipements. Concrètement, nous verrons par exemple que le scénario sur lequel nous nous sommes majoritairement concentrés et que nous avons simulé, implique la machine à laver et le sèche-linge. En outre, pour chaque itération de ce scénario journalier, ces équipements doivent fonctionner au cours de la simulation. Ainsi, pour reproduire un certain aléa en matière d'utilisation ou non de tel ou tel appareil lors d'un jour donné<sup>32</sup>, nous introduisons une variable supplémentaire – `ProbabiliteDeclenchement` – dans la modélisation des équipements à cycles de fonctionnement et à états. Ce paramètre fixé par le modélisateur attribue la probabilité d'utilisation d'un appareil lors du jour considéré.

La première phase du processus de génération de foisonnement d'origine comportementale consiste donc à renseigner dans le scénario modélisé quels sont les équipements utilisés pour un ménage donné. Lister ces appareils, avec leur mode de fonctionnement respectif pour le jour simulé, peut être opéré de manière déterministe par le modélisateur, ou bien, grâce à une méthode statistico-probabiliste plus complexe qui valorise entre autres des données socio-culturelles.

### 3.2.2 Création d'un profil basique de disponibilité par ménage-type en fonction du jour-type

La seconde partie du processus de génération de foisonnement renseigne les habitudes des foyers modélisés relatives à la présence au domicile de leurs occupants.

Partant du constat que les **profils d'occupation**\* normatifs sont trop standardisés et trop peu variés pour garantir une diversité suffisante en matière de scénarios de présence au domicile, nous avons créé une routine de simulation qui produit une variété de profils de disponibilité en fonction des caractéristiques du ménage-type et du jour-type simulé. Une partie de cette routine est prise en charge par la sélection aléatoire des instants caractéristiques entre des bornes paramétrées.

L'idée est de définir et d'allouer aux ménages des périodes permises d'utilisation des appareils domestiques sachant que le niveau de disponibilité d'un ménage varie d'un intervalle à un autre conformément aux activités de ses membres. Ici interviennent donc les influences issues de ménages-types aux caractéristiques diversifiées ainsi que des jours-types traduisant une variété de rythmes journaliers d'activités domestiques. Dans la première approche que nous avons adoptée et que nous détaillons dans ce paragraphe, le profil de disponibilité, construit au niveau du ménage agrégé, s'obtient directement de la forme développée des intervalles infra-journaliers, c'est-à-dire dès lors que l'on a la connaissance du niveau de disponibilité du ménage dans chaque intervalle, et ce, avant le placement des équipements.

Le niveau de disponibilité du ménage est défini en fonction des caractéristiques de ce dernier et de celles du jour-type. Il est en outre attribué intervalle par intervalle. Pour modéliser le niveau de disponibilité, nous avons choisi une représentation à trois états que nous détaillons ci-dessous :

- -1  $\equiv$  présence au domicile des occupants d'un logement à l'état de sommeil : tous les membres du foyer dorment ;

<sup>32</sup>. À défaut d'une méthode raffinée de distribution des tâches journalières lors d'un jour donné prenant en considération toutes les influences que nous avons détaillées dans la partie 3.2.1, nous élaborons ici une première source de variabilité au sein d'un scénario modélisé

- 0  $\equiv$  vacance du logement : tous les occupants du foyer sont absents ;
- 1  $\equiv$  présence des occupants au domicile en mode actif : au moins une personne apte à utiliser des appareils domestiques, c'est-à-dire disposant du niveau adéquat de responsabilité, est éveillée.

Au commencement d'une simulation, la forme développée d'un intervalle  $\iota$  n'est autre chose qu'un tableau de  $l_{Intervalle}(\iota)$  éléments tous égaux au niveau de disponibilité dans cet intervalle, du ménage considéré de manière agrégée. Le profil de disponibilité du ménage correspond alors dans cette première version au résultat de la simple mise bout à bout dans l'ordre chronologique de la forme développée de chacun des intervalles infra-journaliers. Nous proposons au travers de la figure 3.8 un exemple de profil de disponibilité et du **résumé du jour-type\*** pour un ménage actif d'une personne lors d'un jour-type **Semaine**.

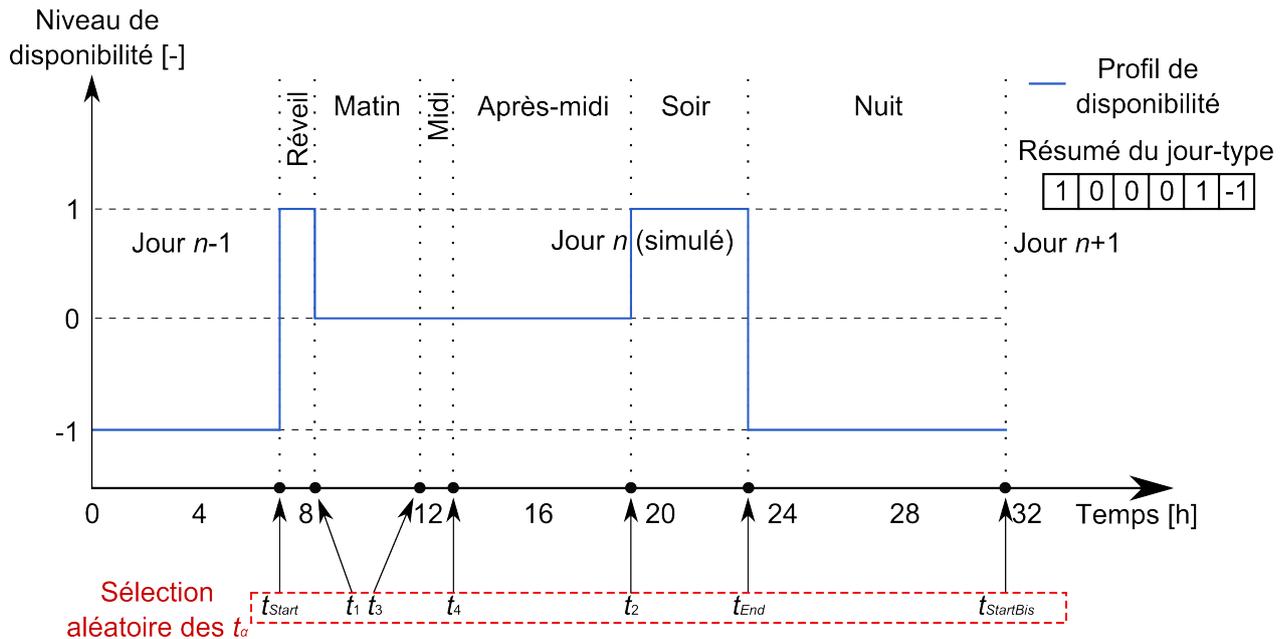


FIGURE 3.8 – Profil de disponibilité et résumé d'un jour-type **Semaine** pour un ménage-type **Actif**

Comme nous l'avons évoqué au paragraphe 3.1.1 page 114, la modalité **Absence** du jour-type n'a pas d'influence sur la sélection des instants caractéristiques puisque, quelles que soient leur valeur, le profil de disponibilité correspondant à cette modalité du jour-type sera identiquement « plat » et d'ordonnée (constante) nulle.

Le croisement des influences du jour-type et du ménage-type sur le profil de disponibilité est un premier résultat intermédiaire du processus de génération de diversité comportementale. Le résumé du jour-type est un tableau de six éléments qui contient le niveau de disponibilité du ménage, au début de la simulation, dans chaque intervalle infra-journalier. Cet objet condensé est utilisé dans la suite du processus pour synthétiser des permissions et des obligations pour chaque macro-tâche au sein de chaque intervalle infra-journalier.

La figure 3.9 propose une variété de profils de disponibilité pour un scénario identique à celui que nous avons précédemment évoqué (ménage-type **Actif** composé d'une personne pour un jour-type **Semaine**).

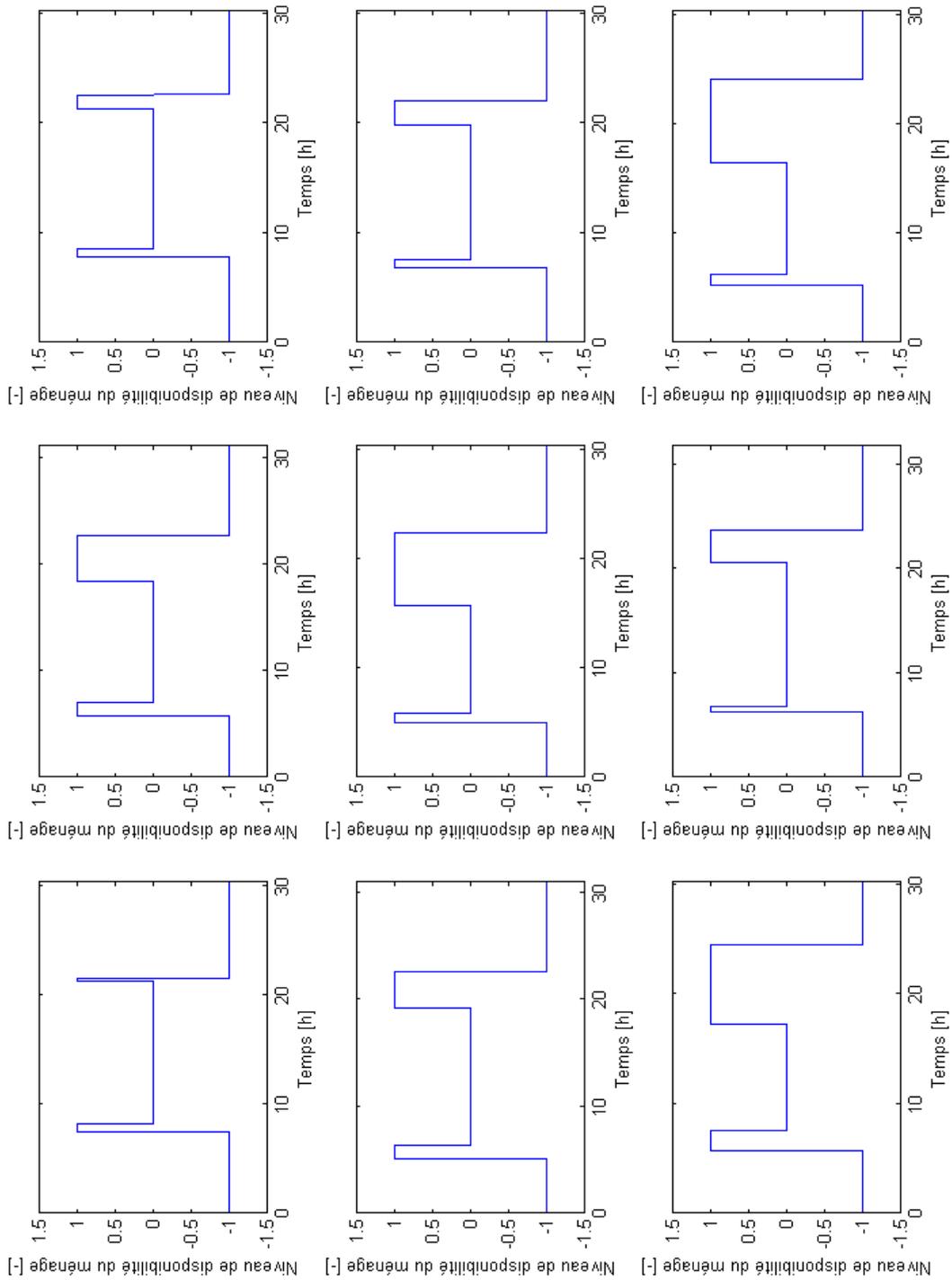


FIGURE 3.9 – Exemples de profils de disponibilité pour un ménage-type **Actif** composé d'une seule personne lors d'un jour-type **Semaine**

### 3.2.3 Méthode évoluée d'obtention des profils de disponibilité

Les profils de disponibilité abordés précédemment ont été construits en assimilant le ménage à une entité formant un tout indivisible. Pour aboutir à ces profils, et plus précisément, pour affecter un niveau de disponibilité au ménage dans chacun des intervalles infra-journaliers, nous avons suivi les trois règles évoquées page 131.

Dans cette seconde approche<sup>33</sup> au contraire, la prise en compte du niveau de disponibilité à chaque instant, qui in fine aboutit au profil de disponibilité, se fait individu (à responsabilité) par individu. Toutefois, l'information concernant la disponibilité du ménage, en tenant compte de l'ensemble de ses membres, conserve son importance. Nous appelons **profil agrégé de disponibilité\*** d'un ménage, le profil construit en effectuant la somme des profils de disponibilité individuels des personnes qui le constitue en appliquant les règles présentées ci-après.

Concrètement dans ce mode de construction des profils de disponibilité, l'affectation à un individu d'un tel profil, qui n'a de sens, rappelons-le, que pour les personnes du ménage ayant un niveau de responsabilité compatible avec l'utilisation des appareils ménagers, obéit aux **règles de cumul (ou d'additivité)\*** ci-dessous :

- le niveau de disponibilité « 1 » témoigne de l'état d'éveil d'une personne et de sa possibilité à utiliser un appareil domestique à un instant donné. **Ce niveau est cumulatif** c'est-à-dire que quel que soit l'instant considéré, le niveau total de disponibilité du ménage tient compte numériquement de la (des) personne(s) éveillée(s) et susceptible(s) de faire l'usage d'un équipement ;
- le niveau de disponibilité « 0 » indique les périodes de la journée durant lesquelles un individu n'est pas présent au sein de son logement. À l'image du précédent, **ce niveau est cumulatif**. Ainsi, à chaque instant, le niveau total de disponibilité du ménage reflète de la (des) personne(s) absente(s) par l'ajout du niveau « 0 » au profil agrégé de disponibilité du ménage pour chacune des personnes absentes ;
- enfin le niveau de disponibilité « -1 » correspond à l'état de sommeil d'une personne. **Ce niveau n'est pas cumulatif** c'est-à-dire que si l'ensemble des individus formant le ménage dorment simultanément à un instant donné, le niveau de disponibilité résultant sera systématiquement minoré par -1. De même, si au sein d'un ménage de trois personnes à responsabilité, deux individus dorment à un instant donné alors que le troisième demeure simultanément éveillé, le niveau de disponibilité résultant à l'échelle du ménage sera égal à 1<sup>34</sup>.

En plus de ces règles, cette méthode d'obtention de profils élaborés de disponibilité conduit à revoir la mécanique de segmentation des jours-types en intervalles infra-journaliers. Plus précisément, il est nécessaire de définir des instants caractéristiques pour chacune des personnes du ménage d'un niveau de responsabilité autorisant le maniement d'équipements domestiques, sans pour autant que le renseignement de paramètres supplémentaires dans le modèle du jour-type ne soit requis. Pour ce faire, il faut tout d'abord comme ce fut le cas dans la première approche, sélectionner un premier jeu<sup>35</sup> d'instants caractéristiques  $t_\alpha$ , appelé **jeu initial d'instants caractéristiques\***, entre les bornes inférieures et supérieures qui ont été paramétrées au niveau du jour-type (cf. paragraphe 3.1.1). Une fois choisis, ceux-ci deviennent alors des bornes inférieures ou supérieures (en fonction du  $t_\alpha$  considéré) des domaines dans lesquels restent à définir les instants caractéristiques des ( $\text{NbPersonnesAResponsabilite} - 1$ )<sup>36</sup> autres jeux. Ces derniers sont appelés **jeux supplémentaires d'instants caractéristiques\***.

La sélection de plusieurs jeux d'instants caractéristiques est schématiquement décrite par la figure 3.10. Sur cette représentation, nous avons indiqué les instants caractéristiques appartenant au jeu initial sélectionné ainsi que les bornes<sup>37</sup> inférieures et supérieures de ceux-ci. De plus, nous y avons fait figurer au moyen de rectangles rouges, les zones dans lesquelles doivent être choisis les instants caractéristiques des

33. Soulignons tout de suite que celle-ci n'a d'intérêt que si le ménage comporte plus d'une personne à responsabilité

34. La personne éveillée prime sur les deux autres à l'échelle du ménage

35. C'est-à-dire les 7 instants caractéristiques suivants :  $t_{Start}$ ,  $t_1$ ,  $t_3$ ,  $t_4$ ,  $t_2$ ,  $t_{End}$  et  $t_{StartBis}$

36. Ce qui correspond aux autres personnes à responsabilité du ménage

37. Ces bornes sont issues du paramétrage au niveau du modèle du jour-type

jeux supplémentaires. Les instants caractéristiques qui composent ces derniers sont notés  $t_{\alpha+}$ . En outre, le cas du choix de  $t_{Start}$  étant analogue à celui de  $t_{StartBis}$  nous proposons une représentation commune pour ces deux processus de sélection (schéma n° 1 de la figure 3.10).

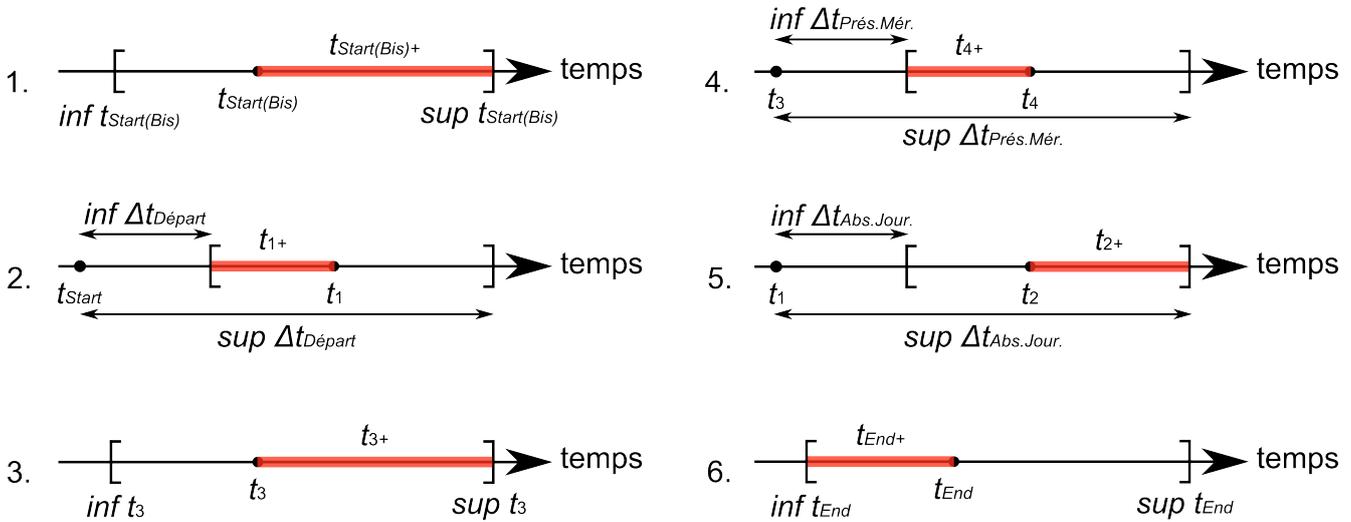


FIGURE 3.10 – Processus de sélection de différents jeux d’instant caractéristiques pour un ménage constitué de plus d’une personne à responsabilité

Une fois sélectionné l’ensemble des jeux d’instant caractéristiques, le niveau de disponibilité de chaque personne à responsabilité, conformément à son statut d’activité, est renseigné au sein des intervalles infra-journaliers. De plus, pour les individus qui sont de type *Inactif*, le processus prend en compte la possibilité d’absences journalières occasionnelles (cf. note 3 de ce chapitre et paragraphe 3.1.2 p. 116). Le profil agrégé de disponibilité du ménage s’obtient enfin en regroupant les profils individuels de chaque membre du foyer tout en appliquant les règles de cumul décrites au préalable. La figure 3.11 propose quelques exemples de profils agrégés de disponibilité pour un ménage composé de quatre personnes à responsabilité dont deux de type *Actif*, une de type *ActifDomicile* et une de type *Inactif* lors d’un jour-type *Semaine*.

La deuxième étape du processus de génération de foisonnement d’origine comportementale consiste à prendre en compte le potentiel temporel journalier d’un ménage pour l’utilisation d’appareils domestiques. Pour cela, nous avons choisi de construire des profils de disponibilité qui prennent en considération les caractéristiques du ménage et du jour simulés.

Une fois cette construction établie, la problématique du déclenchement de l’usage des équipements domestiques peut être abordée. C’est le sujet du paragraphe qui suit.

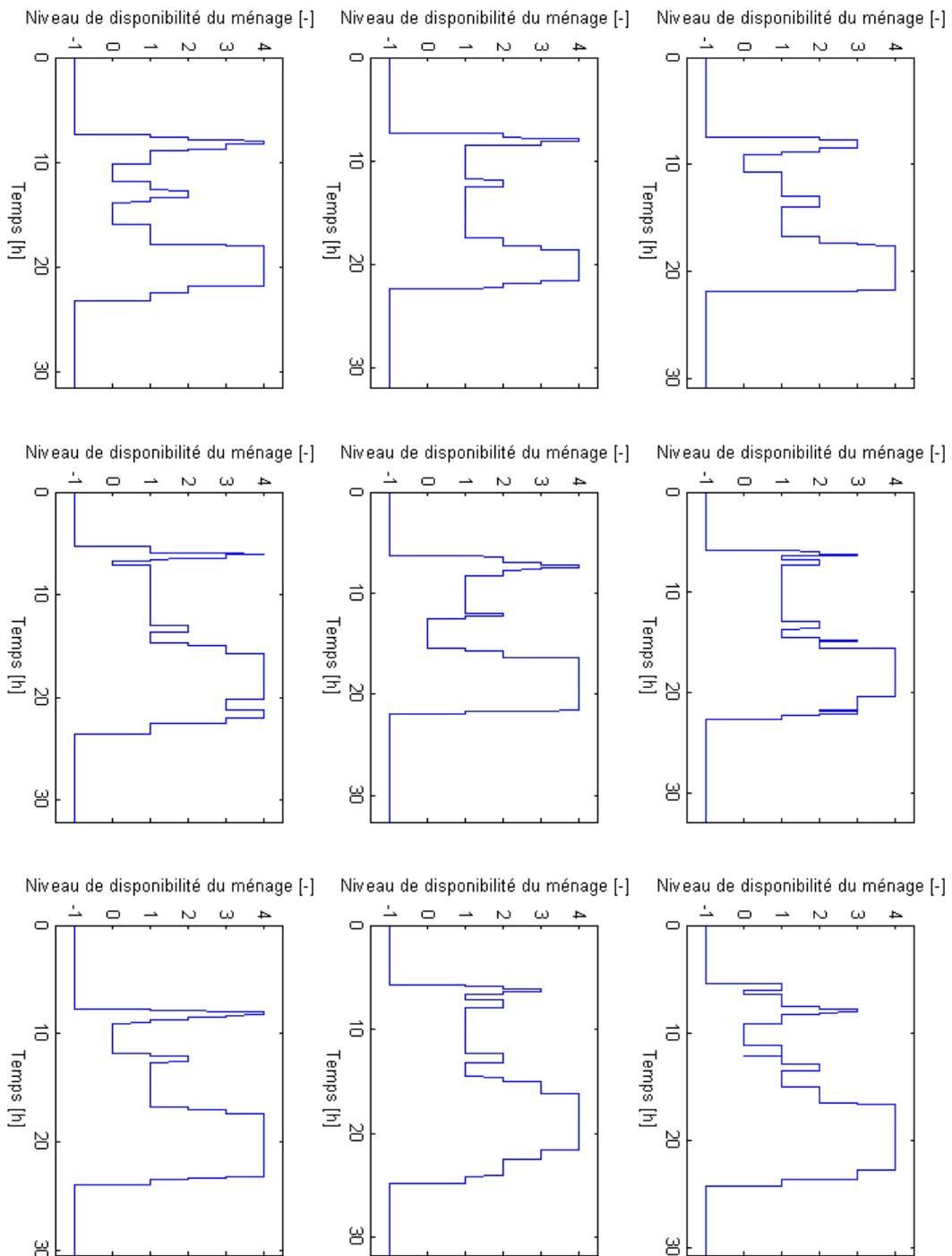


FIGURE 3.11 – Exemples de profils agrégés de disponibilité d'un ménage composé de quatre personnes à responsabilité (deux personnes de type Actif, un individu de type ActifDomicile et un membre de type Inactif) au cours d'un jour-type Semaine

### 3.2.4 Placement des équipements domestiques conformément au profil de disponibilité du ménage-type lors du jour-type simulé

Comme nous l'avons déjà souligné, les macro-tâches domestiques sont réalisées par les occupants des foyers à des moments préférentiels alors que d'autres périodes infra-journalières sont théoriquement plus incompatibles<sup>38</sup> avec l'utilisation d'appareils ménagers. Plutôt que de définir des probabilités de conduite d'une macro-tâche pour chaque intervalle infra-journalier, nous constituons une *Matrice de Macro-Compatibilité* notée *MMAC\** qui renseigne a priori des autorisations et des interdictions au moyen d'un codage binaire (0 représente l'interdiction et 1 l'autorisation). Pour la construire, nous croisons les 6 intervalles infra-journaliers notés  $\iota$  avec les 7 macro-tâches notées  $\underline{a}$  et nous renseignons l'autorisation/l'interdiction dans chacun des cas. Le statut de permission/d'interdiction d'une macro-tâche face à un intervalle sera appelé naturellement *macro-compatibilité de la tâche domestique\**.

Établir la matrice de macro-compatibilité est du ressort du modélisateur. Les choix de celui-ci peuvent reposer sur les pratiques générales de la population (que l'on peut extrapoler à partir de résultats d'enquêtes) ou sur son expérience personnelle par exemple. Dans un premier temps, seule l'influence du jour-type sera prise en compte (nous construisons alors une matrice de macro-compatibilité par jour-type). Nous différencierons ensuite ces autorisations/interdictions en fonction du ménage-type. Un exemple de matrice de macro-compatibilité est proposé au travers du tableau 3.7. Ce dernier indique par exemple que la tâche « Préparation Culinaire et Repas » (PCR) n'est autorisée que durant les intervalles Réveil, Midi et Soir.

	PCR	HC	SO	LA	ED	TAD	LRCM
Réveil	1	1	0	1	1	1	1
Matin	0	0	0	1	1	1	1
Midi	1	0	0	0	0	0	1
Après-midi	0	0	0	1	1	1	1
Soir	1	1	0	1	1	1	1
Nuit	0	0	1	0	0	0	0

TABLEAU 3.7 – Exemple de matrice de macro-compatibilité (jour-type *Semaine*)

Tous les éléments modélisés et décrits depuis le début de ce chapitre ont été identifiés indispensables au placement des équipements au sein des intervalles infra-journaliers ce qui constitue la troisième et dernière phase du processus de génération de foisonnement d'origine comportementale. Pour gérer le déclenchement des équipements au cours d'une journée, nous avons élaboré un algorithme stochastique que nous détaillons dans les paragraphes ci-dessous.

Étant donné un scénario de modélisation (tâches du jour, ménage-type, jour-type), un mécanisme non déterministe de placement des équipements est requis pour générer une diversité temporelle des instants de déclenchement des appareils domestiques. Cet algorithme n'emploie que des éléments mathématiquement manipulables, c'est-à-dire ceux dont nous avons abordé la description dans les pages précédentes. À partir du scénario choisi, l'algorithme cherche à placer dans la journée chaque équipement de manière individuelle, en tenant compte à la fois des liaisons possibles entre appareils domestiques et de la hiérarchie des équipements de tête. Il est à noter qu'au fur et à fur du placement des équipements (i. e. tout au long de la durée de simulation), les objets modélisés relatifs à la temporalité du scénario (profil de disponibilité des ménages, forme résumé des intervalles infra-journaliers) sont constamment mis à jour. Dans les prochains paragraphes, nous indiquerons autant que faire se peut, la valeur initiale prise par les objets modélisés ainsi que les principaux changements qu'ils connaissent, le cas échéant, au cours de la simulation.

38. Néanmoins, tout est possible dans le secteur résidentiel...

L'algorithme que nous avons construit, qui doit être opéré appareil par appareil, est un processus en cinq étapes listées ci-dessous :

1. calcul de la **Matrice de Micro-Compatibilité\*** ;
2. sélection d'un **intervalle infra-journalier éligible\*** pour l'équipement considéré ;
3. recensement des **instants éligibles de déclenchement\*** dans l'intervalle choisi ;
4. sélection d'un instant de déclenchement pour l'équipement considéré parmi l'ensemble des possibles ;
5. mise à jour de l'intervalle.

Nous explicitons chacune de ces étapes dans les sections suivantes.

### 3.2.4.1 Calcul de la matrice de micro-compatibilité

Au départ de la simulation, la matrice de micro-compatibilité notée **MMIC\*** doit être calculée à l'aide de la matrice de macro-compatibilité et du résumé du jour-type **RJT\***. Il s'agit d'un simple calcul membre à membre entre la matrice  $6 \times 7$  et la transposée du vecteur  $1 \times 6$  comme l'indiquent les équations 3.2<sup>39</sup> et 3.3.

$$MMIC = MMAC.*RJT^T \quad (3.2)$$

$$MMIC(\iota, a) = MMAC(\iota, a) \times RJT(\iota), \quad \iota \in [1; 6] \text{ et } a \in [1; 7] \quad (3.3)$$

La matrice **MMIC** s'applique au niveau microscopique, c'est-à-dire à l'échelle du ménage individuel simulé. À l'image de la macro-compatibilité, la **micro-compatibilité d'une tâche domestique\*** désigne le statut d'autorisation ou d'interdiction de cette tâche face à un ou plusieurs intervalle(s) infra-journalier(s).

### 3.2.4.2 Sélection d'un intervalle infra-journalier éligible pour l'équipement considéré

La seconde étape de l'algorithme consiste à choisir un intervalle infra-journalier éligible pour l'équipement simulé, c'est-à-dire un intervalle pouvant accueillir l'instant de déclenchement de l'usage de cet équipement. L'**éligibilité d'un intervalle infra-journalier\***  $\iota$  par rapport à un équipement  $i$  donné, c'est-à-dire le fait que l'intervalle soit éligible ou non pour le déclenchement de l'usage de l'appareil considéré, est représentée par une variable binaire notée **IE**( $\iota, i$ )\* avec la convention numérique suivante : 0  $\equiv$  non éligible, 1  $\equiv$  éligible. L'évaluation de cette variable se déroule suivant un processus en cinq phases :

1. transmission de la compatibilité des **micro-tâches\*** aux équipements sous-jacents ;
2. test de compatibilité en équipements ;
3. test de compatibilité en temps ;
4. codage de la compatibilité de l'équipement pour les intervalles qui lui sont éligibles ;
5. sélection de l'intervalle parmi l'ensemble des possibles.

Nous développons chacune des étapes de ce processus dans les encadrés ci-dessous.

#### Transmission de la compatibilité des micro-tâches aux équipements sous-jacents

Après le calcul de la matrice de micro-compatibilité, nous avons fait le choix de transférer par défaut, en début de simulation, la compatibilité des micro-tâches (tâches domestiques effectuées par un ménage spécifié) à chacun des équipements assurant les usages sous-jacents. Par exemple, la micro-compatibilité du L'Avage s'applique au lave-linge et au sèche-linge.

Comme indiqué précédemment, la compatibilité d'un équipement face aux intervalles infra-journaliers est codée dans un tableau comportant 6 éléments binaires (0  $\equiv$  non compatible, 1  $\equiv$  compatible) à ce stade de la simulation.

39. Dans cette équation, nous utilisons la même notation que MATLAB (« .\* ») pour la multiplication de matrices membre à membre. D'autre part, « <sup>T</sup> » désigne l'opérateur de transposition de matrices et de vecteurs

### Test de compatibilité en équipements

La compatibilité initiale d'un équipement face aux intervalles infra-journaliers doit ensuite être mise à jour en considérant les autres équipements déjà déclenchés dans chacun des intervalles de la journée. Avant d'entrer plus en détails dans les modalités pratiques de cette opération, nous devons apporter des précisions quant aux trois variables que nous avons introduites au sein du paragraphe 3.1.3.1 page 120 que sont l'indice de l'intervalle à considérer ( $\iota_{IAC}$ ), le repère de l'équipement à considérer ( $r_{EAC}$ ) et la durée à considérer ( $d_{AC}$ ).

$\iota_{IAC}$  est un entier, qui pour un équipement de tête ou un équipement sans liaison avec d'autres appareils, est nul. Ainsi, ceci indique que l'ensemble des six intervalles infra-journaliers d'une journée donnée doivent être considérés dans le processus de placement de l'équipement. Pour les équipements avec liaisons, l'indice de l'intervalle à considérer correspond soit à l'intervalle dans lequel l'équipement en relation avec l'appareil considéré a d'ores et déjà été placé, soit à l'intervalle qui correspond à la fin de l'usage de l'équipement lié déjà disposé.

Par exemple pour le couple machine à laver/sèche-linge, l'indice de l'intervalle à considérer pour la machine à laver sera égal à 0 puisque c'est elle qui se place en premier dans une simulation (équipement de tête). Si l'on suppose que l'utilisation du lave-linge a été déclenchée à 17 h 40 à la fin de l'intervalle Après-midi et que la fin de l'usage de cet équipement survient à 19 h 05 au début de l'intervalle Soir (durée d'usage d'1 h 25 dans ce cas), l'intervalle à considérer pour le sèche-linge sera effectivement le Soir ( $\text{IndiceIntervalleAConsiderer} = 5$ ) étant donné que pratiquement<sup>40</sup> cet appareil ne peut pas fonctionner avant cet intervalle.

Pour un équipement de tête ou sans liaison avec d'autres appareils,  $r_{EAC}$  est nul. Pour les équipements séquentiellement liés après un équipement de tête, ce nombre correspond au premier instant de la journée à partir duquel il peut être placé. En ce qui concerne, les équipements liés simultanément, cette valeur est identique à l'instant de déclenchement de l'usage de l'équipement lié correspondant, de manière à pouvoir assurer le lien de simultanéité.

Pour un équipement de tête,  $d_{AC}$  correspond à la durée de l'équipement qui lui est lié séquentiellement après (le cas échéant) de manière à ce que ce dernier puisse être placé dans tous les cas le jour simulé. Cette durée est en particulier à considérer au niveau de l'**IntervalleLimite**. Supposons que pour un jour donné, un cycle de machine à laver et un cycle de sèche-linge doivent être déclenchés. Imaginons qu'il ne reste que l'intervalle infra-journalier Soir pour placer ces deux équipements et que la durée d'usage de sèche-linge est de 45 min, alors la durée à considérer transmise en entrée de la machine à laver sera également de 45 min. Dans le cas le plus « défavorable » (i. e. déclenchement tardif du lave-linge dans l'intervalle Soir), le sèche-linge aura a minima un instant de déclenchement possible qui lui assurera la conduite de son usage à son terme. Pour les autres appareils (équipements liés séquentiellement après ou simultanés), la durée à considérer est nulle.

Pour mener à bien le test de compatibilité en équipements vis-à-vis d'un appareil donné, l'algorithme prend d'abord en considération la valeur de **IndiceIntervalleAConsiderer** pour cet équipement. Si elle est nulle ou si elle vaut 1, l'algorithme teste tous les intervalles infra-journaliers. Au contraire, si **IndiceIntervalleAConsiderer**  $> 1$ , l'algorithme attribuera une incompatibilité en équipements pour les intervalles représentés par les indices compris entre 1 et **IndiceIntervalleAConsiderer** - 1. Ceci permet de limiter les opérations superflues et donc de favoriser la rapidité de l'algorithme. À ce niveau de la procédure, les autres intervalles infra-journaliers sont reconnus a priori éligibles en équipements pour cet appareil avant les contrôles menés sur leurs résumés que nous détaillons ci-dessous.

40. Cette conséquence est valable si l'hypothèse générale suivante est adoptée : dans la liste des tâches d'un jour considéré, si à la fois un cycle de lave-linge et un cycle de sèche-linge sont indiqués, celui du sèche-linge ne pourra être déclenché

Pour chacun des intervalles dont le test en équipements a été validé pour l'appareil considéré à ce stade du processus, son résumé est « balayé » (ou « scanné ») à la recherche d'éventuels équipements liés ou bien du même équipement. Différents cas se présentent alors :

- **le résumé est dans son état initial\*** c'est-à-dire qu'aucun équipement n'est placé dans l'intervalle. Dans ce cas, la compatibilité en équipements est assurée ;
- **des équipements sont déjà placés dans l'intervalle mais il ne s'agit ni de l'équipement considéré à placer ni de ses équipements liés.** Le test de compatibilité en équipements est positif ;
- **le même équipement auto-lié est déjà déclenché dans l'intervalle.** Deux étapes de traitement sont requises :
  1. tout d'abord, il faut savoir combien de fois il est déclenché, soit  $n$  ce nombre ( $n > 0$ ),
  2. ensuite, la **saturation\*** de l'équipement est étudiée. Celle-ci désigne le nombre maximum d'utilisations permises (par l'auteur du scénario simulé) du même équipement auto-lié au sein d'un intervalle<sup>41</sup>. Si la saturation est atteinte, la compatibilité en équipements est nulle. Autrement – la saturation n'est pas atteinte – l'algorithme cherche à savoir si l'équipement peut être placé à nouveau dans l'intervalle conformément à sa durée d'usage et à celle du même appareil déjà déclenché dans l'intervalle. Pour cela sont recueillis l'instant de déclenchement et la durée d'usage de l'équipement déjà déclenché et de l'équipement à placer. Deux cas se présentent alors :
    - (a) l'autorisation de dépassement de l'équipement pour cet intervalle est nulle (i. e. l'intervalle doit contenir entièrement l'équipement sur toute sa durée d'usage). L'algorithme cherche une partie de l'intervalle de longueur suffisante pour le placer. S'il en trouve une, la compatibilité en équipements est vraie, elle est fausse dans le cas contraire ;
    - (b) l'équipement est autorisé à dépasser de l'intervalle. Dans ce cas en plus de la recherche d'une partie de l'intervalle d'une longueur suffisante pour accueillir l'équipement dans sa totalité, l'algorithme évalue la possibilité – le test précédent s'étant révélé négatif – d'inclure dans l'intervalle la **partie non déplaçable de l'usage d'un équipement\*** notée **PNDU\***. Cette dernière correspond à la première séquence continue d'un besoin de disponibilité (non nul) contenue dans la forme développée de l'usage d'un équipement (première suite continue de 1). Elle a une durée notée  $d_{PNDU}$ . Si l'utilisation de l'équipement dans sa totalité ou a minima la partie non déplaçable de son usage peuvent être « accueillis » dans l'intervalle, le test de compatibilité est validé pour cet appareil ;
- **le même équipement n'est pas présent dans l'intervalle mais l'un de ses équipements liés est déjà déclenché.** L'algorithme réitère le processus identique au cas de l'équipement auto-lié en investiguant en plus la modalité de liaison (séquentielle, simultanée) entre les deux équipements et en y tenant compte au cours du test ;
- **dans tous les autres cas**, nous faisons l'hypothèse par défaut que la compatibilité en équipements est vraie. De cette façon, nous évitons de tester de manière exhaustive l'ensemble des autres configurations possibles.

qu'après la fin de l'usage de la machine à laver

41. Par défaut, nous avons fixé ce paramètre égal à 2

### Test de compatibilité en temps

Une fois la compatibilité en équipements testée, l'algorithme considère ensuite les intervalles sous l'aspect temporel. Comme pour le test de compatibilité en équipements, cette étape n'est pas opérée sur l'ensemble des intervalles infra-journaliers. Suivant la valeur de l'`IndiceIntervalleAConsiderer` s'appliquent les mêmes restrictions qu'auparavant :

- considération de l'ensemble des intervalles si `IndiceIntervalleAConsiderer`  $\leq 1$  ;
- analyse des intervalles `IndiceIntervalleAConsiderer` à 6 dans les autres cas.

Ensuite, en fonction de la possibilité de l'équipement à dépasser ou non de l'intervalle, deux cas sont à étudier :

- l'équipement n'est pas autorisé à dépasser de l'intervalle. L'algorithme compare la durée d'usage de l'équipement à la longueur de l'intervalle. Le test est vrai si et seulement si  $d_{Usage} \leq l_{Intervalle}$  ;
- l'équipement peut dépasser de l'intervalle. Le précédent test est réalisé. S'il échoue, l'algorithme compare  $l_{Intervalle}$  et  $d_{PNDU}$ . Le test de compatibilité en temps est alors valide si et seulement si  $d_{PNDU} \leq l_{Intervalle}$ .

Le graphique 3.12 illustre les deux tests que nous venons d'évoquer.

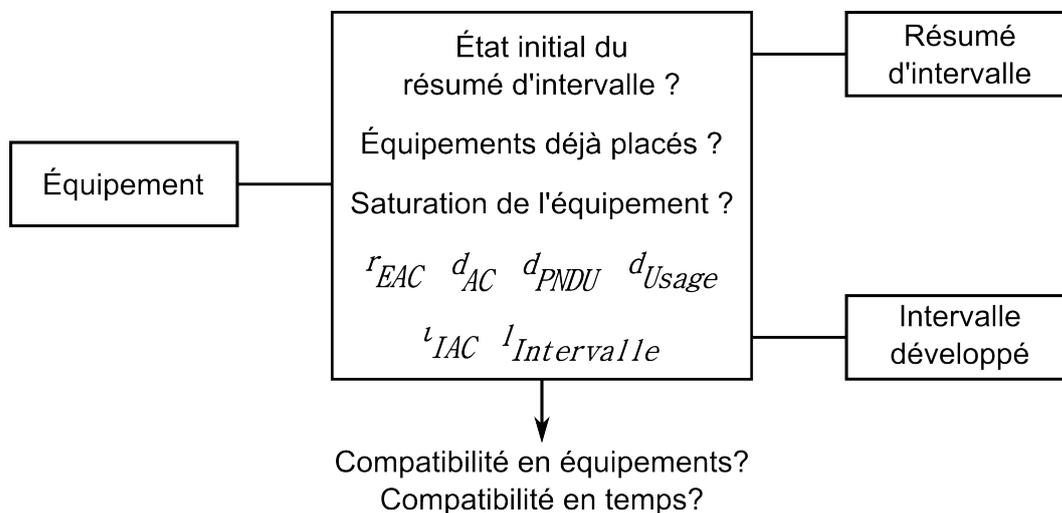


FIGURE 3.12 – Illustration des tests de compatibilité en équipements et en temps

Codage de la compatibilité de l'équipement pour les intervalles qui lui sont éligibles

Les deux précédents tests servent à ajuster la compatibilité par défaut (copie de la compatibilité de la micro-tâche) d'un équipement face aux intervalles infra-journaliers. Nous introduisons les notations suivantes :  $c_i(\iota)$  représente la compatibilité d'un équipement  $i$  par rapport à l'intervalle  $\iota$ ,  $c_{MTP|i}(\iota)$ <sup>42</sup> désigne la compatibilité de la micro-tâche parente ( $MTP^*$ ) de l'équipement  $i$  considéré<sup>43</sup> par rapport à l'intervalle  $\iota$  et enfin  $c_{EE,i}(\iota)$  (respectivement  $c_{ET,i}(\iota)$ ) traduit la compatibilité en équipements (respectivement en temps) de l'équipement  $i$  par rapport à l'intervalle infra-journalier  $\iota$ , compatibilité issue du résultat (0 ou 1) du test correspondant. Avec les notations précédentes, la mise à jour de la compatibilité de l'intervalle  $\iota$  face à l'équipement domestique  $i$  est opérée suivant la relation 3.4.

$$c_i(\iota) = c_{MTP|i}(\iota) \times c_{EE,i}(\iota) \times c_{ET,i}(\iota) \quad (3.4)$$

Sont ainsi mis en lumière les intervalles infra-journaliers effectivement éligibles pour l'équipement considéré conformément à leurs caractéristiques propres. Cette information est synthétisée dans le tableau de compatibilité équipement/intervalles.

Afin d'affiner ce caractère éligible « factuel » de par son codage binaire, nous décidons de le quantifier sous forme de probabilité. Nous noterons  $p_{\iota,i}$ , la propension – exprimée en tant que probabilité – d'un intervalle  $\iota$  à accueillir l'instant de déclenchement d'un appareil domestique  $i$ . Pour y parvenir, nous avons élaboré une méthode qui s'est construite en plusieurs étapes et dont nous ne donnons ci-après que la version finale<sup>44</sup>.

*De la compatibilité équipement/intervalles à des propensions des appareils face aux intervalles infra-journaliers*

La sélection de l'intervalle infra-journalier qui accueillera l'instant de déclenchement de l'usage d'un équipement est fonction du rapport 3.5.

$$\frac{\text{Nombre de cas favorables}^*}{\text{Nombre de cas possibles}^*} \quad (3.5)$$

On appelle cas possible un élément temporel (une « case ») de la forme développée d'un intervalle infra-journalier et cas favorable, un cas possible non précisément connu dans le temps (la recherche des cas favorables se faisant par blocs) qui peut accueillir le déclenchement de l'usage d'un équipement. Le nombre de cas favorables est noté  $n_{CF}$  (`NbCasFavorables` dans notre modèle). Pour un intervalle donné, l'opération de calcul du nombre de cas favorables est conduite si et seulement si la compatibilité en équipements et en temps de l'intervalle pour l'appareil considéré est établie<sup>45</sup>. Plus concrètement, le rapport 3.5 est estimé au moyen du quotient 3.6.

$$\frac{\text{Nombre de cas favorables dans l'intervalle considéré}}{\text{Somme des cas favorables dans l'ensemble des intervalles}} \quad (3.6)$$

Les probabilités  $p_{\iota,i}$  évoquées ci-dessus correspondent au rapport 3.6 calculé pour chaque intervalle éligible à l'équipement considéré. La détermination du nombre de cas favorables dans chacun de ces intervalles éligibles dépend du signe de la différence  $d_{Usage} - l_{Intervalle}$ , de la permission de l'équipement à dépasser de l'intervalle, du caractère limite de chacun des intervalles traités pour l'équipement étudié, de la durée de la partie non déplaçable de l'usage de l'équipement ( $d_{PNDU}$ ) et de la *durée tampon de l'usage\** de l'équipement notée  $d_{Tampon}$ . Cette dernière désigne la longueur de la suite continue de zéros, si elle existe, qui succède directement à la partie non déplaçable de l'usage d'un équipement dans la représentation de son usage sous forme développée. Nous illustrons ces propos par l'exemple présenté dans le tableau 3.8.

42. On note  $MTP | i$  le lien hiérarchique existant entre l'équipement  $i$  et sa micro-tâche parente

43. De manière équivalente, l'équipement  $i$  est sous-jacent à la micro-tâche

	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$
$d_{\beta, Durée}$ [min]	8	12	20	0	5
$d_{\beta, Disp}$ [-]	1	0	0	0	1

TABLEAU 3.8 – Exemple d’usage d’un équipement avec des  $d_{PNDU}$  et  $d_{Tampon}$  non nulles (8 et 32 min respectivement)

L’usage d’un équipement présenté dans le tableau 3.8 se décompose comme suit :

- une partie non déplaçable d’une durée de 8 min ( $d_{PNDU} = d_{1, Durée}$ ),
- une partie tampon d’une durée de 32 min qui regroupe  $d_2$  et  $d_3$  puisque pendant le fonctionnement de l’appareil assurant l’usage, l’intervention humaine n’est pas requise ( $d_{3, Disp} = 0$ ) et
- 5 min de disponibilité sont nécessaires une fois terminé le fonctionnement de l’équipement ( $d_{5, Durée} = 5$  et  $d_{5, Disp} = 1$ ).

L’ensemble des configurations possibles de placement d’un équipement dans un intervalle sont répertoriées dans le tableau 3.9<sup>46</sup>. Chacune d’elles aboutit à un nombre bien particulier de cas favorables dont le calcul est explicité au sein de l’annexe B, dans le tableau B.1. Nous détaillons ci-après la lecture de ce tableau en suivant les critères qui différencient les cas et nous mentionnons chaque configuration au travers des valeurs de ces critères.

Tout d’abord, recenser le nombre de cas favorables pour le placement d’un équipement consiste à passer en revue tous les intervalles reconnus jusqu’alors éligibles après la transmission aux équipements sous-jacents de la compatibilité des micro-tâches parentes et les tests de compatibilité en équipements et en temps. Pour un même équipement, différents cas peuvent donc s’appliquer en fonction de l’intervalle traité c’est-à-dire de l’intervalle éligible considéré.

Ensuite et avant tout calcul, le nombre de cas favorables pour chaque intervalle est fixé à zéro. En effet étant donné que les formules de calcul de cas favorables peuvent aboutir à des résultats négatifs, elles sont systématiquement conditionnées par des tests de façon à empêcher de telles situations absurdes. Les cas de placement ne passant pas les tests sont donc logiquement des configurations dans lesquelles aucun cas favorable n’est envisageable d’où la mise à zéro par défaut du nombre de cas favorables de chaque intervalle.

44. Les changements successifs de paradigme s’expliquent par le besoin d’une prise en compte de situations plus complexes, en matière de placement d’équipements, que celles auxquelles nous avons initialement songé

45. À ce stade n’est donc pas testé par exemple le cas suivant :  $d_{Usage} > l_{Intervalle}$  avec un équipement non autorisé au dépassement d’intervalle. Le nombre de cas favorables est automatiquement fixé à 0 dans ce cas

46. Pour faciliter sa lecture, nous adoptons les conventions d’écriture suivantes :  $IT^*$   $\equiv$  Intervalle Traité,  $IL^*$   $\equiv$  Intervalle Limite et  $IAC^*$   $\equiv$  Intervalle À Considérer

De plus et de manière générale, les cas de placement sont répartis en deux catégories principales : d'un côté, I et II concernent les équipements liés puisque l'indice de l'intervalle à considérer,  $\iota_{IAC}$ , est non nul. À l'opposé, III et IV s'appliquent aux équipements de tête.

Quatre critères viennent ensuite s'ajouter aux précédentes considérations :

1. le premier (Critère 1) consiste à savoir si l'intervalle traité est l'intervalle limite. Si c'est le cas, l'équipement dans la totalité de sa durée d'usage doit impérativement être placé dans l'intervalle observé. L'autorisation de dépassement d'intervalle n'est donc pas testée (cas I.A, I.B et III) ;
2. le second (Critère 2) détermine le fait que l'intervalle traité est ou non l'intervalle à considérer. Si tel est le cas, il faut tenir compte du repère de l'équipement à considérer,  $r_{EAC}$ , pour le calcul du nombre de cas favorables de manière explicite (cas I.A, II.A.1, II.A.2 et II.A.3.2) ou implicite (cas II.A.3.1) par restriction de la zone de recherche dans l'intervalle. À noter que ce critère n'est évidemment pas pertinent dès lors que l'indice de l'intervalle à considérer est nul (cas III et IV) ;
3. le critère suivant (Critère 3) consiste à étudier le statut de l'équipement par rapport à l'intervalle traité : est-il autorisé ou non à dépasser de cet intervalle ? Ce critère n'est pas systématiquement testé puisque les deux précédents peuvent être suffisamment contraignants pour déterminer à leur niveau le nombre de cas favorables de placement de l'usage de l'équipement (cas I.A, I.B et III). Si l'équipement n'est pas autorisé à dépasser de l'intervalle, le calcul de cas favorables tiendra compte du fait qu'il devra être intégralement disposé dans l'intervalle. Au contraire, si une partie de l'usage de l'équipement peut être reportée sur l'intervalle suivant, il faut estimer quelle partie de la forme développée de l'usage de l'équipement peut être accueillie dans l'intervalle traité. Ceci constitue le dernier critère ;
4. le Critère 4 permet de finaliser, si besoin est, la différenciation entre les configurations de placement d'un équipement. Il consiste à identifier la part de la forme développée de l'usage de l'équipement qui peut être incluse dans l'intervalle. Le test est en particulier affiné lorsque l'équipement est autorisé à dépasser de l'intervalle. Si c'est le cas, trois configurations sont possibles :
  - (a) l'intégralité de la forme développée de l'usage de l'équipement peut être accueillie dans l'intervalle. Le nombre de cas favorables synthétise dans ce cas les possibilités d'accueil complet de l'usage de l'équipement dans l'intervalle ainsi que les opportunités de placement liées à son possible report dans l'intervalle suivant (cas II.A.2, II.B.2 et IV.B) ;
  - (b) la forme développée de l'usage de l'équipement ne peut pas être contenue entièrement dans l'intervalle mais l'intégralité de la partie non déplaçable et de la partie tampon de l'usage de l'équipement peuvent l'être (cas II.A.3.1, II.B.3.1 et IV.C.1) ;
  - (c) la forme développée de l'usage de l'équipement ne peut pas être contenue entièrement dans l'intervalle et seule la partie non déplaçable de l'usage de l'équipement peut l'être (cas II.B.3.2 et IV.C.2).

		Critère 1	Critère 2	Critère 3	Critère 4
	Code cas	$IT = IL$	$IT = IAC$	Autorisation dépassement intervalle	Accueil de la forme développée de l'usage de l'équipement dans l'intervalle
$\iota_{IAC} \geq 1$	I.A	oui	oui	non testée	non testé
	I.B	oui	non	non testée	non testé
	II.A.1	non	oui	non	intégralité de $d_{Usage}$
	II.A.2	non	oui	oui	intégralité de $d_{Usage}$
	II.A.3.1	non	oui	oui	intégralité de $d_{PNDU} + d_{Tampon}$
	II.A.3.2	non	oui	oui	intégralité de $d_{PNDU}$
	II.B.1	non	non	non	intégralité de $d_{Usage}$
	II.B.2	non	non	oui	intégralité de $d_{Usage}$
	II.B.3.1	non	non	oui	intégralité de $d_{PNDU} + d_{Tampon}$
	II.B.3.2	non	non	oui	intégralité de $d_{PNDU}$
$\iota_{IAC} = 0$	III	oui	non pertinent	non testée	non testé
	IV.A	non	non pertinent	non	intégralité de $d_{Usage}$
	IV.B	non	non pertinent	oui	intégralité de $d_{Usage}$
	IV.C.1	non	non pertinent	oui	intégralité de $d_{PNDU} + d_{Tampon}$
	IV.C.2	non	non pertinent	oui	intégralité de $d_{PNDU}$

TABLEAU 3.9 – Configurations possibles pour le calcul du nombre de cas favorables dans un intervalle infra-journalier pour le placement d'un équipement

Recenser le nombre de cas favorables ne suffit cependant pas pour assurer un placement convenable de l'équipement. En effet, en plus de la connaissance de l'éventail des possibles, il faut également savoir où se situent globalement ces instants dans les intervalles. Concrètement pour un intervalle donné, il faut déterminer si l'usage de l'équipement est contraint au report dans un intervalle infra-journalier postérieur à l'intervalle traité ou non. Si la nécessité de report de l'utilisation de l'équipement est effective, les cas favorables au placement de l'équipement sont naturellement situés à la fin de l'intervalle traité et il faut ainsi passer en revue l'intervalle en partant de  $t_{Final}$ . Dans le cas contraire, ils s'organisent à partir du début de l'intervalle, il faut donc commencer les recherches à partir de  $t_{Initial}$ .

Dans notre modèle, l'information traduisant le report nécessaire est représentée par un vecteur de taille 6 noté **ReportNecessaire**, qui comporte le sens de recherche des cas favorables dans chacun des intervalles infra-journaliers. La convention suivante a été choisie :  $0 \equiv$  recherche à partir du début de l'intervalle,  $1 \equiv$  recherche à partir de la fin de l'intervalle.

Sur le même modèle que l'évaluation du nombre de cas favorables dans les intervalles infra-journaliers, nous détaillons dans le tableau 3.10, l'impact des différentes configurations de recherche des cas favorables sur l'affectation de la valeur de **ReportNecessaire** $[\iota]$ , c'est-à-dire, intervalle par intervalle. Nous précisons ci-dessous les configurations de calcul des éléments du vecteur **ReportNecessaire**.

Code cas	$IT = IAC$	Autorisation dépassement intervalle	Accueil de la forme développée de l'usage de l'équipement dans l'intervalle	ReportNecessaire [ $\iota$ ]
A.1	oui	non	intégralité de $d_{Usage}$	0
A.2	oui	oui	intégralité de $d_{Usage}$	0
A.3.1	oui	oui	intégralité de $d_{PNDU} + d_{Tampon}$	1
A.3.2	oui	oui	intégralité de $d_{PNDU}$	1
B.1	non	non	intégralité de $d_{Usage}$	0
B.2	non	oui	intégralité de $d_{Usage}$	0
B.3.1	non	oui	intégralité de $d_{PNDU} + d_{Tampon}$	1
B.3.2	non	oui	intégralité de $d_{Tampon}$	1

TABLEAU 3.10 – Évaluation des composantes du vecteur **ReportNecessaire** selon les configurations possibles de recherche des cas favorables dans les intervalles infra-journaliers

Tout d'abord, l'impact du premier critère de différenciation n'est pas explicitement visible dans le tableau 3.10. Il s'agit de tester si l'intervalle traité est l'intervalle à considérer. Ceci influence en fait les tests sur la valeur prise par la variable  $r_{EAC}$  : 0 si l'intervalle traité n'est pas l'intervalle à considérer (il n'y a pas d'équipement à prendre en considération), une valeur strictement positive dans le cas contraire.

Le second critère d'influence correspond à la permission de l'équipement de dépasser ou non de l'intervalle traité. Ce critère n'est pas à lui seul suffisant pour évaluer une composante de **ReportNecessaire**. Il faut en effet connaître quelle est la part de la forme développée de l'usage de l'équipement qui peut être accueillie dans l'intervalle.

La portion de la forme développée de l'usage de l'équipement pouvant être placée dans l'intervalle constitue donc le dernier critère. Couplée à la précédente, cette information détermine la valeur des composantes de **ReportNecessaire** et plusieurs cas se présentent :

- si la forme développée de l'usage de l'équipement peut être intégralement accueillie dans l'intervalle  $\iota$ , le report n'est pas nécessaire et ce quel que soit le statut de l'équipement par rapport au dépassement d'intervalle (i. e. qu'il soit autorisé ou non à dépasser) : **ReportNecessaire** [ $\iota$ ] = 0 (cas A.1, A.2, B.1 et B.2) ;
- dans le cas contraire, le report est nécessaire quelle que soit la part de la forme développée de l'usage de l'équipement qui peut être admise dans l'intervalle ( $d_{PNDU} + d_{Tampon}$  ou  $d_{Tampon}$  seulement) : **ReportNecessaire** [ $\iota$ ] = 1 (cas A.3.1, A.3.2, B.3.1 et B.3.2).

## Sélection de l'intervalle parmi l'ensemble des possibles

Les probabilités  $p_{l,i}$  calculées précédemment <sup>47</sup> permettent de transformer la compatibilité binaire des équipements face aux intervalles infra-journaliers en pondérations qui indiquent « l'attrance » d'un intervalle vis-à-vis d'un équipement. La sélection de l'intervalle de déclenchement de l'usage d'un appareil ménager valorise donc cette série de pondérations. Ces dernières définissent une **distribution discontinue de probabilités**.

Dans notre modèle, nous avons donc mis en place une fonction capable d'assurer la sélection, sous la contrainte de la distribution de probabilité mentionnée ci-avant, d'un intervalle infra-journalier dans lequel a lieu le déclenchement de l'usage de l'équipement considéré.

### 3.2.4.3 Recensement des instants éligibles de déclenchement dans l'intervalle sélectionné

Après avoir choisi un intervalle pour l'équipement considéré, l'algorithme scanne au sein de cet intervalle tous les instants éligibles au déclenchement de l'usage de cet équipement. La notion d'instant éligible englobe celle de cas favorable car cet instant, en plus de pouvoir potentiellement accueillir le lancement d'un appareil, est précisément connu dans le temps.

Concrètement, la forme développée de l'usage de l'équipement est mise en regard de celle de l'intervalle retenu pour son déclenchement, toutes deux étant construites au pas de temps de la simulation. En fonction :

- du nombre de cas favorables calculé précédemment et qui tient compte des caractéristiques en durée de l'usage de l'équipement et de l'intervalle (prise en considération notamment de  $d_{Usage}$ ,  $d_{PN DU}$ ,  $d_{Tampon}$  et de  $l_{Intervalle}$ ),
- de la présence de l'équipement dans l'intervalle si celui-ci est auto-lié,
- de la présence dans l'intervalle d'un ou plusieurs de ses possibles équipements liés,
- de la nature de ces éventuelles liaisons,
- du caractère limite de l'intervalle,
- de l'autorisation ou non de l'équipement à dépasser de cet intervalle,
- de la nécessité ou non de l'équipement à être reporté sur l'intervalle suivant,

la forme développée de l'intervalle, ou seulement une ou plusieurs de ses parties, est (sont) considérée(s) par le processus de recherche des instants éligibles pour le déclenchement de l'usage de l'équipement.

Quel que soit le domaine de définition sur lequel s'opère cette recherche, elle demeure séquentielle : l'intervalle développé est parcouru dans le sens chronologique (si le report de l'équipement n'est pas nécessaire) ou dans l'ordre inverse si l'équipement est contraint au report. Pour chaque élément de la forme développée de l'usage de l'équipement (ou simplement sa partie non déplaçable s'il est autorisé à dépasser de l'intervalle sélectionné), le niveau correspondant de disponibilité requise est comparé à celui de son « vis-à-vis » dans la forme développée de l'intervalle.

Les représentations 1 à 4 de la figure 3.13 illustrent le processus précédemment décrit en faisant l'hypothèse que l'équipement n'est pas forcé au report et qu'aucune contrainte liée à  $r_{EAC}$  ou en matière d'équipements liés ne vienne complexifier son placement.

Les illustrations de la figure 3.13 ne témoignent que du principe général de la recherche des instants éligibles. En effet, suivant la multiplicité de la combinatoire des critères précédemment évoqués, le processus de balayage prend des formes diverses. Nous traitons synthétiquement en annexes, au sein du tableau B.2, chacune des configurations de recherche des instants éligibles au déclenchement de l'usage d'un équipement dans un intervalle.

47. Pour rappel, celles-ci sont estimées suivant l'équation 3.6 et peuvent être calculées dès lors que le nombre de cas favorables pour chacun des intervalles est connu

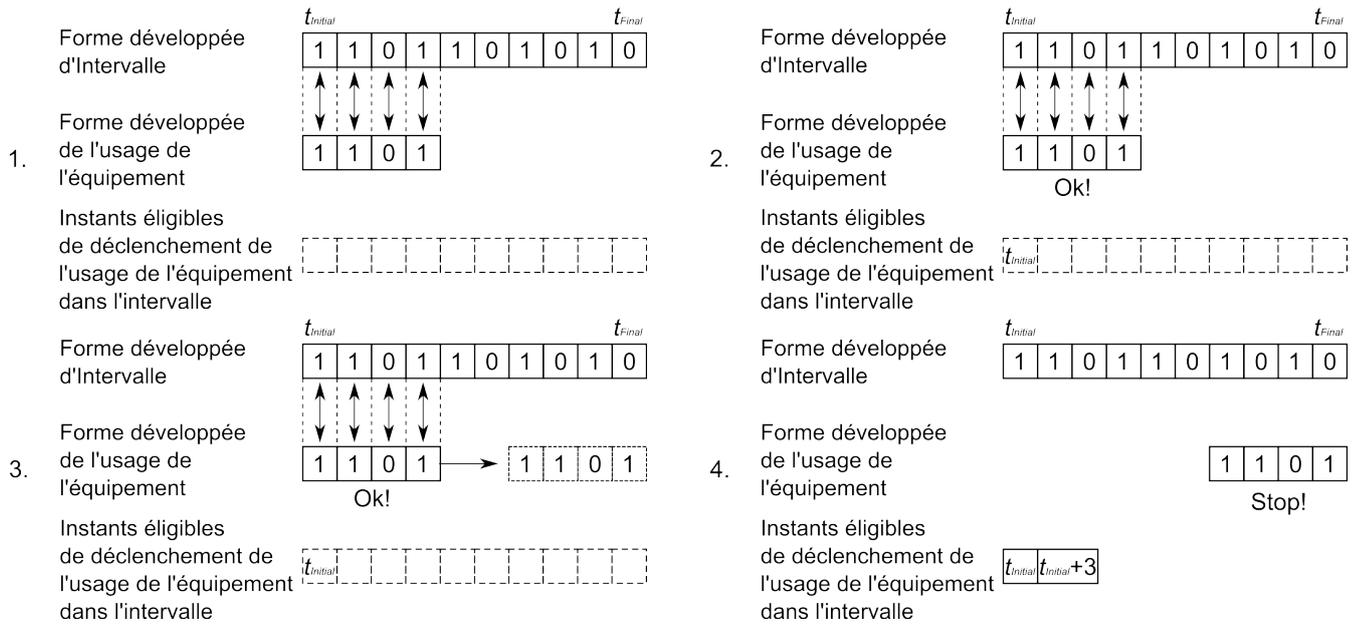


FIGURE 3.13 – Étapes 1 à 4 du processus de recensement des instants éligibles au sein d'un intervalle pour le déclenchement de l'usage d'un équipement

**3.2.4.4 Sélection d'un instant de déclenchement pour l'usage de l'équipement considéré parmi l'ensemble des possibles**

Cette étape consiste à choisir par un tir aléatoire un instant de déclenchement pour l'usage de l'équipement considéré parmi l'ensemble des moments éligibles identifiés à l'étape précédente.

**3.2.4.5 Mise à jour de l'intervalle**

Sitôt l'instant de déclenchement sélectionné pour l'usage de l'équipement considéré, une double mise à jour doit être menée : au niveau de l'intervalle développé d'une part et de son résumé d'autre part.

La première consiste à prendre en compte le besoin de disponibilité pour l'usage de l'équipement par rapport au potentiel de disponibilité (humaine) dans l'intervalle. Une soustraction membre à membre est donc effectuée entre la forme développée de l'intervalle et celle de l'usage de l'appareil. En accord avec les caractéristiques de l'équipement placé, le processus est opéré soit sur toute la durée de l'usage de celui-ci comme l'illustrent les représentations 1 à 3 de la figure 3.14<sup>48</sup>, ou seulement sur sa partie non déplaçable et sur tout ou partie de sa durée tampon le cas échéant.

48. L'instant encadré sur le schéma 1 de cette figure correspond au moment de déclenchement choisi par l'algorithme pour l'usage de l'équipement considéré

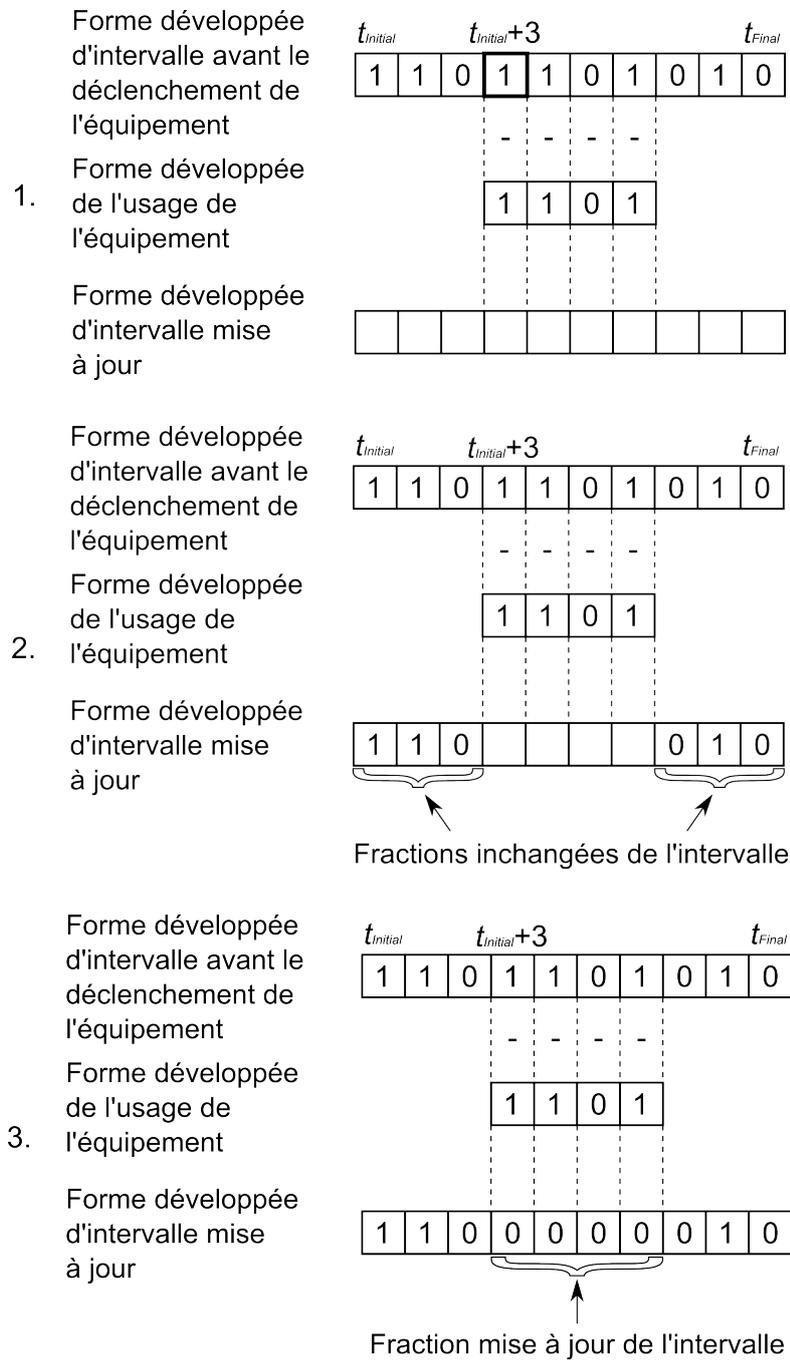


FIGURE 3.14 – Étapes 1 à 3 de la mise à jour de l'intervalle développé suite au placement de l'usage d'un équipement

La mise à jour du résumé de l'intervalle infra-journalier, qui a été retenu pour accueillir l'instant de déclenchement de l'usage de l'équipement, consiste à y ajouter l'événement [instant de déclenchement, signature d'équipement] en respectant l'ordre chronologique qui caractérise le résumé. Pour cela, une fonction dédiée a été développée dans notre modèle. Celle-ci inscrit l'événement au sein du résumé d'intervalle conformément à la position qui lui correspond en matière de temporalité. Les illustrations 1 à 4 de la figure 3.15 schématisent le travail de cette fonction.

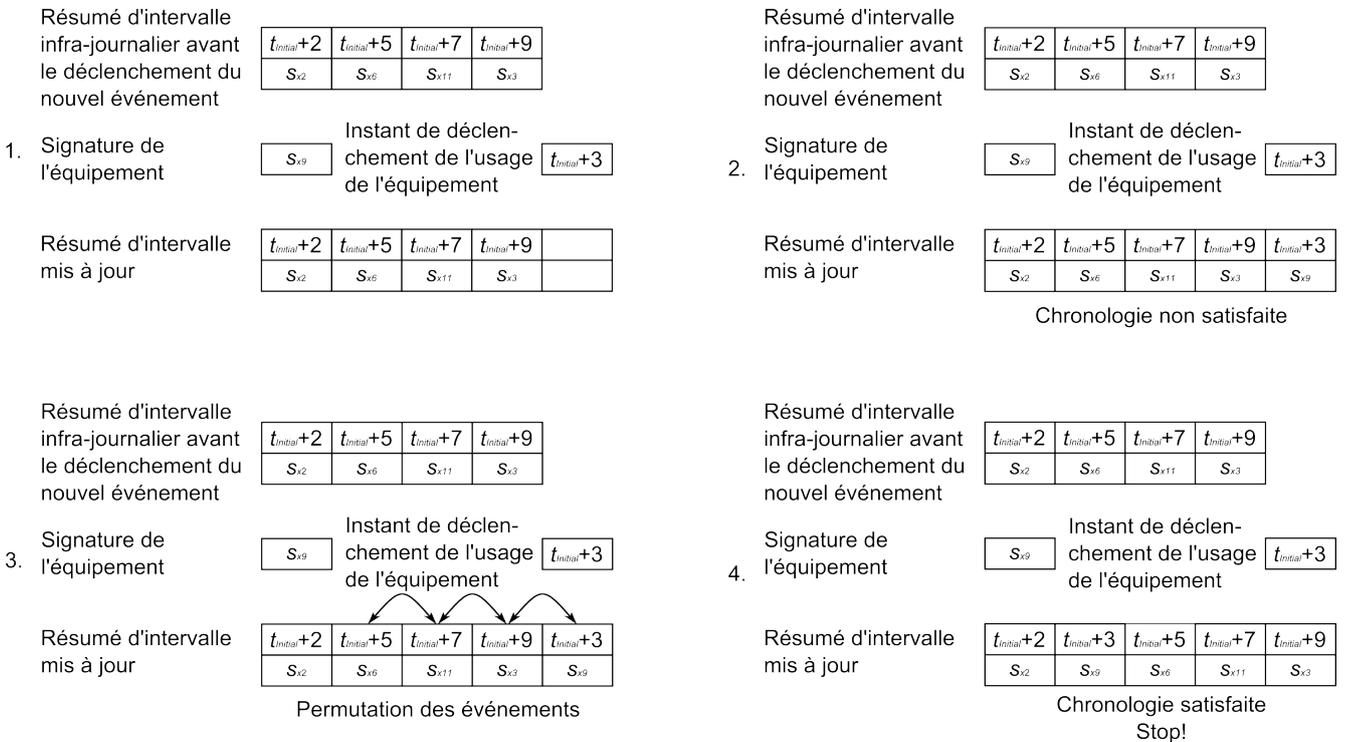


FIGURE 3.15 – Étapes 1 à 4 de la mise à jour d'un résumé d'intervalle suite au placement de l'usage d'un équipement

L'algorithme de placement des équipements au sein d'une journée constitue la brique opérationnelle indispensable pour la reconstitution de la courbe de charge électrique des équipements domestiques. Celui-ci permet notamment de simuler une première série de scénarios de modélisation.

Nous détaillerons dans la partie 4.2 du chapitre 4 à la fois les scénarios que nous avons construits et la manière avec laquelle ils sont élaborés mais aussi les résultats en puissance qui ont été obtenus par simulation.

Avant de conclure ce chapitre, nous proposons dans la partie 3.3 une mise en perspective de l'algorithme de placement des équipements domestiques, et ce, à deux niveaux. Dans un premier temps, il s'agit de prendre du recul par rapport à notre travail en le synthétisant et ensuite nous replaçons notre méthode vis-à-vis des modèles de la littérature.

### 3.3 Mise en perspective de l'algorithme de placement des équipements domestiques

Comme nous l'avons suggéré ci-dessus, cette partie a deux objectifs majeurs. Le premier consiste à résumer ce qui a été décrit dans l'ensemble de ce chapitre et pour ce faire, nous proposons une représentation graphique de synthèse. En effet, dans les pages précédentes, nous avons décrit l'algorithme dans le détail, description accompagnée d'un certain nombre de schémas indispensables à la compréhension de chaque procédure logique ou de calcul mise en jeu dans le processus de génération de foisonnement comporte-

mental. Cela dit, chacune des illustrations mentionnées au préalable ne représente au mieux qu'une des étapes de l'algorithme<sup>49</sup> sans mettre en perspective ce dernier dans son ensemble. Il manque donc une illustration qui compile la globalité du processus : des données d'entrée de l'algorithme aux courbes de charge électriques. C'est pourquoi nous proposons la figure 3.16 qui expose l'intégralité de l'algorithme et le replace dans son contexte c'est-à-dire en regard de notre objectif de reconstitution de la courbe de charge résidentielle.

La figure 3.16 se lit « naturellement » de haut en bas et représente la partie centrée sur l'algorithme de la chaîne de simulation que nous avons construite au cours de nos travaux de recherche.

Partant d'une distribution des tâches journalières pour un couple ménage-type/jour-type (cf. partie 3.2.1), d'un découpage du jour simulé en intervalles infra-journaliers amenant à l'établissement d'un profil de disponibilité (paragraphe 3.2.2 et 3.2.3) et de règles permettant de construire la matrice de macro-compatibilité des tâches domestiques (voir l'introduction de la section 3.2.4), la matrice de micro-compatibilité peut être mise en place (paragraphe 3.2.4.1) tout comme la forme modélisée des équipements simulés et des intervalles infra-journaliers (voir les parties 3.1.1 à 3.1.3). Ensuite, la compatibilité équipement/intervalles est construite par défaut pour chaque appareil modélisé en entrée<sup>50</sup> d'algorithme. L'objectif de ce dernier (sélectionner un instant de déclenchement pour l'usage de chaque équipement domestique considéré dans une simulation) implique d'abord la mise à jour de cette compatibilité équipement/intervalles au moyen de tests (en équipements et en temps) et ensuite la conversion de cette information – initialement codée sous forme binaire – en probabilités (cf. partie 3.2.4.2). Arrivé à cette étape, il ne suffit plus que de choisir parmi les possibles un intervalle compatible avec le déclenchement de l'usage de l'appareil étudié, puis au sein de cet intervalle, de recenser les instants éligibles et d'en choisir un parmi eux (cf. sections 3.2.4.3 et 3.2.4.4).

Des étapes de mise à jour des données manipulées sont ensuite nécessaires afin d'être en mesure de répéter les précédentes phases de l'algorithme, ceci pour chacun des autres appareils domestiques modélisés dans le scénario considéré (voir la partie 3.2.4.5). Cette dernière étape correspond à la prise en compte dans la modélisation, des interactions pouvant exister entre les appareils domestiques au sein d'un foyer.

Au final, pour chaque équipement nous obtenons simultanément un scénario d'utilisation de cet appareil sous forme d'une distribution des instants de déclenchement de l'usage de celui-ci pour le scénario modélisé ainsi que les courbes de charge journalières associées.

Après cette prise de recul quant aux développements présentés dans ce chapitre, nous proposons maintenant de mettre en perspective notre méthode de génération de foisonnement d'origine comportementale par rapport aux autres méthodologies utilisées dans les modèles de la littérature, modèles que nous avons analysés au sein du chapitre 1.

Tout d'abord, faisons un bref rappel de ce chapitre premier (cf. partie 1.4) et plus particulièrement en ce qui concerne la manière avec laquelle nous avons alors défini et décrit les méthodes d'obtention de diversité proposées par les auteurs que nous avons lus. En laissant de côté les modèles qui n'explicitent pas le foisonnement<sup>51</sup> (celui-ci peut notamment être inclus dans des données d'entrée sans que d'autres procédures ne viennent améliorer cette représentation déterministe), il demeure quatre catégories de méthodes qui le génèrent véritablement. Suivant l'échelle de représentation de la diversité<sup>52</sup> que nous avons établie dans le chapitre d'analyse bibliographique, le foisonnement ainsi modélisé est respectivement de type déterministe scénarisé, aléatoire mathématique, probabilité empirique et probabilité scénarisée. L'algorithme auquel nous avons abouti produit justement un foisonnement de cette dernière catégorie.

La différence essentielle entre notre méthode et les modèles de génération de foisonnement qui s'inscrivent dans cette classe de qualité supérieure<sup>53</sup> réside dans le fait que dans notre cas, l'aspect scénarisé ne

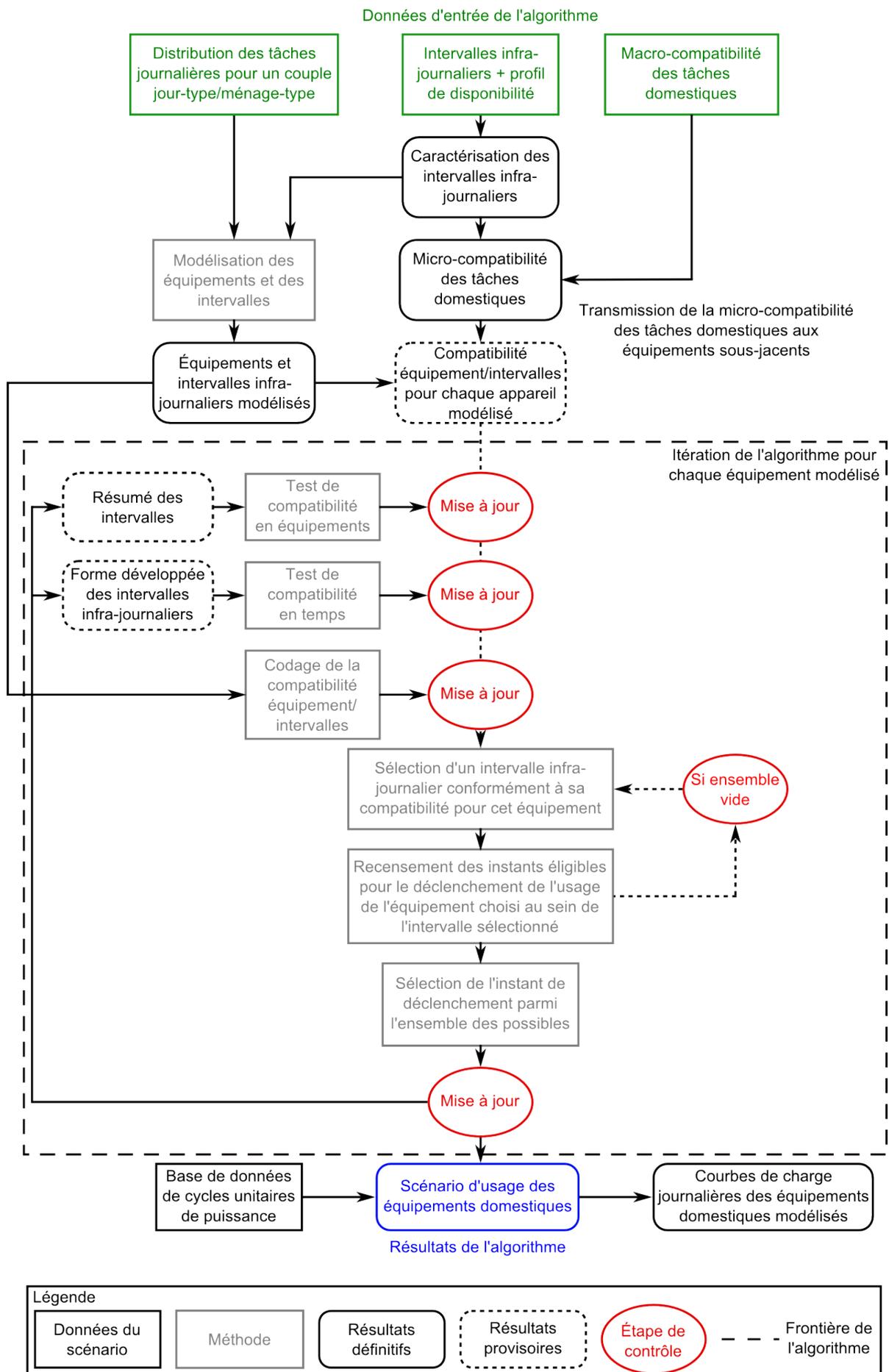
49. Pour la plupart, il ne s'agit qu'une des phases d'un processus correspondant à une unique étape de l'algorithme

50. À ce niveau, il s'agit d'une donnée temporaire

51. C'est le cas notamment des approches d'Aigner et al. [ASK84], de Bartels et al. [BFGL92] et de Train et al. [THW85]

52. Introduite à la page 17, celle-ci est reprise au travers de la figure 1.15

53. Font partie de cette catégorie les modèles de Walker et Pokoski [WP85], Capasso et al. [CGLP94], Armstrong et al.



Légende

Données du scénario	Méthode	Résultats définitifs	Résultats provisoires	Étape de contrôle	Frontière de l'algorithme
---------------------	---------	----------------------	-----------------------	-------------------	---------------------------

FIGURE 3.16 – Synthèse de l'algorithme de placement des équipements domestiques et mise en perspective dans son contexte

provient pas de mesures « prêtes à l'emploi » comme peuvent l'être des distributions de probabilité d'une activité domestique ou bien de l'usage d'un appareil ménager. Ces distributions peuvent être elles-mêmes construites à partir de la compilation de données recueillies dans des cahiers de bord rédigés par des ménages sélectionnés lors d'enquêtes publiques<sup>54</sup>. **Dans notre approche au contraire, nous aboutissons à des distributions analogues grâce à une série de règles de comportement et de contraintes qui sont à la fois construites par le modélisateur et par les opérations de calcul de l'algorithme.**

D'autre part, comme il l'a été rappelé plus haut, notre méthode tient à la fois compte de la disponibilité des membres des foyers mais aussi des liens pouvant exister entre les équipements domestiques au sein d'un logement. Cette dernière caractéristique paraît originale puisqu'elle ne semble pas avoir été retenue par les auteurs de la littérature dont nous avons consulté les travaux.

Définir plus précisément les propriétés du foisonnement synthétisé par notre algorithme impose la réalisation d'un grand nombre de simulations et ce, pour différents scénarios. Dans le chapitre 4, nous initions cette étude sur les résultats des simulations que nous avons menées durant la thèse. Nous tenterons en particulier d'établir une caractérisation de la diversité d'origine comportementale qui affecte les appels de puissance des ménages modélisés.

## Conclusion du chapitre

Dans ce troisième chapitre, nous avons abordé la modélisation d'un certain nombre de concepts de la méthode de reconstitution de courbe de charge présentée au chapitre 2. Plus précisément, la modélisation a été axée de sorte de pouvoir représenter la diversité d'origine comportementale qui se traduit notamment au niveau du ménage par des préférences en matière d'instant de déclenchement de l'usage des équipements. Pour ce faire, nous avons élaboré un algorithme stochastique capable de reproduire l'activation des appareils ménagers au sein d'un foyer et à l'échelle journalière. L'utilisation des équipements ainsi modélisée se veut réaliste autant que faire se peut. La temporalité du déclenchement des appareils tient bien entendu compte du facteur humain (présence au domicile des occupants des foyers dans un état d'éveil compatible avec le maniement des équipements) mais aussi des liens existants entre les appareils eux-mêmes. Ce chapitre a donc été l'occasion de détailler très précisément les différents éléments modélisés ainsi que les procédures permettant de les faire interagir.

Finalement, nous avons synthétisé nos travaux de modélisation et nous les avons mis en regard des autres modèles de la littérature.

La modélisation des concepts et l'algorithme stochastique mettent en place les éléments indispensables à la construction de scénarios qui peuvent ensuite être simulés. Ceci constitue donc une application directe des travaux détaillés dans ce chapitre. La construction des scénarios et l'analyse de leurs résultats en puissance sont deux thématiques qui seront abordées dans le dernier chapitre de ce manuscrit. De plus, il y sera également question de la recherche d'une méthode efficace pour étudier le phénomène de foisonnement qui impacte les courbes d'appels de puissance à un niveau agrégé.

D'autre part, de manière à alimenter les scénarios modélisés, nous avons cherché à identifier des comportements réels d'usage des équipements. Pour y parvenir, nous avons effectué différents travaux de classification de courbes de charge que nous présentons dans le chapitre 4.

[ASR<sup>+</sup>09], Widén et Wäckelgård [WW09] et Richardson et al. [RTIC10]

54. Aux États-Unis, il s'agit de l'« American Time Use Survey » [UNI12], au Canada l'étude se nomme « Social Survey – Time Use » [STA10], au Royaume-Uni l'enquête est appelée « Time Use Survey » [ECO00], enfin en France c'est l'INSEE qui réalise l'« Enquête Emploi du Temps » [INS10a]

## Chapitre 4

# Classification de courbes de charge et étude du foisonnement

*« La connaissance s'acquiert par l'expérience, tout le reste n'est que de l'information. »*  
Albert Einstein

### Introduction du chapitre

L'objectif de reconstitution de la courbe de charge électrique résidentielle implique un travail qui va au-delà de l'analyse bibliographique, des développements conceptuels et de leur mise en application que nous avons respectivement présentés au sein des trois précédents chapitres. En effet, il nous a semblé essentiel dans ce travail de thèse d'y associer une composante expérimentale qui permette de traiter des données simulées et empiriques de manière à améliorer notre connaissance de la courbe de charge. Nous développons dans le présent chapitre les différentes recherches que nous avons menées dans ce domaine.

Trois problématiques découlent de l'objectif de la thèse. La première concerne l'observation et la quantification du foisonnement, phénomène qui impacte les courbes agrégées d'appels de puissance. Notre objectif ici est d'aboutir au critère de convergence  $\varepsilon$  que nous avons introduit au paragraphe 2.3.3.1 du chapitre 2. Pour rappel, celui-ci renseigne de l'obtention d'un niveau de foisonnement suffisant entre courbes de charge agrégées et indique ainsi le nombre de simulations unitaires nécessaires pour y aboutir. Dans cette optique, nous avons réalisé une revue des indicateurs caractéristiques de la courbe de charge. Simultanément, nous nous sommes intéressés à la recherche d'une mesure pour quantifier l'écart éventuel entre courbes d'appels de puissance. Ces deux sujets sont abordés dans la partie 4.1.

La seconde problématique consiste en l'étude et en l'analyse des résultats de notre modèle obtenus par simulation. Ceci permet d'initier une première évaluation de notre méthode de synthèse de courbe de charge notamment en ce qui concerne sa capacité de génération de diversité d'origine comportementale. Pour ce faire, nous détaillons dans la section 4.2 les scénarios que nous avons modélisés au cours de nos recherches. Nous décrivons dans le même paragraphe, quelques-uns des résultats en puissance obtenus par simulation.

Pour dépasser la simple étude descriptive des résultats précédemment mentionnés, nous avons mis en œuvre une partie des techniques et des méthodes abordées dans la section 4.1. Nous avons ensuite pu mener une analyse très fine et originale des courbes de charge générées par notre modèle. Ceci est présenté dans le paragraphe 4.3.

Enfin, la troisième problématique cherche à mettre en lumière les ressemblances entre courbes de charge unitaires lorsque l'observation visuelle n'est plus pertinente, c'est-à-dire dès que sont considérées une dizaine de courbes environ. Plus précisément, ce travail vise à l'identification de « comportements types » concernant l'utilisation des équipements électriques domestiques. Dans la section 4.4, nous détaillons notre travail de classification de courbes d'appels de puissance qui a été mené pour l'usage de la télévision.

## 4.1 Caractérisation et comparaison de courbes de charge

Au cours de notre analyse des modèles de la littérature, nous avons très vite constaté qu'un des enjeux importants sur lesquels les auteurs portaient leur attention et orientaient leurs efforts était celui de la comparaison de courbes de charge, qu'elles soient mesurées, simulées ou reconstituées. In fine, les auteurs recherchent un moyen de tester les résultats de leur modèle, résultats qui sont le plus souvent mis en regard de données empiriques.

Dans la plupart des cas, un unique indicateur, calculé entre différentes courbes de charge est adopté : ceci permet en une seule valeur de résumer l'information disponible mais simultanément ce genre de représentation ne restitue que partiellement la richesse des données. Un certain nombre d'indicateurs de ce type existent dans la littérature et nous les présenterons dans la partie 4.1.1. Autant que faire se peut, nous les illustrerons par des exemples soulignant leurs avantages mais aussi leurs limites.

Au regard de ces derniers éléments, il nous est apparu indispensable de rechercher une distance plus adéquate à l'étude de courbes de charge que les indicateurs précédemment évoqués. Nous traiterons cette question dans la partie 4.1.2.

Quantifier la proximité entre courbes de charge peut être conduit de diverses manières dans la mesure où, en fonction du contexte, différents éléments d'intérêt peuvent être pris en considération. Parmi eux, les deux dimensions qui nous semblent les plus pertinentes sont d'une part, le niveau d'agrégation des courbes de charge noté  $N_{agr}$ , pour lesquelles la comparaison est menée (courbes de charge unitaires – *CDCU* – ou plus particulièrement courbes de charge moyennes foisonnées – *CDCMF* – construites à partir de l'agrégation de plusieurs *CDCU*) et d'autre part l'intérêt éventuellement différent qui est porté à la courbe de charge en fonction de l'heure de la journée (typiquement la pointe de consommation du soir peut être suivie de façon privilégiée).

### 4.1.1 Indicateurs statistiques existants pour la comparaison de courbes de charge

Comme nous l'avons abordé au travers du chapitre 1 et au paragraphe 2.3.6.2 du chapitre 2, nous détaillons ci-dessous les coefficients statistiques que nous avons identifiés lors de nos recherches et qui s'appliquent aux courbes de charge.

#### 4.1.1.1 Indicateurs statistiques classiques

Pour caractériser une courbe de charge, il est possible de faire appel aux très nombreux éléments de statistique descriptive tels que le minimum, le maximum, le mode, la médiane, la moyenne, l'écart-type... Nous n'avons pas jugé utile de rappeler ici leur définition respective et nous invitons le lecteur à se référer au chapitre 5 de [SAP06] si besoin. À la différence des indicateurs que nous verrons par la suite, ces éléments peuvent être calculés pour une courbe de charge unique : par abus de langage, nous les appellerons les **indicateurs unaires**\*. Du fait de leur simplicité, nous ne rappellerons pas leur définition mais nous illustrons quelques uns d'entre eux au moyen de la figure 4.1.

Malgré leur grand nombre et leur diversité, ces éléments pris individuellement ne semblent pas suffisants pour comparer des courbes de charge entre elles : en effet le plus souvent, ils n'offrent qu'une vision très partielle des appels de puissance. Toutefois, lorsqu'ils sont regroupés sur une même représentation graphique, leur intérêt s'avère supérieur puisque de cette façon ils proposent un moyen rapide pour illustrer la dispersion des valeurs de puissance qui constituent les courbes de charge observées. La figure 4.2 illustre nos propos en présentant les boîtes à moustaches de quelques courbes de charge journalières. Nous rappelons qu'une boîte à moustaches est construite des éléments suivants [SAP06] :

**la boîte** contient 50 % des valeurs de la série, valeurs comprises entre le premier et le troisième quartile<sup>1</sup> de la série (respectivement notés  $qle_1$  et  $qle_3$ ). La médiane est représentée par un trait horizontal au travers de la boîte (la moyenne peut également être indiquée au moyen d'une croix) ;

1. Rappel : à l'image de la médiane qui sépare une population en deux échantillons de taille identique, les quartiles segmentent une population en quatre échantillons d'effectif égal

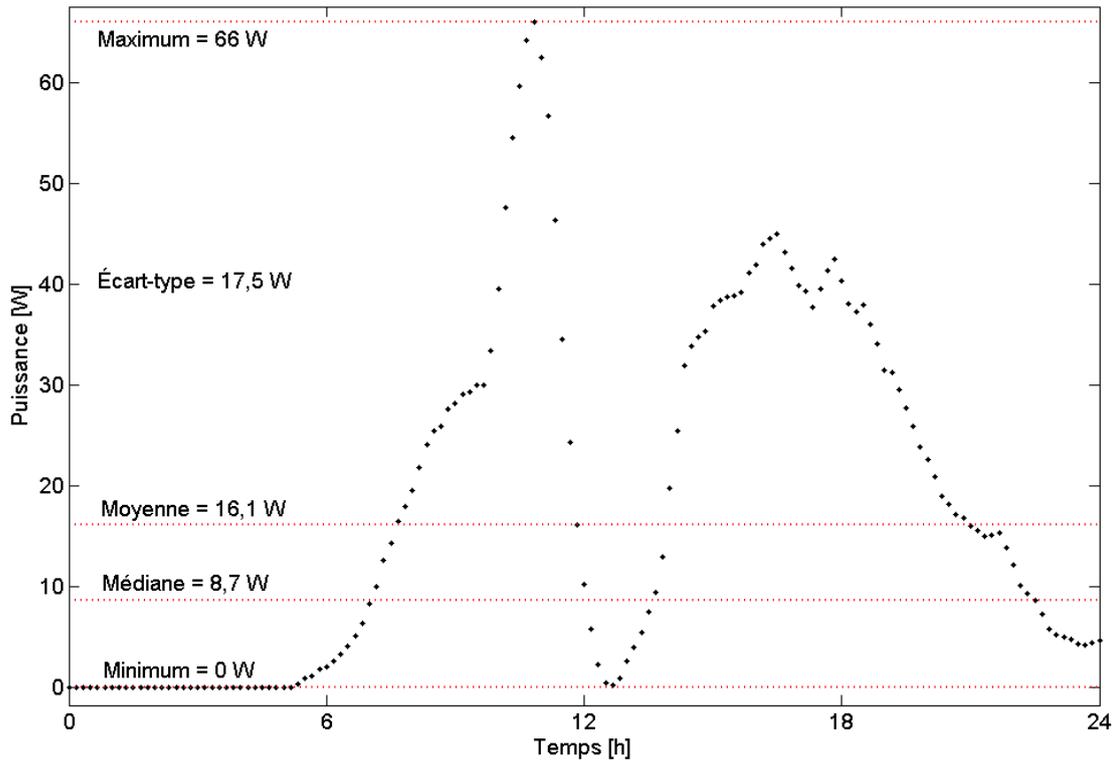


FIGURE 4.1 – Illustration sur une courbe de charge journalière de quelques indicateurs de statistique descriptive

**les moustaches** donnent une idée de l'étendue des valeurs de la série. À partir des extrémités supérieure et inférieure de la boîte, les limites des moustaches sont distantes d'une fois et demie l'espace interquartile, c'est-à-dire  $1,5 \times (q_{le3} - q_{le1})$ , s'il existe des valeurs dans la série étudiée au-delà de ces limites ; sinon dans le cas contraire, les moustaches s'arrêtent aux extremums de la série ;

**les valeurs hors norme** sont celles qui excèdent les limites des moustaches ; elles peuvent être schématisées par des étoiles ou des croix.

Après cette courte revue des indicateurs unaires, nous allons désormais nous intéresser aux **indicateurs binaires\***, c'est-à-dire ceux dont le calcul n'a de sens que lorsque plusieurs courbes de charge sont comparées entre elles deux à deux. Précisons que ces indicateurs sont bien adaptés à l'étude de courbes de charge puisqu'ils sont traditionnellement utilisés pour la description de séries temporelles. Dans la suite de ce paragraphe, nous faisons l'hypothèse que ces grandeurs sont calculées pour comparer une courbe de charge mesurée et une courbe de charge estimée (calculée ou simulée).

### Erreur absolue moyenne

L'erreur absolue moyenne ou **Mean Absolute Error (MAE)\*** est une grandeur qui caractérise l'écart entre deux séries temporelles. Celle-ci s'exprime dans l'unité des éléments considérés (le watt pour des courbes de charge). L'erreur absolue moyenne entre une courbe de charge mesurée  $P_{mes}$  et une courbe de charge estimée  $P_{est}$  se calcule conformément à la relation 4.1.

$$MAE = \frac{1}{n} \cdot \sum_{t=1}^n \left| P_{mes}(t) - P_{est}(t) \right| \quad (4.1)$$

Dans la précédente équation,  $n$  correspond au nombre d'instant  $t$  sur lesquels le calcul est opéré (par exemple, pour une courbe de charge journalière au pas horaire,  $n$  sera égal à 24).

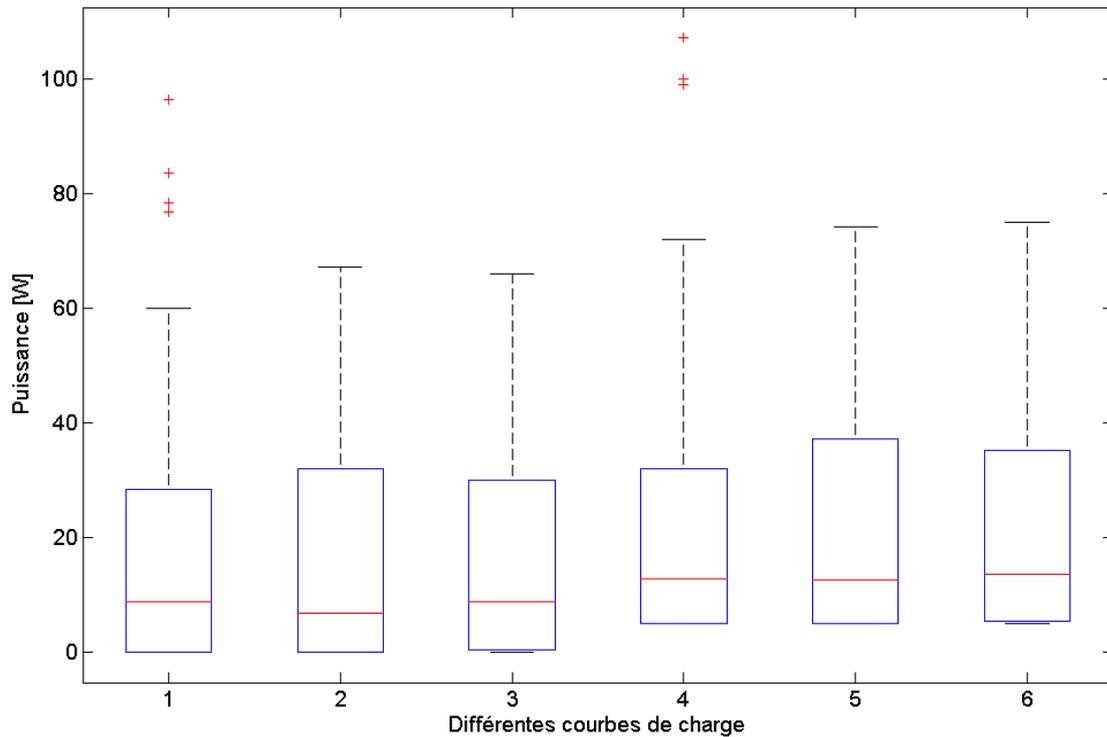


FIGURE 4.2 – Représentation de plusieurs courbes de charge journalières sous forme de boîtes à moustaches

Dans le cas de courbes de charge, le *MAE* donne donc une information d'« éloignement » sur l'axe vertical entre deux représentations d'appels de puissance. Puisqu'il s'agit d'une valeur moyenne, celle-ci est affectée par les valeurs extrêmes des écarts. De plus, cette grandeur ne traduit pas la manière avec laquelle les écarts se répartissent sur l'axe temporel des appels de puissance. L'exemple de la figure 4.3 est en cela très significatif puisqu'il montre deux couples de courbes de charge ([Courbe de charge de référence, Courbe de charge 1] et [Courbe de charge de référence, Courbe de charge 2])<sup>2</sup> avec des *MAE* identiques (dans ce cas il vaut 60 W) mais avec une répartition des écarts très différente. Les courbes de charge du premier couple diffèrent en chaque point alors que celles du second couple sont très voisines hormis en deux instants qui correspondent aux pointes de consommation.

### Erreur absolue moyenne en pourcentage

L'erreur absolue moyenne en pourcentage (*MAPE*) est très similaire au précédent indicateur à la différence près que cette grandeur n'a pas d'unité : elle s'exprime en pourcentage. Pour « éliminer » l'unité, une référence est définie : dans le cas de courbes de charge, la référence est souvent prise sur l'existant c'est-à-dire sur des données passées mesurées. Soulignons de plus que malgré sa proximité mathématique avec le *MAE*, si deux couples de courbes de charge ont le même *MAE*, elles n'auront pas nécessairement le même *MAPE* (c'est notamment le cas des couples de courbes présentées au travers de la figure 4.3). Le *MAPE* entre deux courbes de charge est calculé suivant l'équation 4.2

$$MAPE = \frac{1}{n} \cdot \sum_{t=1}^n \left| \frac{P_{est}(t) - P_{mes}(t)}{P_{mes}(t)} \right| \quad (4.2)$$

où comme précédemment,  $P_{est}(t)$  est la valeur estimée de l'appel de puissance à l'instant  $t$  ( $t = 1, \dots, n$ ) alors que  $P_{mes}(t)$  est la valeur mesurée. Pour chaque pas de temps, la valeur absolue de l'écart entre

2. Sur la figure 4.3, ces courbes de charge sont respectivement notées  $CDC_{ref}$ ,  $CDC_1$  et  $CDC_2$

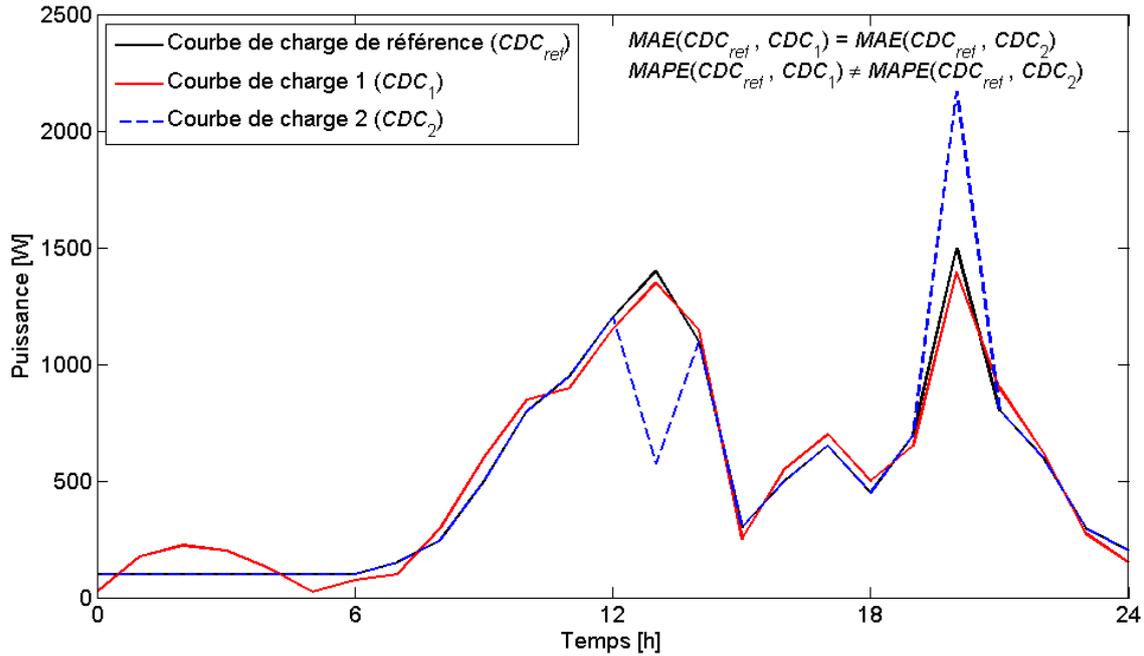


FIGURE 4.3 – Illustration de l'utilisation du  $MAE$  et du  $MAPE$  pour la comparaison de courbes de charge journalières

puissance mesurée et puissance estimée est calculée. Cette grandeur est ensuite divisée par la puissance mesurée à l'instant  $t$  prise pour référence : un écart relatif entre mesure et estimation est donc obtenu. Le  $MAPE$  indique donc la valeur moyenne en pourcentage de l'écart relatif précédemment explicité. Plus le  $MAPE$  entre deux courbes de charge est faible et plus celles-ci sont proches. Les limites du  $MAPE$  sont les mêmes que celles du  $MAE$ . Néanmoins, le fait de construire les écarts relatifs à chaque pas de temps crée une pondération des différences entre courbes de charge. L'information portée par le  $MAPE$  est donc plus riche que celle donnée par le  $MAE$ .

### Erreur quadratique moyenne

L'erreur quadratique moyenne ou **Root Mean Square Error ( $RMSE$ )**\* est un indicateur qui se construit sur le modèle de l'écart-type. En effet, il correspond à la racine carrée de la moyenne des carrés résiduels entre courbes de charge (la moyenne étant estimée sur la longueur totale des courbes de charge comparées<sup>3</sup>), où les carrés résiduels désignent la différence, élevée à la puissance 2, de la valeur des courbes de charge à un instant  $t$  donné. L'erreur quadratique moyenne se calcule au moyen de l'équation 4.3 [TUF10].

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{t=1}^n (P_{est}(t) - P_{mes}(t))^2} \quad (4.3)$$

Le  $RMSE$  est un indicateur qui s'exprime dans l'unité des objets manipulés, le watt pour les courbes de charge. En comparaison du  $MAE$ , il se distingue par le fait que les différences entre courbes de charge sont d'autant plus accentuées puisqu'elles sont élevées à la puissance 2. À nouveau, plus le  $RMSE$  sera faible et plus l'écart moyen entre courbes de charge sera limité. Malgré tout, cet indicateur souffre des mêmes limites que le  $MAE$  du fait qu'il s'agit d'une moyenne, il est très influencé par les valeurs extrêmes des différences entre courbes de charge.

3. Il va sans dire que le calcul du  $RMSE$  ne vaut que si les courbes de charge comparées sont de même longueur

### Indice d'accord

L'indice d'accord ou **Index of Agreement (IA)**\* proposé par Willmott dans [WIL81] et discuté par la suite notamment dans [WIL82], est un indicateur adimensionnel borné par 0 et 1. Cet indice a avant tout pour vocation d'évaluer le degré de fiabilité (ou ce qui revient au même l'absence d'erreurs) d'un modèle de prédiction. La reconstitution de la courbe de charge et la comparaison des résultats obtenus à des données mesurées se situent donc tout à fait dans le champ d'application de cette grandeur. Cette mesure normalisée permet d'effectuer simplement des comparaisons entre modèles : un indice d'accord de 1 correspond à une parfaite concordance entre mesures et prédictions alors que la valeur nulle de cet indicateur témoigne de l'incapacité prédictive du modèle considéré. L'indice d'accord se calcule à l'aide de l'équation 4.4

$$IA = 1 - \frac{\sum_{t=1}^n (P_{est}(t) - P_{mes}(t))^2}{\sum_{t=1}^n (|P_{est}(t) - \overline{P_{mes}}| + |P_{mes}(t) - \overline{P_{mes}}|)^2} \quad (4.4)$$

dans laquelle  $\overline{P_{mes}}$  désigne la valeur moyenne de la puissance mesurée pour  $t = 1, \dots, n$  (les autres notations ayant déjà été commentées au préalable). Des variantes de l'indice d'accord existent et sont abordées dans [WAD<sup>+</sup>85]. Elles offrent notamment la possibilité de pondérer certaines valeurs prédites par rapport aux autres. Plus complexe que le *RMSE*, cette valeur souffre cependant du fait qu'il s'agit d'un quotient de moyennes, les valeurs extrêmes des écarts entre prédictions et réalisations influent donc grandement sur la valeur finale de cet indicateur. À notre connaissance, cet indicateur n'a pas été employé pour caractériser des modèles de courbe de charge.

Les différents éléments précédemment exposés sont directement issus du monde statistique : du fait qu'ils puissent être appliqués à des séries temporelles, ils sont a fortiori utilisables sur des courbes de charge électriques.

En plus de ces éléments, divers indicateurs ont été spécialement définis pour l'étude de courbes d'appels de puissance. Nous les exposons dans les prochains paragraphes.

#### 4.1.1.2 Facteur de charge

Le facteur de charge de la puissance maximale est un rapport adimensionnel et borné par 0 et 1. Cet indicateur se construit en divisant la valeur moyenne d'une courbe de charge échantillonnée à différents instants  $t$  sur une période notée  $\Delta t$  par la valeur maximale de cette série temporelle. Typiquement, il peut s'agir de courbes de charge journalières connues au pas horaire. Dans la suite de ce paragraphe et sans exclure tout autre possibilité, nous considérerons qu'il s'agit de courbes de charge obtenues par simulation. Le facteur de charge de la puissance maximale noté  $F_{P_{max}}(\Delta t)$  se calcule à l'aide de l'équation 4.5.

$$F_{P_{max}}(\Delta t) = \frac{\overline{P(\Delta t)}}{P_{max}(\Delta t)} \quad (4.5)$$

Dans cette équation,  $\Delta t$  est la durée de simulation considérée sur laquelle  $P_{max}$  et  $\overline{P}$ , respectivement pointe de puissance et puissance moyenne, peuvent être aisément déterminées. Le facteur de charge témoigne, dans une certaine mesure, de la dispersion des valeurs de puissance qui composent la courbe de charge étudiée. Plus précisément, il renseigne du caractère chahuté ou non des appels de puissance : plus grande sera la valeur du facteur de charge (i. e. proche de 1) et davantage la courbe de charge sera dite « plate ». Au contraire une courbe de charge ayant un facteur de charge faible (proche de zéro) pourra être considérée comme très chahutée. Nous illustrons ces deux cas au moyen de la figure 4.4.

Le facteur de charge peut être calculé sur la durée totale d'une simulation ou sur un ou plusieurs intervalle(s) temporel(s) plus restreint(s). Bien entendu, ce rapport peut être également calculé sur une (des) période(s) équivalente(s) de mesure. La comparaison de plusieurs facteurs de charge entre courbes

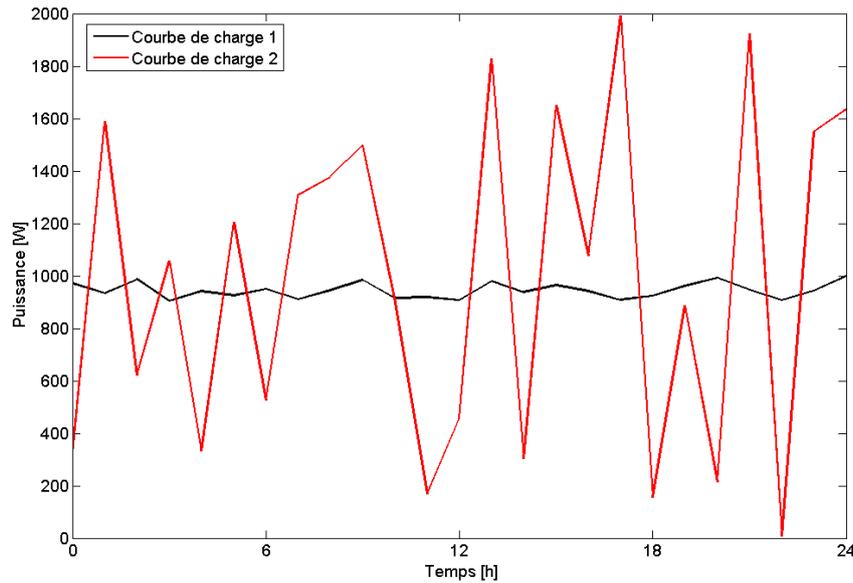


FIGURE 4.4 – Illustration du facteur de charge calculé sur une période  $\Delta t = 24$  h dans le cas d’une courbe de charge « plate » (courbe de charge 1,  $F_{P_{max}} = 0,945$ ) et d’une courbe de charge « chahutée » (courbe de charge 2,  $F_{P_{max}} = 0,498$ )

d’appels de puissance de différentes natures (courbes de charge mesurées et simulées par exemple) offre une vision globale mais limitée de la forme de ces courbes comparées.

#### 4.1.1.3 Coefficient de foisonnement

Le coefficient de foisonnement ou coincidence factor<sup>4</sup> est un indicateur adimensionnel et borné dans l’intervalle  $[0; 1]$  [LAN85, LAN87, PAT89]. Historiquement, il s’est imposé tout naturellement dès le début de l’histoire des réseaux de distribution d’électricité [HUG83]. Il renseigne de l’hétérogénéité avec laquelle un groupe de  $n$  individus – des consommateurs domestiques par exemple – font l’usage de l’énergie électrique, tous les équipements étant confondus<sup>5</sup>. Plus précisément, cet indicateur noté  $K_{f,n}(\Delta t)$  compare la pointe de puissance des  $n$  consommateurs sur une période  $\Delta t$ ,  $P_{max,n}(\Delta t)$  appelée **pointe concomitante\*** des  $n$  utilisateurs, avec la somme des pointes individuelles,  $P_{max,j}$ , de chaque individu  $j$  ( $j = 1, \dots, n$ ), sur la même période. Le coefficient de foisonnement s’obtient par la relation 4.6.

$$K_{f,n}(\Delta t) = \frac{P_{max,n}(\Delta t)}{\sum_{j=1}^n P_{max,j}(\Delta t)} \quad (4.6)$$

Si cet indicateur est proche de zéro, cela signifie que le foisonnement (i. e. la diversité de la demande en puissance) est élevé : les appels de puissance sont assez peu concomitants et la charge (puissance) totale maximale du réseau est relativement faible par rapport à la somme de ses charges partielles (une charge partielle désigne ici la puissance maximale appelée par un utilisateur raccordé). Inversement, un coefficient de foisonnement proche de l’unité est synonyme d’une répartition médiocre des instants d’usage des appareils électriques : une grande partie des charges partielles sont appelées à un niveau proche de leur maximum et de manière simultanée si bien que la pointe sur le réseau s’approche de la somme des

4. À noter que dans la littérature anglo-saxonne, c’est davantage le diversity factor, i. e. l’inverse du coincidence factor qui est utilisé

5. Cette version est la plus commune du coefficient de foisonnement ; cependant, il est possible d’en définir d’autres notamment en considérant les équipements individuellement [ORP99]

charges qui le constituent. Une illustration du coefficient de foisonnement est proposée au travers de la figure 4.5. La comparaison du coefficient de foisonnement des courbes de charge mesurées et modélisées donne là encore une vue d'ensemble de la vraisemblance de la modélisation des appels de puissance.

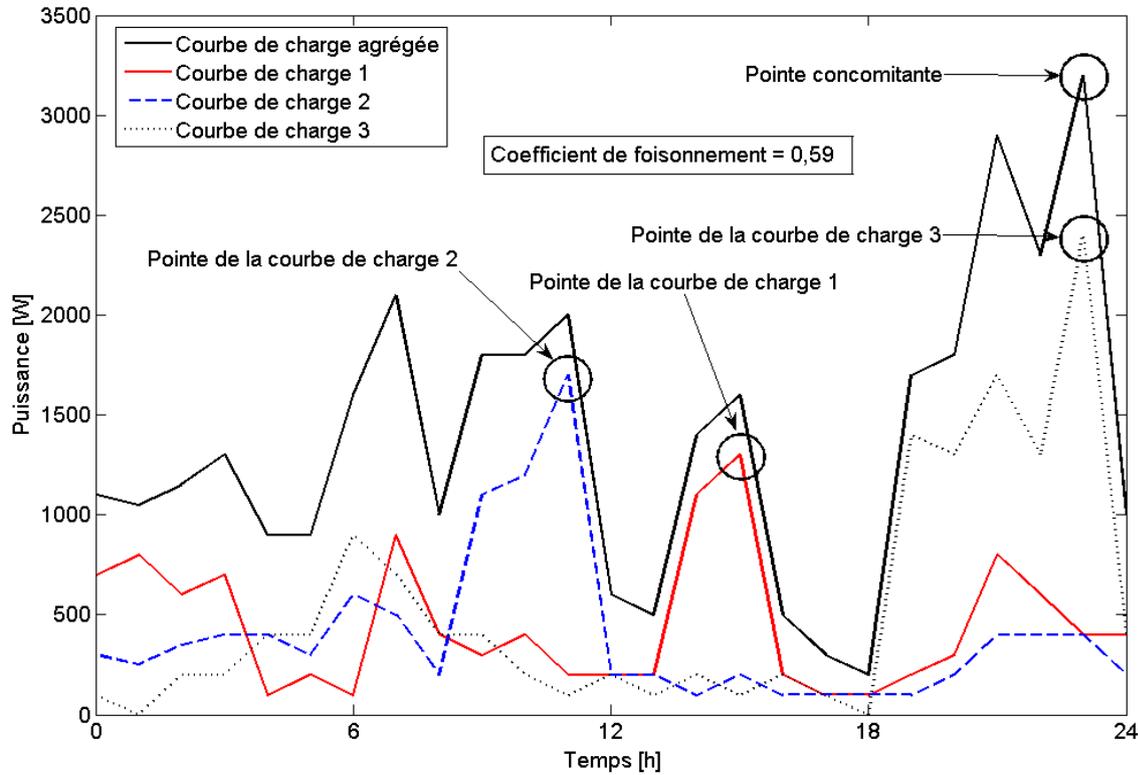


FIGURE 4.5 – Illustration du coefficient de foisonnement d'une courbe de charge agrégée résultante de la somme de trois courbes de charge individuelles

#### 4.1.1.4 Normalized Variation Factor

Le Normalized Variation Factor (*NVF*), qui peut se traduire par facteur normalisé de variation<sup>6</sup>, est un indicateur largement utilisé dans les publications scientifiques que nous avons analysées au chapitre 1 [WP85, TRA92, CGLP94, WLW<sup>+</sup>09]. Le *NVF* entre une courbe de charge estimée  $P_{est}$  et une courbe de charge mesurée  $P_{mes}$  est défini par l'équation 4.7.

$$NVF = \frac{1}{n} \cdot \frac{\sum_{t=1}^n (P_{est}(t) - P_{mes}(t))^2}{\left(\frac{1}{n} \cdot \sum_{t=1}^n P_{mes}(t)\right)^2} \quad (4.7)$$

Le *NVF* est le rapport de la moyenne des carrés des écarts entre puissance mesurée et puissance estimée d'une part et le carré de la puissance moyenne mesurée prise comme référence d'autre part. Plus cette fraction est petite, meilleure est l'estimation de la courbe de charge par rapport à la référence choisie. De par sa construction, le *NVF* n'est borné que par valeur inférieure (0) mais il n'est pas possible d'en définir un majorant numérique. C'est notamment pour cette raison que dans [CGLP94], cet indicateur est calculé d'une part entre courbes de charge mesurée et simulée pour un jour donné et d'autre part entre courbes

6. Cette appellation n'a pas été rencontrée dans la littérature

de charge mesurées de deux jours consécutifs. Ceci permet de se donner une idée de la variation inter-journalière possible de cet indicateur et donc d'estimer les écarts entre mesure et simulation relativement à cette donnée.

Le *NVF* clôture cette synthèse sur les indicateurs caractéristiques des courbes de charge.

#### 4.1.2 Prise en compte de la dimension temporelle de la courbe de charge

Les indicateurs statistiques présentés dans la partie 4.1.1 ne sont pas suffisants pour un travail de comparaison de courbes de charge en particulier lorsque la problématique du foisonnement est abordée. **Leur principal défaut réside dans le fait que seules les différences de niveau de puissance sont considérées alors qu'en réalité la problématique du déphasage entre courbes s'ajoute à cette dimension.** En conséquence, nous avons cherché une distance adéquate à la comparaison de courbes de charge c'est-à-dire un moyen de quantifier à la fois l'écart en puissance et en temporalité qui peuvent exister entre deux courbes d'appels de puissance.

Dans un premier temps et pour mieux comprendre les développements de ce paragraphe et de la partie 4.4, nous rappelons la définition de dissimilarité et de distance. Soient trois objets  $\underline{o}$  (notés  $o_1$ ,  $o_2$  et  $o_3$ ) appartenant à un ensemble  $\underline{\Omega}$ , une dissimilarité est une application  $\underline{\varphi}$  de  $\Omega \times \Omega$  dans  $\underline{\mathbb{R}}^+$  telle que [SAP06] :

$$\begin{cases} \varphi(o_1, o_2) = \varphi(o_2, o_1) \\ \varphi(o_1, o_2) \geq 0 \quad \varphi(o_1, o_1) = 0 \end{cases} \quad (4.8)$$

Une distance  $\underline{\delta}$  est un cas particulier d'une dissimilarité puisqu'elle doit satisfaire les quatre conditions suivantes [SAP06] :

$$\begin{cases} \delta(o_1, o_2) = \delta(o_2, o_1) \\ \delta(o_1, o_2) \geq 0 \\ \delta(o_1, o_2) = 0 \Leftrightarrow o_1 = o_2 \\ \delta(o_1, o_2) \leq \delta(o_1, o_3) + \delta(o_3, o_2) \end{cases} \quad (4.9)$$

Par distance adéquate à la comparaison de courbes de charge, nous désignons une mesure qui puisse à la fois traduire les ressemblances (ou les dissemblances) sur l'axe vertical (valeurs de puissance) comme sur l'axe horizontal (échelle du temps). Un grand nombre de distances sont référencées dans la littérature et il n'est pas question ici d'en dresser une liste exhaustive. Nous rappelons ci-dessous l'expression de quatre d'entre elles [AND02]. Celles-ci sont calculées pour deux vecteurs  $x$  et  $y$  de taille  $n$  (nous notons  $\theta$  l'indice des éléments de ces vecteurs) :

– distance de Minkowski :

$$\delta_{\text{Mink}}(x, y) = \sqrt[p]{\sum_{\theta=1}^n (|x_{\theta} - y_{\theta}|)^p} \quad p \in \mathbb{N} \quad (4.10)$$

– distance de Manhattan :

$$\delta_{\text{Manh}}(x, y) = \sum_{\theta=1}^n |x_{\theta} - y_{\theta}| \quad (4.11)$$

La distance de Manhattan est un cas particulier de la distance de Minkowski pour  $p = 1$ .

– distance euclidienne :

$$\delta_{\text{eucl}}(x, y) = \sqrt{\sum_{\theta=1}^n (x_{\theta} - y_{\theta})^2} \quad (4.12)$$

La distance euclidienne est un cas particulier de la distance de Minkowski pour  $p = 2$ .

– distance maximale :

$$\delta_{\max}(x, y) = \max_{\theta \in [1; n]} |x_{\theta} - y_{\theta}| \quad (4.13)$$

Les distances classiques comme celles évoquées ci-avant ne considèrent que les différences sur l'axe vertical si bien que deux courbes de charge identiques, qui présentent une forte dynamique, et seulement déphasées de quelques minutes seront considérées comme « éloignées » d'après la valeur prise par ces distances. **En définitive, nous cherchons un moyen d'apprécier simultanément les décalages temporels comme en amplitude entre différentes courbes d'appels de puissance.**

Conscients que nous partageons cette problématique avec d'autres domaines scientifiques, nous nous sommes intéressés aux techniques du traitement du signal. En particulier, la méthode du [Dynamic Time Warping \(DTW\)](#)\* a attiré notre attention.

### 4.1.3 Apports d'une méthode appliquée aux séries temporelles : le Dynamic Time Warping

#### 4.1.3.1 Historique

La méthode de Dynamic Time Warping, déformation temporelle dynamique, est une technique élaborée au début des années 1970 [SC71, SC78]. Initialement, elle a été utilisée pour la reconnaissance vocale (signature sonore de mots « cibles ») mais très vite de nombreuses autres applications en astronomie, biologie, économie, médecine... ont vu le jour [MÜL07]. Son but est de déceler les ressemblances entre deux séries temporelles et ce malgré des déformations éventuelles sur l'axe des abscisses. L'algorithme cherche ainsi à trouver le « meilleur » alignement possible entre deux séries temporelles, c'est-à-dire celui qui minimise le coût de déformation. Le champ d'application de cette méthode est très vaste et plusieurs raffinements ont été apportés à l'algorithme de base que nous explicitons au paragraphe 4.1.3.2. Ces améliorations sont décrites à la section 4.1.3.3. Le décalage et la déformation sur l'axe temporel sont bien au cœur de notre problème de courbe de charge.

#### 4.1.3.2 Algorithme de base

Soient deux séries temporelles  $\underline{Q}$  et  $\underline{V}$  de taille respective  $n$  et  $m$ . Elles sont définies de la manière suivante :

$$\underline{Q} \equiv \{q_1, q_2, \dots, q_n\} \quad \underline{V} \equiv \{v_1, v_2, \dots, v_m\}$$

La recherche de l'alignement optimal entre les séquences  $\underline{Q}$  et  $\underline{V}$  s'opère en construisant la matrice  $M$  de taille  $n \times m$  où l'élément  $(\theta, \psi)$  correspond à la distance  $\delta(q_{\theta}, v_{\psi})$  et donc à l'alignement entre les points  $q_{\theta}$  et  $v_{\psi}$  des séquences  $\underline{Q}$  et  $\underline{V}$  respectivement. Traditionnellement la distance euclidienne ( $\delta(q_{\theta}, v_{\psi}) = (q_{\theta} - v_{\psi})^2$ ) est utilisée mais toute autre métrique peut être employée (notamment celles évoquées dans la partie 4.1.2). La figure 4.6 illustre l'alignement optimal entre deux séries temporelles de type périodique obtenu par l'algorithme [Derivative Dynamic Time Warping \(DDTW\)](#)\* (voir le paragraphe correspondant dans la partie 4.1.3.3).

On appelle [chemin de déformation](#)\* (ou warping path) noté  $\underline{W}$ , une séquence contigüe, au sens des contraintes décrites ci-après, de  $K$  éléments de la matrice  $M$  ( $K$  est la longueur du chemin de déformation) qui définit un alignement des séquences  $\underline{Q}$  et  $\underline{V}$ .  $\underline{W}$  est formé par l'ensemble des points  $w_{\kappa} = (\theta, \psi)_{\kappa}$  ( $\kappa$  tel que  $1 \leq \kappa \leq K$  désigne l'indice des éléments du chemin de déformation  $\underline{W}$ ), comme l'explique l'équation 4.14.

$$\underline{W} \equiv \{w_1, w_2, \dots, w_{\kappa}, \dots, w_K\} \quad \text{avec} \quad \max(m, n) \leq K < m + n - 1 \quad (4.14)$$

Le chemin de déformation doit respecter les trois contraintes de base ci-dessous :

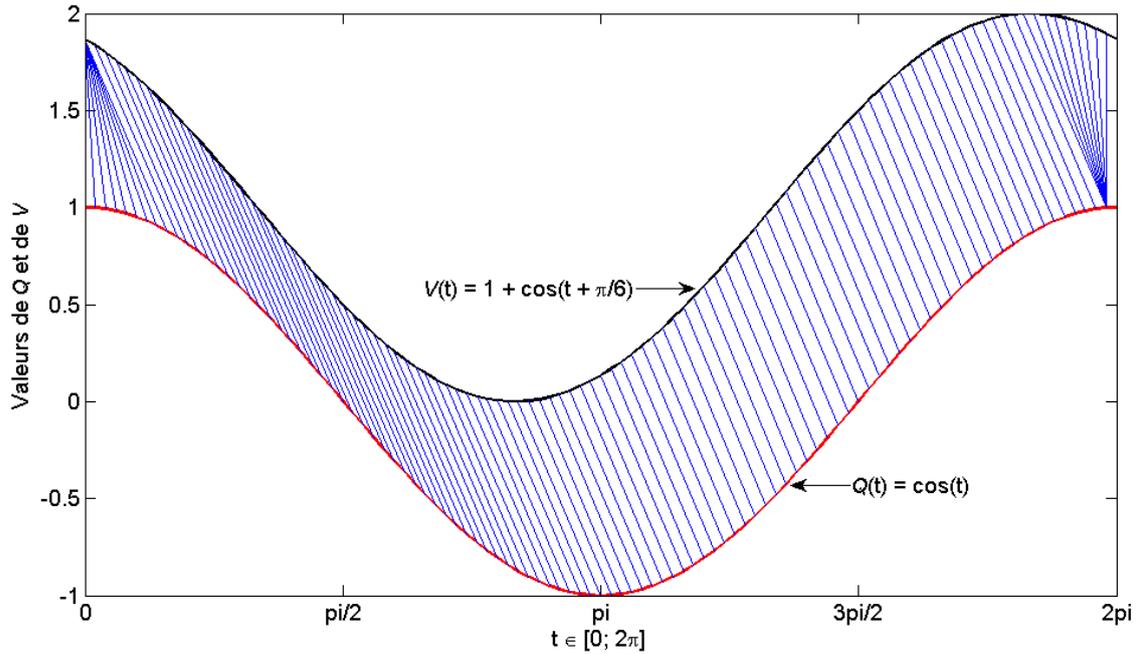


FIGURE 4.6 – Exemple d’alignement optimal entre deux séries temporelles  $Q$  et  $V$  grâce à l’algorithme  $DDTW$

**conditions limites :**  $w_1 = (1, 1)$  et  $w_K = (n, m)$ . De cette manière, le chemin débute à un coin de la matrice  $M$  et doit se terminer dans le coin opposé ;

**continuité :**  $w_\kappa = (\theta, \psi)$  et  $w_{\kappa-1} = (\theta', \psi')$ , les relations suivantes doivent être respectées :  $(\theta - \theta') \leq 1$  et  $(\psi - \psi') \leq 1$ . Concrètement la progression dans le chemin de déformation à partir d’un de ses éléments se fait exclusivement vers les cellules directement adjacentes de la matrice (les déplacements horizontaux, verticaux et en diagonale d’une seule case sont donc permis) ;

**monotonicité :**  $w_\kappa = (\theta, \psi)$  et  $w_{\kappa-1} = (\theta', \psi')$ , les relations suivantes doivent en outre être respectées :  $(\theta - \theta') \geq 0$  et  $(\psi - \psi') \geq 0$ . Cette contrainte force à ce que les éléments de  $W$  soient répartis de manière monotone sur l’axe temporel.

Déterminer le chemin de déformation optimal revient à un problème de minimisation sur l’ensemble des chemins de déformation potentiels  $W$  qui remplissent les trois contraintes précédentes. De façon générale, une fois choisie la métrique  $\delta w_\kappa = \delta(q_\theta, v_\psi)$  qui mesure la distance entre les points de deux séries temporelles, reste à solutionner l’équation 4.15 [BC94].

$$DTW(Q, V) = \min_W \left[ \sum_{\kappa=1}^K \delta w_\kappa \right] \quad (4.15)$$

Si la métrique choisie est la distance euclidienne, l’équation 4.15 devient la relation 4.16 [KP99].

$$DTW(Q, V) = \min_W \frac{\sqrt{\sum_{\kappa=1}^K w_\kappa}}{K} \quad (4.16)$$

Cette dernière équation tient compte de la longueur des chemins de déformation potentiels (dénominateur) dans un but de normalisation de la valeur  $DTW$ . Soulignons également que s’il n’y a pas de déphasage temporel entre deux séries comparées, la distance  $DTW$  entre deux courbes sera de même

valeur, au facteur  $1/K$  près, que celle de la métrique utilisée pour calculer les écarts entre les points des deux séquences dans l'algorithme *DTW* (dans le cas où cette mesure est estimée sur la totalité des longueurs des séquences).

Le *DTW* constitue une distance à part entière pour quantifier la similarité/dissimilarité entre séries temporelles. Elle peut en outre être employée comme métrique dans des algorithmes de classification [PKG11].

Un exemple de chemin de déformation est présenté à la figure 4.7.

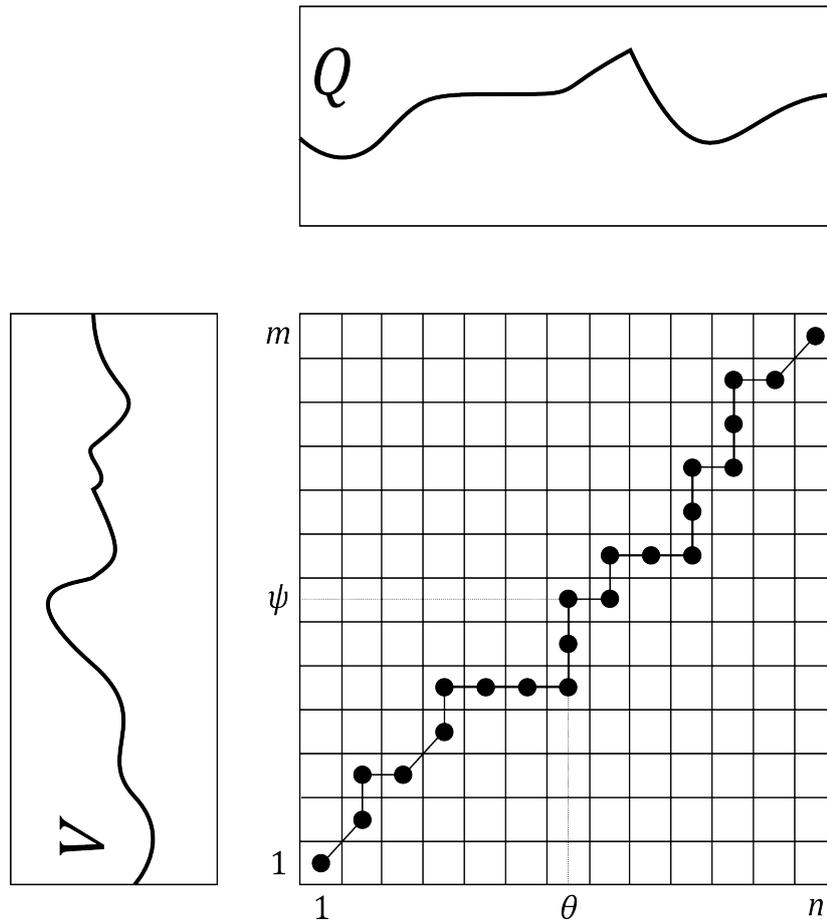


FIGURE 4.7 – Exemple de chemin de déformation entre deux séquences dont l'alignement est étudié [KP99]

Pour trouver le chemin de déformation qui minimise l'équation 4.15, une démarche récursive est adoptée. Celle-ci débute à partir de la fin du chemin en  $w_K = (n, m)$  et en cherchant à minimiser à chaque progression dans le chemin la distance cumulée  $\underline{\gamma}(\theta, \psi)$ , somme de la distance  $\delta(q_\theta, v_\psi)$  et du minimum de la distance cumulée estimée pour chacune des cellules adjacentes de la matrice  $M$ . Ce processus récursif est explicité par l'équation 4.17.

$$\gamma(\theta, \psi) = \delta(q_\theta, v_\psi) + \min\{\gamma(\theta - 1, \psi); \gamma(\theta, \psi - 1); \gamma(\theta - 1, \psi - 1)\} \quad (4.17)$$

Une étude approfondie sur le choix des contraintes de déplacement pour la construction de chemin de déformation a été menée dans [SC78]. Elle indique notamment que le déplacement « standard<sup>7</sup> » n'est pas le plus efficace pour le cas d'application en question.

7. Au sein de la matrice  $M$ , les seuls déplacements possibles à partir d'un élément donné sont ceux qui aboutissent aux cellules directement adjacentes tout en respectant les contraintes de conditions limites, de continuité et de monotonicité

### 4.1.3.3 Variantes de l'algorithme de base

Dans ce paragraphe, nous présentons brièvement deux améliorations qui ont été apportées au Dynamic Time Warping. Tout d'abord, dans le [Segmented Dynamic Time Warping \(SDTW\)\\*](#), les séquences temporelles comparées sont assimilées à une suite de segments de droite. D'autre part dans le Derivative Dynamic Time Warping, une métrique différente est choisie pour le calcul du chemin de déformation. Ces deux améliorations nous semblent potentiellement adaptées à notre sujet puisque d'une part les courbes de charge, et en particulier celles qui proviennent de mesures, sont constituées le plus souvent par une suite discrète de points « éloignés<sup>8</sup> » : une représentation des courbes de charge sous forme de ligne brisée (succession de segments de droite) s'impose alors tout naturellement, représentation qui est utilisée dans le Segmented Dynamic Time Warping. D'autre part, les améliorations apportées par le Derivative Dynamic Time Warping s'appliquent bien à la dynamique des courbes de charge. En effet, lorsque plusieurs d'entre elles sont comparées, il est à la fois important de quantifier les différences sur l'axe temporel et sur l'axe des puissances  $y$  compris lorsque les disparités sont faibles et/ou locales. L'algorithme *DDTW* pourrait par exemple typiquement s'appliquer à l'étude sur la courbe de charge de la pointe de consommation.

#### Segmented Dynamic Time Warping

Cette version de l'algorithme s'appuie sur une représentation nouvelle des séries temporelles [KP98, KP99]. Cette dernière a été motivée par le besoin d'améliorer la rapidité du processus de calcul de l'algorithme *DTW*. Concrètement, le *SDTW* impose tout d'abord de représenter les séries temporelles sous forme de séquences de segments de droite.

Soit  $\underline{S}$  une série temporelle échantillonnée sur  $t$  points. On appelle séquence notée  $\mathbf{S}$ , la version segmentée de  $S$  sur  $\mathbf{T}$  segments. Chacun d'eux est défini au moyen de deux points de coordonnées respectives  $\{SXL; SYL\}$  et  $\{SXR; SYR\}$ .  $\mathbf{S}$  peut donc être représentée comme un  $n$ -uplet de vecteurs de longueur  $\mathbf{T}$  avec  $n$  valant 4 ou 5 selon que sont attribués ou non des poids  $\underline{SW}$  à chacun des segments de la séquence. La séquence  $\mathbf{S}$  est représentée comme suit

$$\mathbf{S} \equiv \{SXL, SXR, SYL, SYR, (SW)\}$$

Le  $\theta^{\text{ème}}$  élément de la séquence  $\mathbf{S}$  est donc le segment qui relie les points  $(SXL_\theta, SYL_\theta)$  et  $(SXR_\theta, SYR_\theta)$  et qui est affecté ou non du poids  $SW_\theta$ . Une représentation schématique d'une séquence  $\mathbf{S}$  est présentée à la figure 4.8.

Le cas échéant, la gestion des poids des segments est opérée de manière dynamique. Tout d'abord, ceux-ci sont initialisés à 1. Ensuite, toute modification de l'un des poids des segments entraîne une renormalisation de l'ensemble du jeu de poids de sorte que la somme des produits des poids par la longueur de leur segment respectif soit égale à la longueur totale de la séquence. Concrètement, l'équation 4.18 doit être toujours vérifiée.

$$\sum_{\theta=1}^{\mathbf{T}} SW_\theta \times (SXR_\theta - SXL_\theta) = SXR_{\mathbf{T}} - SXL_1 \quad (4.18)$$

Ce procédé fait en sorte que l'on considère ainsi le poids relatif d'un segment par rapport aux autres et non pas son poids absolu. Augmenter le poids d'un segment va automatiquement réduire le poids de l'ensemble des autres segments de la séquence. Un autre avantage souligné par Keogh et Pazzani [KP98] est que de cette manière, le poids total associé à une séquence ne change pas, quel que soit le nombre de segments qui la représentent.

Une fois cette représentation des séquences établie et mise en pratique, la distance  $\delta$  entre deux séquences  $\mathbf{A}$  et  $\mathbf{B}$  se calcule au moyen de la relation 4.19 :

$$\delta(\mathbf{A}, \mathbf{B}) = \sum_{\theta=1}^{\mathbf{T}} AW_\theta \times BW_\theta \times \left| (AYL_\theta - BYL_\theta) - (AYR_\theta - BYR_\theta) \right| \quad (4.19)$$

8. En effet dans les campagnes de mesure, il est usuel de suivre des consommations d'énergie soit à pas de temps fixe soit à incrément – un enregistrement se fait tous les  $x$  Wh par exemple – de sorte que les courbes de charge reconstituées sont définies à une résolution faible (10 min dans le meilleur des cas en général)

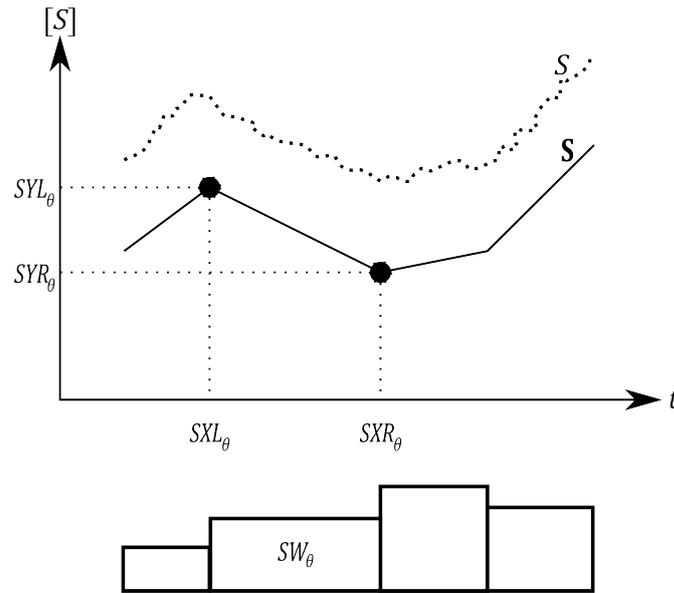


FIGURE 4.8 – Représentation d’une séquence  $\mathbf{S}$  à partir de la série temporelle  $S$

Cette mesure perçoit la manière avec laquelle les segments de droite sont proches d’un point de vue parallélisme tout en tenant compte de l’éventuelle pondération de chacun d’entre eux.

### Derivative Dynamic Time Warping

Dans cette approche, le changement essentiel par rapport à l’algorithme initial du *DTW* réside dans le choix de la métrique [KP01]. En lieu et place de la distance euclidienne est implémentée une distance  $\delta$  qui correspond au carré de la différence des dérivées estimées aux points  $q_\theta$  et  $v_\psi$  de deux séquences considérées  $Q$  et  $V$ . Ce choix est motivé par le fait que l’algorithme *DTW* est capable d’identifier des différences (accélération ou décélération) entre deux séries sur l’axe temporel. Lorsqu’en plus de ces écarts, des différences en matière de niveau des séries existent, l’algorithme a tendance à échouer en particulier si les divergences sont locales.

Dans le *DDTW*, l’algorithme cherche donc à mettre en phase les séquences en fonction de leurs variations à chaque pas de temps d’où l’emploi des dérivées de chacune d’entre elles. La dérivée de la série temporelle  $Q$  en un point  $\theta$  est calculée au moyen de la relation 4.20.

$$\frac{dQ}{dt}(t = \theta) = \frac{(q_\theta - q_{\theta-1}) + ((q_{\theta+1} - q_{\theta-1})/2)}{2} \quad 1 < \theta < n \quad (4.20)$$

De par la forme de l’équation 4.20, la dérivée ne peut pas être calculée aux extrémités de la série temporelle. La dérivée au premier point est fixée égale à celle calculée au point 2. De même, la dérivée au point  $n$  de  $Q$  est celle calculée au point  $n - 1$ .

La distance implémentée  $\delta(q_\theta, v_\psi)$  dans l’algorithme *DDTW* pour un point  $(\theta, \psi)$  de la matrice  $M$  est calculée conformément à la relation 4.21 :

$$\begin{aligned} \delta(q_\theta, v_\psi) &= \left( \frac{dQ}{dt}(t = \theta) - \frac{dV}{dt}(t = \psi) \right)^2 \quad \text{avec } 1 < \theta < n, 1 < \psi < m \\ &= \frac{(q_\theta - q_{\theta-1}) + ((q_{\theta+1} - q_{\theta-1})/2)}{2} - \frac{(v_\psi - v_{\psi-1}) + ((v_{\psi+1} - v_{\psi-1})/2)}{2} \end{aligned} \quad (4.21)$$

Cette partie propose un tour d’horizon des éléments de caractérisation d’une courbe de charge électrique. Abordés ici d’un point de vue théorique, nous mettons en œuvre certains d’entre eux (plus particulièrement l’algorithme Dynamic Time Warping) dans le paragraphe 4.3 pour l’analyse de courbes de charge

issues de notre modèle. Avant cela, nous détaillons dans la partie 4.2 quels ont été les scénarios modélisés par notre méthode de reconstitution des appels de puissance. En outre, nous proposons une étude succincte des courbes de charge obtenues par simulation.

## 4.2 Scénarios modélisés par notre méthode et étude descriptive des résultats en puissance

L'étude et l'analyse des résultats produits par notre modèle de courbe de charge ne peuvent être opérées non sans avoir au préalable défini et décrit les scénarios que nous avons modélisés : ce travail préliminaire est proposé au sein de la section 4.2.1. Puis, parmi les scénarios établis nous en avons choisi deux plus particulièrement pour visualiser les courbes de charge électriques reconstituées. Ainsi le paragraphe 4.2.2 traite des résultats de la simulation du *scénario principal*\*. La seconde étude, proposée quant à elle dans la section 4.2.3, est relative aux conclusions que nous avons tirées de la simulation d'une variante du précédent scénario.

### 4.2.1 Construction de scénarios en vue de leur simulation

Dans cette partie, nous commencerons par préciser dans le paragraphe 4.2.1.1 quels ont été les principaux objectifs que nous avons cherché à atteindre en matière de simulation. Nous déclinerons ces attentes en termes de scénarios à établir. Puis dans la section 4.2.1.2, nous aborderons plus en détails, le cas de référence que nous avons construit, appelé scénario principal, et dont nous avons opéré la simulation. Malgré sa simplicité, celui-ci nous a permis d'obtenir des résultats suffisamment fournis pour mener un certain nombre d'études et d'analyses sur les courbes de charge simulées. Enfin, dans la partie 4.2.1.3, nous examinerons les autres simulations que nous avons conduites. Plus particulièrement, il s'agit de deux variantes du scénario principal.

#### 4.2.1.1 Objectifs liés à la simulation de scénarios

Nous avons détaillé, au sein du chapitre 2, la méthode bottom-up et agrégative que nous avons construite pour parvenir au résultat final qu'est la courbe de charge électrique du parc résidentiel national. Toutefois, cet objectif ne peut être atteint qu'en ayant effectué au préalable un certain nombre de calculs intermédiaires et obtenu les résultats sur des *segments*\* du parc résidentiel national.

Le chapitre 2 a également souligné l'importance des processus itératifs au sein de notre méthode. Elle y fait appel à différents niveaux notamment de sorte de pouvoir répéter par simulation le même scénario un certain nombre de fois. Typiquement ces procédés répétitifs servent entre autres à la reconstruction explicite du foisonnement d'origine comportementale. Ainsi la modélisation d'un segment du parc contenant  $n$  n-uplets identiques revient à la paramétrisation d'un seul n-uplet qui sera simulé  $n$  fois.

Le travail de modélisation mené au sein de la thèse n'a fait qu'initialiser une tâche de plus grande ampleur. En effet, la méthode de reconstitution de courbe de charge – concrètement l'algorithme de déclenchement des équipements domestiques – n'a pu être testée que sur un segment résidentiel très simplifié. Plus précisément, nous nous sommes concentrés dans un premier temps sur les appels de puissance de quelques équipements d'électricité spécifique. Ce travail initial visait à tester le bon fonctionnement de l'algorithme que nous avons implémenté. Il avait également pour vocation de permettre l'observation et la quantification du foisonnement d'origine comportementale. Différentes études et analyses ont été conduites en ce sens et sont présentées au sein des paragraphes 4.2 et 4.3.

Ensuite, nous avons cherché à tester des variantes du cas de référence (scénario principal). La première variante nous a permis d'étudier l'impact en puissance de la modification du cycle unitaire d'un appareil. Cette variante présente la particularité d'aboutir à la même consommation d'énergie que le scénario d'origine. La seconde variante a été l'occasion d'étoffer le scénario principal modélisé en y ajoutant des équipements supplémentaires.

#### 4.2.1.2 Scénario principal de modélisation d'un segment du parc résidentiel

La définition d'un segment du parc résidentiel revient à choisir et à construire les  $n$ -uplets auxquels on s'intéresse pour la modélisation en puissance. Dans notre cas et comme nous l'avons précisé au préalable, l'aspect bâtiment n'est pas considéré dans nos travaux de modélisation. En revanche, il subsiste les dimensions ayant trait aux ménages et bien entendu aux équipements domestiques.

Dans une première approche de modélisation, nous n'avons retenu que les éléments essentiels de caractérisation des ménages et des appareils électriques. Le lecteur remarquera donc qu'ils ne sont pas aussi nombreux et variés que ceux détaillés dans la méthode conceptuelle (parties 2.1.1.2 et 2.1.1.3 du chapitre 2). Cette première approche de modélisation a abouti à l'élaboration du scénario principal, scénario que nous détaillons dans ce paragraphe.

Du fait des spécificités de l'environnement informatique dans lequel nous avons implémenté l'algorithme, ce dernier est fractionné en différents étapes qui s'enchaînent de façon chronologique au cours d'une simulation. Cette succession est établie au sein d'un script qui prend comme paramètres le nombre d'itérations souhaitées de simulation du même scénario et la durée de la simulation. Sur ce dernier point, nous avons systématiquement choisi pour nos recherches une durée de simulation différente de 24 h<sup>9</sup>. En effet, malgré le rythme calendaire qui conditionne sensiblement le comportement journalier des individus, celui-ci n'épouse pas strictement cette régularité temporelle. Ainsi par exemple, comme nous l'avons fait observer au sein du paragraphe 3.1.1, l'heure de couchage des individus d'une population hétérogène lors d'un jour de semaine suit une distribution qui s'étale sur deux jours calendaires consécutifs (couchage avant et après minuit). En choisissant une telle valeur de durée de simulation, il est certain avec la paramétrisation choisie, que la borne inférieure de l'intervalle Nuit,  $t_{End}$ , appartienne à la simulation.

Nous proposons au sein de l'annexe C une description des étapes de la simulation d'un scénario effectué à l'aide de notre modèle disponible dans l'environnement DYMOLA.

En ce qui concerne le scénario principal, deux équipements sont considérés : le lave-linge et le sèche-linge. Nous ne nous intéressons donc qu'aux seules macro-tâche et micro-tâche LAVage. Au niveau de la macro-tâche, nous devons renseigner la compatibilité de celle-ci vis-à-vis de chaque intervalle infra-journalier. Nous avons choisi d'autoriser la conduite du LAVage dans les intervalles Réveil, Matin, Après-midi et Soir alors que nous l'interdisons dans les deux autres intervalles. Cette donnée de macro-compatibilité est ensuite croisée avec le résumé du jour-type. Ceci permet de construire la donnée de micro-compatibilité pour l'activité LAVage du ménage-type considéré dans ce scénario.

À ce stade, il faut paramétrer les équipements simulés. Soulignons ici que les cycles unitaires de puissance utilisés dans les simulations, ont été représentés par des constructions assez simples basées sur des couples {paliers de puissance ; durée de phases} à l'image de ce que propose Stamminger [STA09]. Ces représentations se voulaient davantage vraisemblables qu'exactes<sup>10</sup> et elles présentaient en outre le mérite de la simplicité.

Les figures 4.9 et 4.10 illustrent la modélisation des appels de puissance que nous avons adoptée pour la machine à laver et le sèche-linge. Le tableau 4.1 synthétise quant à lui les valeurs des paramètres que nous avons choisies et qui ont un impact sur le calcul des appels de puissance des appareils et/ou sur la sélection de l'instant de déclenchement de leur usage<sup>11</sup>. Le tableau 4.2 synthétise les valeurs<sup>12</sup> utilisées dans la modélisation des cycles unitaires de puissance des deux équipements simulés.

9. Plus précisément, nous avons opté pour 100 000 s soit un peu moins de 28 h

10. À ce moment de la thèse, nous ne disposions pas encore d'une connaissance empirique approfondie des cycles unitaires de puissance de ces appareils en fonction de leur mode de fonctionnement

11. Dans le tableau 4.1, nous ne listons pas l'ensemble des paramètres du modèle d'équipement à cycle de fonctionnement. En effet, l'influence effective de chacun d'eux n'a pas pu être implémentée dans cette version de la modélisation

12. Rappelons une fois encore que cette représentation a été choisie pour sa simplicité

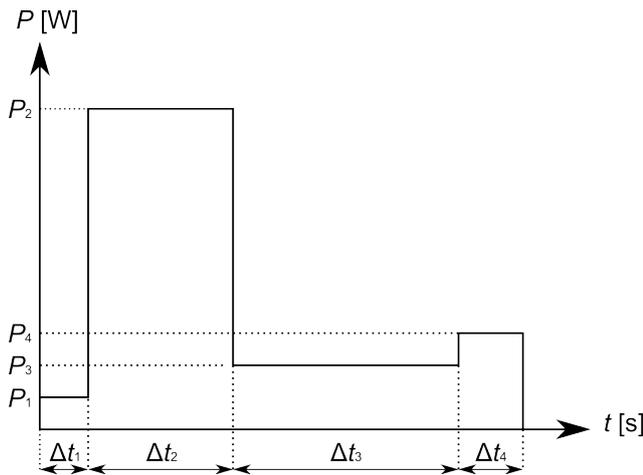


FIGURE 4.9 – Modélisation du cycle unitaire de puissance de la machine à laver pour le scénario principal

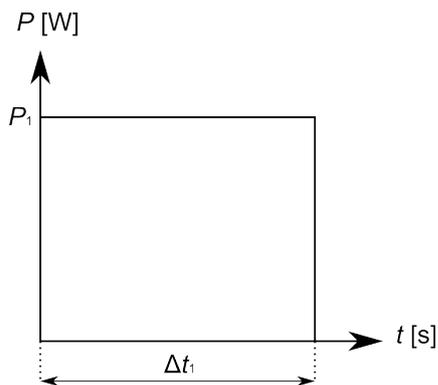


FIGURE 4.10 – Modélisation du cycle unitaire de puissance de sèche-linge pour le scénario principal

À ce niveau, deux précisions importantes doivent être mentionnées. Tout d'abord en ce qui concerne la modélisation du cycle unitaire de puissance de la machine à laver qui comporte quatre phases, une erreur dans notre implémentation informatique repérée tardivement, a conduit à la suppression de la phase d'essorage (couple  $\{P_4; \Delta P_4\}$ ) dans les résultats de la simulation du scénario principal.

D'autre part, la restitution des courbes de charge simulées est réalisée suivant un pas de temps variable qui est sélectionné de manière dynamique par le logiciel de simulation. Il en résulte que les transitions instantanées de puissance présentes dans la modélisation des cycles unitaires ne sont « plus que » des variations de même amplitude sur une durée égale au pas de temps de restitution du logiciel des courbes de charge unitaires (200 s dans notre cas)<sup>13</sup>.

En définitive, les cycles unitaires de puissance de machine à laver et de sèche-linge que nous retrouvons dans nos courbes de charge journalières simulées sont ceux qui sont présentés par la figure 4.11.

Paramètre [unité]	Valeurs des paramètres	
	Machine à laver	Sèche-linge
$P_{nom}$ [W]	1200	1000
$P_{veille}$ [W]	5	2
Mode de fonctionnement	Cycle à 40 °C	Cycle coton
$d_1, Durée$ [min]	2	4
$d_1, Disp$ [-]	1	1
$d_2, Durée$ [min]	0	0
$d_2, Disp$ [-]	0	0
$d_3, Durée$ [min]	65	40
$d_3, Disp$ [-]	0	0
$d_4, Durée$ [min]	5	17
$d_4, Disp$ [-]	0	0
$d_5, Durée$ [min]	8	3
$d_5, Disp$ [-]	1	1
$AD_{Réveil}$ [-]	0	0
$AD_{Matin}$ [-]	0	0
$AD_{Midi}$ [-]	0	0
$AD_{Après-midi}$ [-]	0	0
$AD_{Soir}$ [-]	1	1
$AD_{Nuit}$ [-]	0	0
Équipement auto-lié	Oui	Oui
Équipement lié	Sèche-linge	Aucun
Temporalité de la liaison	Séquentiellement après	Sans objet

TABLEAU 4.1 – Valeurs pour le scénario principal des paramètres à renseigner au niveau des modèles d'équipements domestiques à cycle de fonctionnement

Paramètre [unité]	Valeurs des couples {paliers de puissance; durée de phases}	
	Machine à laver	Sèche-linge
$P_1$ [W]	$0,1 \times P_{nom}$	$0,5 \times P_{nom}$
$P_2$ [W]	$1 \times P_{nom}$	-
$P_3$ [W]	$0,2 \times P_{nom}$	-
$P_4$ [W]	$0,3 \times P_{nom}$	-
$\Delta t_1$ [min]	$0,05 \times d_3, Durée$	$1 \times d_3, Durée$
$\Delta t_2$ [min]	$0,15 \times d_3, Durée$	-
$\Delta t_3$ [min]	$0,65 \times d_3, Durée$	-
$\Delta t_4$ [min]	$0,15 \times d_3, Durée$	-

TABLEAU 4.2 – Éléments de modélisation des cycles unitaires de puissance utilisés dans le scénario principal

13. Bien évidemment, il est possible de reconstituer les cycles unitaires originaux à l'aide d'un post-traitement. Cependant, puisqu'un pas de temps de restitution des courbes de charge d'électricité spécifique égal à 10 min nous semble adéquat, cette perte d'information n'est pas cruciale. Nous avons donc travaillé sur les courbes de charge unitaires restituées de la sorte (pas de temps de sortie de simulation égal à 200 s)

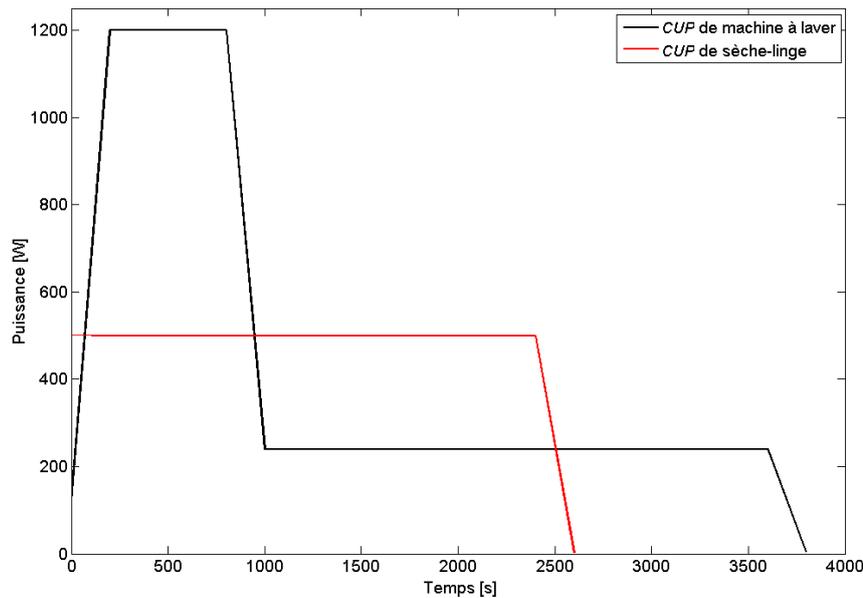


FIGURE 4.11 – *CUP* de machine à laver et de sèche-linge issus des simulations du scénario principal

Dans le reste du manuscrit, chaque fois que nous ferons référence aux simulations du scénario principal pour la machine à laver et le sèche-linge et sans que nous fassions d'autres précisions, il s'agira de courbes de charge avec ces cycles unitaires de puissance.

Enfin, en plus des résultats en puissance qui sont fournis par le modèle, nous avons choisi de conserver la temporalité d'utilisation des appareils domestiques, c'est-à-dire leur instant de déclenchement respectif. Ceci permet par exemple de tester l'impact en puissance du changement de technologie et/ou de cycle unitaire de puissance (passage du *CUP* initial à un autre type de cycle). Le paragraphe 4.2.1.3 décrit un exemple d'un tel scénario.

#### 4.2.1.3 Autres scénarios modélisés et simulés

Dans le paragraphe précédent, nous avons abordé la construction du scénario principal modélisé mettant en jeu deux équipements domestiques. Dans cette partie, nous détaillons les autres scénarios que nous avons considérés dans nos travaux de thèse. Ceux-ci sont construits de façon très similaire au scénario principal.

##### Simulation d'une variante du scénario principal : modification du cycle unitaire de puissance à iso-consommation

Il est naturellement possible de conduire la simulation des appareils domestiques en modifiant n'importe lequel des paramètres que nous avons mentionnés dans la partie 4.2.1.2. Comme nous l'avons souligné au préalable, nous avons construit le scénario principal de sorte de pouvoir récupérer l'ensemble des instants de déclenchement des équipements modélisés. De cette manière, il est possible de rejouer le même scénario en ne modifiant que l'allure du (des) cycle(s) unitaire(s) de puissance, la temporalité des appels étant strictement la même.

Nous avons justement voulu savoir quelle pouvait être la modification en puissance sur la courbe de charge moyenne foisonnée du passage d'un cycle à un autre en restant à iso-consommation énergétique.

Pour ce faire et à partir du cycle unitaire de puissance original de machine à laver, nous avons construit un autre *CUP*, engendrant la même consommation énergétique, et dont la valeur maximale de puissance

est deux fois moins importante que celle du cycle original tout en étant appelée deux fois plus longtemps.

La figure 4.12 illustre d'une part le fait que l'on considère ici le  $CUP_1$ , cycle unitaire de puissance original du scénario principal, sous une « version instantanée » notée  $^{inst}CUP_1$ <sup>14</sup> (nous détaillons ci-après comment nous aboutissons à cette représentation). D'autre part, cette même figure témoigne de la substitution du  $^{inst}CUP_1$  par le  $^{inst}CUP_2$  engendrant la même consommation d'énergie et tel que :

$$\begin{cases} P_{max,CUP_1} = P_{max,^{inst}CUP_1} \\ P_{max,^{inst}CUP_2} = \frac{1}{2} \times P_{max,^{inst}CUP_1} \\ \Delta t_{P_{max,^{inst}CUP_2}} = 2 \times \Delta t_{P_{max,^{inst}CUP_1}} \end{cases} \quad (4.22)$$

Dans l'équation 4.22,  $P_{max,CUP_1}$  (respectivement  $P_{max,^{inst}CUP_1}$  et  $P_{max,^{inst}CUP_2}$ ) désigne la puissance maximale du  $CUP_1$  (respectivement du  $^{inst}CUP_1$  et du  $^{inst}CUP_2$ ) et  $\Delta t_{P_{max,^{inst}CUP_1}}$  (respectivement  $\Delta t_{P_{max,^{inst}CUP_2}}$ ) la durée pendant laquelle est appelée la puissance maximale du cycle  $^{inst}CUP_1$  (respectivement  $^{inst}CUP_2$ ).

Faisons également remarquer que pour assurer la stricte iso-consommation entre les deux scénarios à l'échelle journalière, il est indispensable de mettre à zéro le niveau non nul de puissance de veille choisi dans le scénario principal (il est en effet de 5 W pour la machine à laver, cf. tableau 4.1).

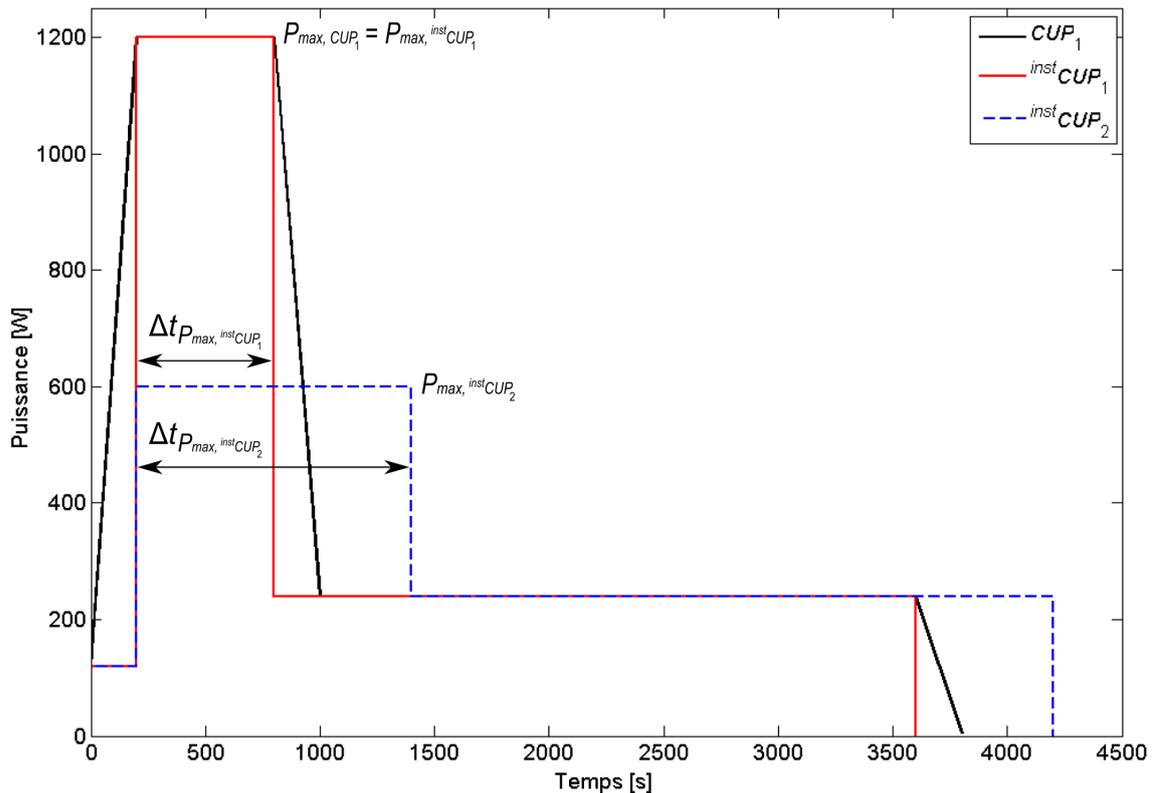


FIGURE 4.12 – Modification du cycle unitaire de puissance de machine à laver à iso-consommation : assimilation du  $CUP_1$  au  $^{inst}CUP_1$  et passage du  $^{inst}CUP_1$  au  $^{inst}CUP_2$

Construire la version instantanée d'un cycle unitaire de puissance est une opération très simple : pour cela, il suffit d'interpoler un cycle unitaire original au moyen d'un pas de temps très fin (par exemple

14. Se lit version instantanée du cycle unitaire de puissance 1

1 s). En outre, nous avons transformé les transitoires de puissance (d'un niveau  $P_1$  à un niveau  $P_2$ ) en variations immédiates. De cette manière, nous aboutissons à de véritables créneaux pour lesquels il est bien plus simple d'appliquer un ratio<sup>15</sup> donné de consommation entre différents cycles modélisés.

Nous étudierons en détails les résultats de ce test au sein du paragraphe 4.2.3.

**Modélisation d'un scénario étoffé**

Le scénario principal ne comportant que deux équipements domestiques d'électricité spécifique, nous avons cherché à étoffer les possibilités de simulation en y ajoutant de nouveaux équipements. Ainsi, nous avons considéré, en plus du lave-linge et du sèche-linge, le lave-vaisselle et la télévision.

Au delà de la macro et de la micro-tâche LAVage qui sont impliquées dans ce scénario au travers de la machine à laver et du sèche-linge, le scénario étoffé met en jeu les macro et micro-tâches Préparation Culinaire et Repas par l'intermédiaire du lave-vaisselle ainsi que Loisirs, Repos, Communications et Mobilité qui est la tâche parente de la télévision.

Dans ce scénario, la paramétrisation de la machine à laver et du sèche-linge et la modélisation de leur cycle unitaire de puissance sont comparables<sup>16</sup> aux choix que nous avons faits pour le scénario principal. En ce qui concerne la télévision et le lave-vaisselle, les figures 4.13 et 4.14 ainsi que les tableaux 4.3 et 4.4 synthétisent les éléments de leur modélisation respective.

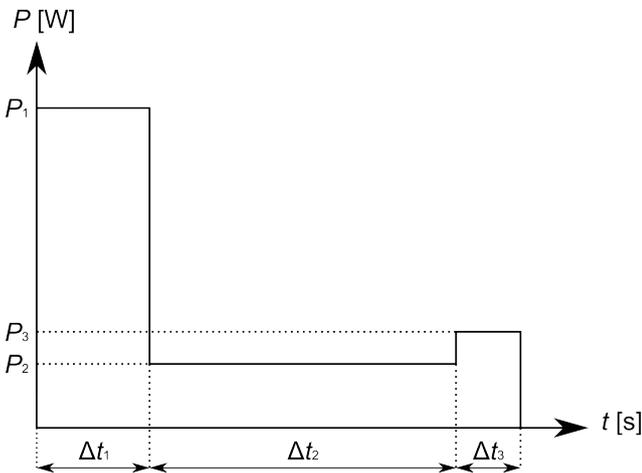


FIGURE 4.13 – Modélisation du cycle unitaire de puissance de lave-vaisselle pour le scénario étoffé

Paramètre [unité]	Valeurs des paramètres	
	Lave-vaisselle	Télévision
$P_{nom}$ [W]	2000	200
$P_{veille}$ [W]	4	6
Mode de fonctionnement	Cycle à 55 °C	Marche (on)
$d_1, Durée$ [min]	10	1
$d_1, Disp$ [-]	1	1
$d_2, Durée$ [min]	0	0
$d_2, Disp$ [-]	0	0
$d_3, Durée$ [min]	75	45
$d_3, Disp$ [-]	0	0
$d_4, Durée$ [min]	60	0
$d_4, Disp$ [-]	0	0
$d_5, Durée$ [min]	5	0
$d_5, Disp$ [-]	1	0
$AD_{Réveil}$ [-]	1	0
$AD_{Matin}$ [-]	0	0
$AD_{Midi}$ [-]	1	0
$AD_{Après-midi}$ [-]	0	0
$AD_{Soir}$ [-]	1	0
$AD_{Nuit}$ [-]	0	0
Équipement auto-lié	Oui	Non
Équipement lié	Aucun	Aucun <sup>17</sup>
Temporalité de la liaison	Sans objet	Sans objet

TABLEAU 4.3 – Valeurs pour le scénario étoffé des paramètres à renseigner au niveau des modèles d'équipements domestiques à cycle de fonctionnement et à états

15. Dans le cas de deux cycles engendrant la même consommation, ce rapport est naturellement égal à 1. Il serait égal à 0,5 si la consommation d'un cycle est égale à la moitié de celle de l'autre...

16. Nous ne jugeons pas utile de préciser les quelques différences existantes entre les deux scénarios en ce qui concerne la paramétrisation de la machine à laver et du sèche-linge

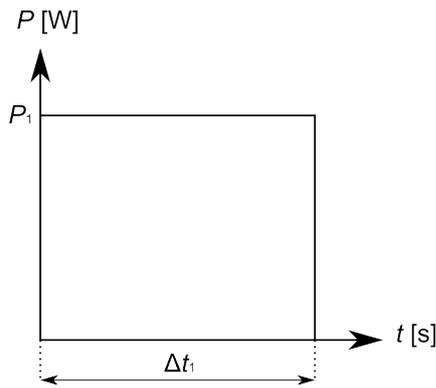


FIGURE 4.14 – Modélisation du cycle unitaire de puissance de télévision pour le scénario étoffé

Paramètre [unité]	Valeurs des couples {paliers de puissance ; durée de phases}	
	Lave-vaisselle	Télévision
$P_1$ [W]	$1 \times P_{nom}$	$1 \times P_{nom}$
$P_2$ [W]	$0,2 \times P_{nom}$	-
$P_3$ [W]	$0,3 \times P_{nom}$	-
$\Delta t_1$ [min]	$0,2 \times d_3, Durée$	$1 \times d_3, Durée$
$\Delta t_2$ [min]	$0,65 \times d_3, Durée$	-
$\Delta t_3$ [min]	$0,15 \times d_3, Durée$	-

TABLEAU 4.4 – Éléments de modélisation des cycles unitaires de puissance du lave-vaisselle et de la télévision utilisés dans le scénario étoffé

N'ayant conduit que trop peu de simulations de ce scénario étoffé, nous n'étudierons pas les résultats obtenus. En effet, au vu du nombre de courbes de charge individuelles disponibles (quelques-unes seulement) pour ces nouveaux équipements, la connaissance qu'il serait possible d'en extraire serait véritablement limitée, en particulier en matière d'évaluation du foisonnement.

Pour clôturer la revue des différents scénarios modélisables dans l'état final de nos travaux, remarquons ici que nous pourrions imaginer des cas d'études plus complexes, mettant en jeu à la fois des équipements d'électricité spécifique (tels que ceux que nous avons déjà considérés mais aussi d'autres comme les équipements de froid domestique par exemple), des équipements thermiques<sup>18</sup> (au moyen d'un couplage<sup>19</sup> avec un modèle de bâtiment équipé de ses systèmes énergétiques) tout en considérant les membres des ménages de manière individuelle (contrairement aux scénarios détaillés jusqu'alors dans lesquels, les ménages simulés sont traités de manière agrégée en tant qu'entités indivisibles).

#### 4.2.2 Étude des résultats du scénario principal

Le scénario principal, que nous avons abordé en détails dans le paragraphe 4.2.1.2, met en jeu rappelons-le deux équipements domestiques : la machine à laver et le sèche-linge. Ceux-ci sont déclenchés dans la journée simulée au sein de ménages de type *Inactif*. Au total, nous avons regroupé les résultats de 6500 simulations issues du scénario principal et du *scénario principal assimilé*<sup>20</sup>. Dans la suite, nous nommerons *scénario principal au sens large*\* le regroupement des résultats issus des deux scénarios évoqués dans la note 20. Pour des raisons de commodité d'écriture, nous le désignerons en tant que scénario principal *ASL*\*. Lorsque cela s'avérera nécessaire, nous ferons référence au scénario principal, au sens strict du terme, en le nommant « scénario principal *stricto sensu* ».

17. À l'heure de tester ce scénario étoffé, nous n'avons pas implémenté d'équipements pouvant être liés au téléviseur d'où la valeur de cette modalité

18. La modélisation de tels équipements est disponible dans notre environnement de simulation

19. Au cours des travaux de thèse, un test de couplage entre le bâti et notre modèle d'électricité spécifique a été opéré avec succès

20. Comme nous pourrions le constater au moment de mettre en œuvre l'algorithme du Dynamic Time Warping sur les courbes de charge moyennes foisonnées correspondantes aux résultats de la modélisation (paragraphe 4.3 p. 185), les simulations ont été opérées sur le scénario principal *stricto sensu*, c'est-à-dire tel que détaillé dans la partie 4.2.1.2 ainsi que sur un scénario très proche de ce dernier appelé scénario principal assimilé. En effet, la *seule* différence réside au niveau de la valeur du paramètre `ProbabiliteDeclenchement` – cf. page 130 – pour le sèche-linge. Dans le scénario principal *stricto sensu*, ce paramètre vaut 1 pour les deux appareils considérés (autrement dit à chaque simulation ils sont utilisés au sein du ménage considéré) alors que dans le scénario principal assimilé, `ProbabiliteDeclenchement` est inférieur à 1 pour le sèche-linge si bien que dans certaines simulations, cet équipement n'a pas été déclenché

Les figures 4.15 et 4.16 proposent quelques courbes de charge unitaires issues des simulations du scénario principal ASL, pour la machine à laver d'une part et le sèche-linge d'autre part. Sur ces graphiques, nous notons  $CDCU_n$ , la courbe de charge unitaire issue de la  $n^{\text{ème}}$  simulation du scénario principal ASL.

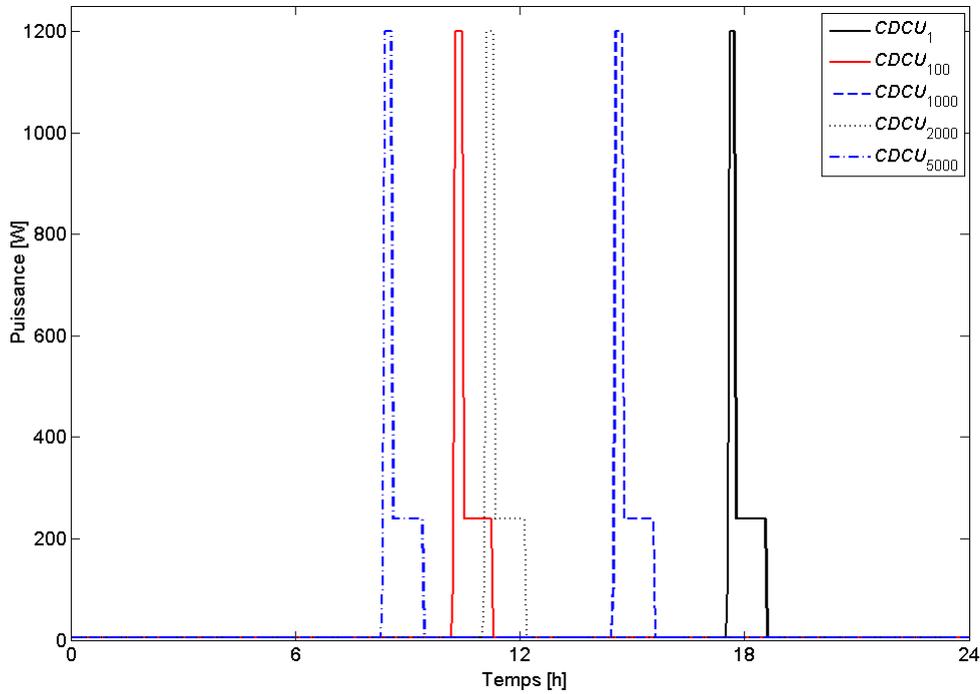


FIGURE 4.15 – Exemples de quelques courbes de charge unitaires de machine à laver issues des simulations du scénario principal ASL

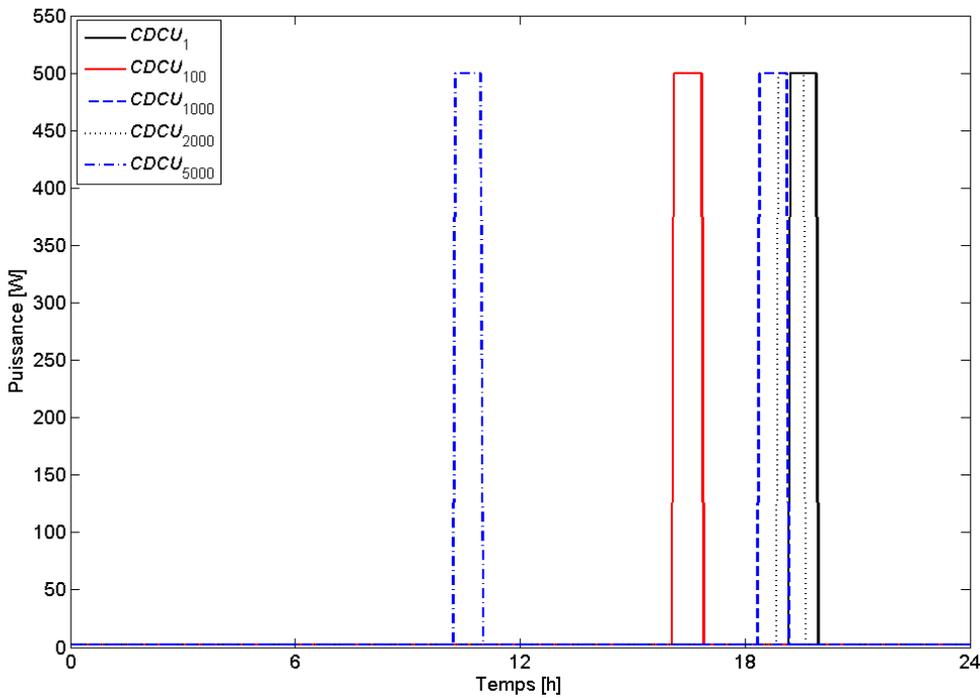


FIGURE 4.16 – Exemples de quelques courbes de charge unitaires de sèche-linge issues des simulations du scénario principal ASL

Comme le suggèrent les illustrations 4.15 et 4.16, l'étude des résultats de simulation n'a d'intérêt que si l'on travaille à une échelle agrégée<sup>21</sup> (cette remarque vaut de manière générale et donc a fortiori pour le scénario principal ASL considéré ici). Autrement dit, il nous faut impérativement construire des représentations en puissance correspondantes à un ensemble de simulations.

Gardant à l'esprit l'objectif de visualiser et d'analyser le foisonnement, nous avons construit des courbes de charge moyennes foisonnées (*CDCMF*) au pas 10 min. Chacune d'elles est calculée en sommant un certain effectif de courbes de charge unitaires (*CDCU*) et en divisant le résultat de cette somme par l'effectif. Ce dernier correspond au niveau d'agrégation de la courbe de charge moyenne foisonnée. Nous proposerons au paragraphe 4.3 différentes méthodes pour construire ces courbes de charge moyennes foisonnées. À ce stade, il n'est pas indispensable de préciser quelle méthode a été employée ici.

Pour étudier les résultats du scénario principal ASL, nous avons décidé de construire des *CDCMF* à partir d'effectifs successivement croissants de *CDCU* et à intervalles réguliers. Nous avons choisi 100 comme effectif de base si bien que nous avons établi 65 *CDCMF* de machine à laver et autant pour le sèche-linge à partir des résultats de simulation du scénario principal ASL. Nous notons  $CDCMF_{100}$  la courbe de charge moyenne foisonnée calculée à partir de 100 courbes de charge unitaires,  $CDCMF_{200}$  représente la *CDCMF* construite à partir de 200 *CDCU* etc.

Les figures 4.17 et 4.18 proposent respectivement quelques courbes de charge moyennes foisonnées de machine à laver et de sèche-linge pour le scénario principal ASL et à différents niveaux d'agrégation.

Nous décrivons ci-après les principaux enseignements de la lecture graphique des courbes de charge moyennes foisonnées de lave-linge et de sèche-linge construites à partir de simulations unitaires. Les commentaires que nous allons en faire se veulent expliciter au mieux nos observations sur les 5 *CDCMF* exposées pour chaque appareil. Ces remarques sont donc davantage générales et transverses qu'exactes au cas par cas.

Si l'on s'intéresse tout d'abord aux *CDCMF* de la machine à laver, nous pouvons premièrement remarquer que chacune d'elles se décompose en deux parties durant lesquelles les appels de puissance sont notables : la matinée d'un côté et l'après-midi et la soirée de l'autre (en matière de durée, la seconde partie de la courbe de charge est environ 1,5 fois plus longue que la première). Cette séparation nette entre deux périodes où se concentre l'utilisation du lave-linge s'explique tout simplement par la paramétrisation du scénario principal ASL. Plus particulièrement, ce qui s'observe facilement ici est l'influence de la valeur de macro-compatibilité du LAVage pour l'intervalle Midi<sup>22</sup>.

Partant du niveau de veille (5 W) paramétré dans le scénario considéré, la puissance moyenne foisonnée de la machine à laver croît à partir de 5 h du matin pour atteindre son maximum global (entre 70 et 80 W) aux alentours de 11 h. Ce maximum est également appelé pointe concomitante journalière de l'équipement. Après cela, la puissance appelée revient à son niveau initial (5 W ou à une puissance très légèrement supérieure) au moment de la période méridienne. Nous pouvons ici souligner l'impact du niveau d'agrégation de la courbe de charge moyenne foisonnée sur la pointe concomitante de l'équipement. **Pratiquement, l'appel maximal de puissance décroît avec l'augmentation du niveau d'agrégation de la courbe de charge moyenne foisonnée.**

En début d'après-midi, les appels de puissance repartent à la hausse pour atteindre, peu avant 16 h, un plateau situé entre 35 et 55 W. Ce plateau prend fin autour de 19-20 h. À partir de là, la puissance agrégée moyenne est décroissante jusqu'au terme de la journée<sup>23</sup> où le niveau appelé équivaut alors à

21. À noter toutefois qu'au niveau individuel, une analyse pertinente demeure : il s'agirait, dans chacun des cas, de s'assurer de la cohérence de l'utilisation des appareils domestiques et de la temporalité des tâches ménagères via-à-vis du scénario de présence et de disponibilité des occupants du ménage considéré. Une telle étude, aussi exhaustive soit elle, n'apporte cependant rien de nouveau en matière de connaissance du foisonnement. Celui-ci rappelons-le, est d'un intérêt majeur dans le cadre de nos recherches

22. Nous avons vu page 168 que dans le scénario principal, nous n'autorisons pas la conduite de la macro-tâche LAVage durant l'intervalle méridien

23. Bien que nous ayons simulé sur une période plus longue que 24 h, nous avons choisi de restituer les résultats de notre modèle à l'échelle journalière (calendaire) pour des raisons de commodité de lecture

celui du début de la simulation.

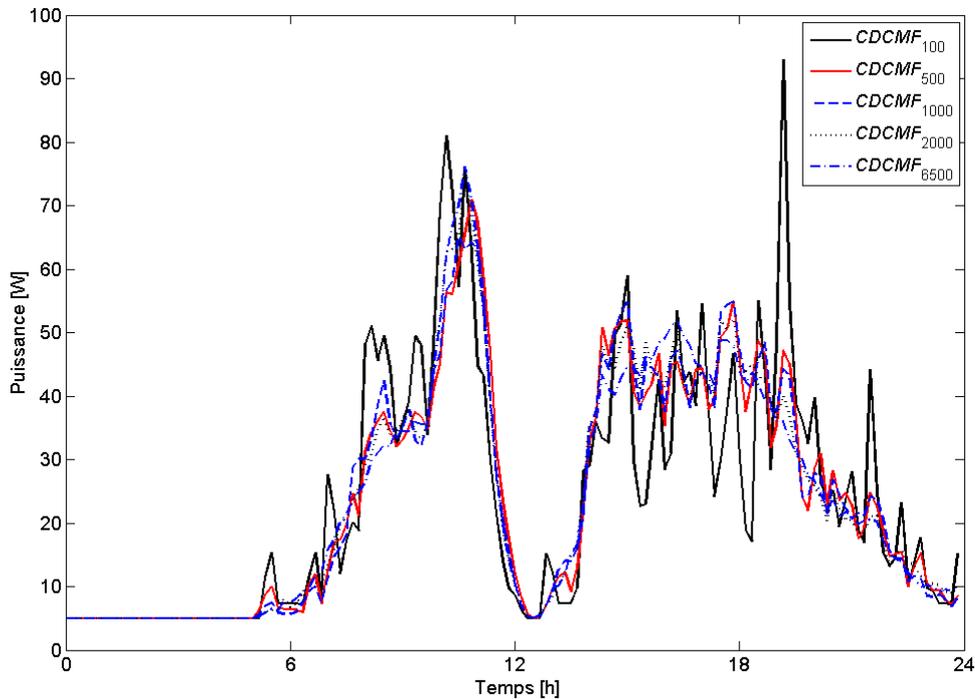


FIGURE 4.17 – Illustration de quelques courbes de charge moyennes foisonnées de machine à laver calculées à partir de courbes de charge unitaires simulées pour le scénario principal ASL

Les courbes de charge moyennes foisonnées du sèche-linge sont également formées de deux parties durant lesquelles une puissance notable (i. e. supérieure au niveau de veille) est requise par l'appareil. Plus précisément, il s'agit d'abord d'une courte période juste avant midi (environ 2 h) et d'une plus longue ( $\approx 10$  h) après la pause déjeuner et jusqu'à la fin de la journée. Là encore se matérialise l'influence de la macro-compatibilité du LAVage.

Initialisées par une période durant laquelle la puissance appelée est égale au niveau de veille des appareils simulés (2 W), les courbes de charge moyennes foisonnées du sèche-linge présentent ensuite une augmentation très progressive de la puissance appelée qui débute entre 7 h et 7 h 30. Cette évolution croissante aboutit à un maximum local de l'ordre de 10 W aux alentours de 11 h. Sitôt ce pic passé, la puissance appelée ensuite décroît pour revenir à son niveau initial à la mi-journée. En début d'après-midi, l'augmentation de la puissance requise est tout aussi progressive que dans la matinée. Néanmoins, elle se poursuit plus longuement puisqu'un maximum local (environ 50-55 W) est atteint aux environs de 19-20 h. Après cela, la puissance appelée baisse pour parvenir vers 21 h à un minimum local (plus ou moins 40-45 W). Les appels de puissance repartent une dernière fois à la hausse pour atteindre une heure plus tard le maximum global ( $> 55$  W). Ensuite, la puissance électrique appelée par les sèche-linge baisse jusqu'à un niveau compris entre 15 et 20 W. Le non retour à la puissance initiale s'explique par le fait que nous avons choisi 24 h comme intervalle de restitution des résultats du modèle<sup>24</sup>.

24. Dans nos résultats sur la durée totale de simulation (100 000 s), la puissance agrégée moyenne continue de baisser jusqu'à un niveau très légèrement supérieur (3 W) au niveau initial de puissance : certains lave-linge continuent donc de fonctionner la nuit comme nous l'avons autorisé dans le scénario principal – cf. tableau 4.1

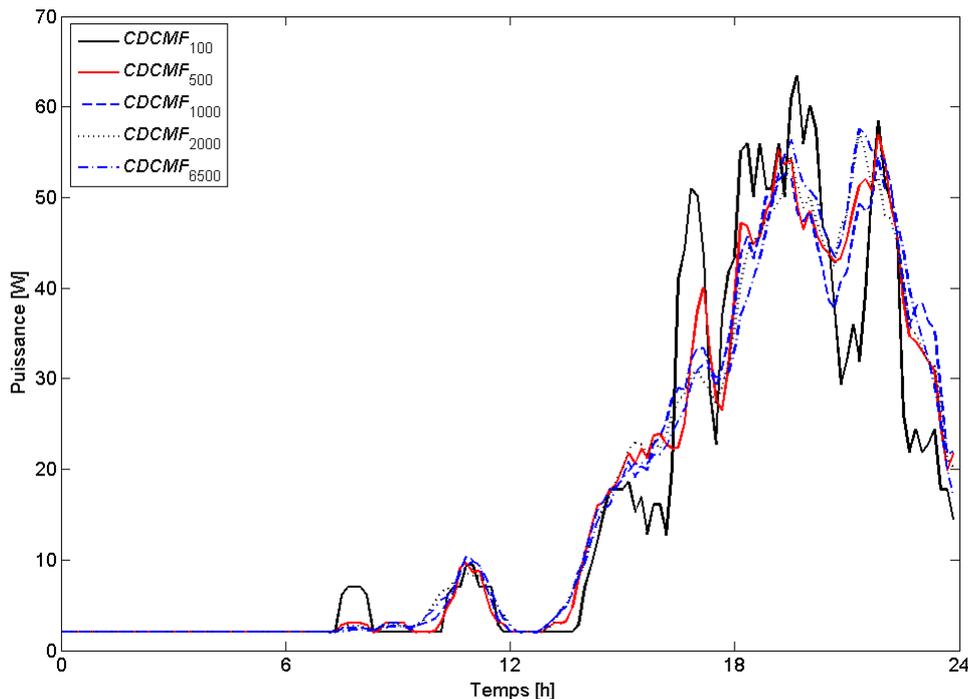


FIGURE 4.18 – Illustration de quelques courbes de charge moyennes foisonnées de sèche-linge calculées à partir de courbes de charge unitaires simulées pour le scénario principal ASL

Dans les deux précédents cas, une diversité de courbes d'appels de puissance donne lieu à des courbes de charge moyennes foisonnées de plus en plus lissées et qui semblent converger au fur et à mesure de l'augmentation du nombre de simulations. Ce phénomène est lié au **degré de simultanéité\***<sup>25</sup> (ou de concomitance) des besoins individuels en électricité pour ces appareils lorsqu'une population conséquente est observée, autrement dit lorsque l'on agrège un certain nombre de courbes de charge unitaires de consommateurs divers. Ainsi apparaît l'effet moyennant – le foisonnement<sup>26</sup> – sur les courbes de charge électriques.

Étudier ces courbes de charge de manière « classique » revient à calculer les indicateurs usuels,  $NVF$  et autres<sup>27</sup>, sur celles-ci à l'image de ce qu'ont proposé Walker et Pokoski [WP85], Capasso et al. [CGLP94] ou encore Widén et Wäckelgård [WW09] (cf. chapitre 1). Les tableaux 4.5<sup>28</sup> et 4.6 synthétisent les valeurs de quelques-uns des indicateurs unaires et binaires estimés pour les  $CDCMF$  de lave-linge et de sèche-linge dont nous venons de faire la discussion.

À l'image de la lecture des tableaux 4.5 et 4.6 et comme nous l'avons mentionné au préalable, l'étude de ces indicateurs usuels de la courbe de charge nous semble relativement limitée. **En particulier, les grandeurs classiques des courbes d'appels de puissance ne constituent pas un critère d'atteinte du foisonnement et la dimension temporelle des courbes de charge n'est pas**

25. Le degré de simultanéité des appels de puissance électrique pour un certain appareil est équivalent à la probabilité de déclenchement de cet équipement

26. Dans un réseau de distribution d'électricité, ce phénomène s'observe au niveau d'un poste source desservant un certain nombre de clients qui appartient ou non au secteur résidentiel

27. Ceux-ci ont été abordés dans la partie 4.1.1

28. Dans ce tableau, nous introduisons deux notations supplémentaires :  $P_{min}$  et  $P_{med}$  qui représentent respectivement la puissance minimale et médiane. D'autre part, le coefficient de foisonnement est relatif au niveau d'agrégation ( $N_{agr}$ ) des  $CDCMF$  plutôt qu'au traditionnel nombre de consommateurs/clients ( $n$ ) qui s'adapte plus particulièrement aux courbes de charge mesurées

Équipement	$N_{agr}$ [-]	$P_{min}$ [W]	$P_{max}$ [W]	$\bar{P}$ [W]	$P_{med}$ [W]	$K_f, N_{agr}$ [-]	$F_{P_{Max}}$ [-]
machine à laver	100	5,00	92,98	21,51	14,02	0,08	0,23
	500	5,00	71,01	21,44	12,94	0,06	0,30
	1000	5,00	65,20	21,38	13,97	0,05	0,33
	2000	5,00	73,31	21,39	13,57	0,06	0,29
	6500	5,00	76,22	21,38	14,08	0,06	0,28
sèche- linge	100	2,00	63,42	15,32	6,15	0,13	0,24
	500	2,00	56,95	15,32	5,49	0,11	0,27
	1000	2,00	54,79	15,37	5,40	0,11	0,28
	2000	2,00	57,48	15,39	6,48	0,11	0,27
	6500	2,00	57,58	15,39	6,40	0,12	0,27

TABLEAU 4.5 – Valeurs de quelques indicateurs unaires pour des  $CDCMF$  simulées de lave-linge et de sèche-linge

Équipement	$N_{agr}$ des $CDCMF$ comparées		$NVF$ [-]	$MAE$ [W]	$MAPE$ [-]	$RMSE$ [W]
machine à laver	100	500	0,16	4,97	0,19	8,49
	500	1000	0,01	1,50	0,07	2,26
	1000	2000	0,01	1,46	0,07	2,27
	2000	6500	0,01	1,22	0,05	2,02
sèche- linge	100	500	0,15	3,48	0,27	5,94
	500	1000	0,01	1,00	0,08	1,73
	1000	2000	0,02	1,12	0,07	1,91
	2000	6500	0,01	0,87	0,05	1,42

TABLEAU 4.6 – Valeurs de quelques indicateurs binaires entre  $CDCMF$  simulées de lave-linge et de sèche-linge

**prise en compte.** En conséquence et pour aller au delà des précédents constats descriptifs, nous avons cherché à mettre en œuvre l’algorithme du Dynamic Time Warping pour analyser des courbes de charge. Nous nous sommes efforcés de tirer profit de cette méthode pour étudier et décomposer le foisonnement affectant les courbes d’appels de puissance précédemment exposées. Nous présenterons ces travaux dans la partie 4.3.

Avant d’en arriver là, nous décrivons dans le paragraphe ci-dessous les résultats en puissance de la simulation d’une variante du scénario principal ASL.

### 4.2.3 Étude des résultats en puissance du passage d’un cycle unitaire à un autre à iso-consommation

Précédemment détaillé au paragraphe 4.2.1.3, le scénario dont nous nous proposons maintenant de décrire les résultats, s’inscrit comme étant une première étude prospective parmi celles qui peuvent être conduites au moyen de notre méthode de reconstitution de courbe de charge.

Pour faciliter la compréhension de la suite du paragraphe, nous notons  $^{inst}$ scénario principal ASL\* le scénario principal ASL dont les cycles unitaires de puissance de machine à laver sont considérés de manière instantanée et dont le niveau de veille a été fixé à 0 W. Les cycles unitaires de puissance dans ce scénario correspondent au  $^{inst}CUP_1$  de la figure 4.12. De même, la variante à iso-consommation de  $^{inst}$ scénario principal ASL\* est notée  $^{inst}$ variante iso-consommation scénario principal ASL\* : dans ce scénario, les cycles unitaires de puissance correspondent au  $^{inst}CUP_2$  de la figure 4.12.

À l’image de la courte étude descriptive des résultats du scénario principal ASL que nous avons présentée au paragraphe précédent, nous allons ici exposer quelques courbes de charge moyennes foisonnées résultantes de la simulation de cette variante à iso-consommation du scénario principal ASL ( $^{inst}$ variante

iso-consommation scénario principal ASL<sup>\*</sup>). Il est ici intéressant d'établir diverses courbes de charge moyennes foisonnées (pour plusieurs niveaux d'agrégation distincts) et de les comparer avec leurs homologues construites avec les résultats du <sup>inst</sup>scénario principal ASL<sup>\*</sup>. Les illustrations de la figure 4.19 mettent en parallèle les *CDCMF* calculées pour les scénarios précédemment mentionnés, et ce, pour quatre niveaux d'agrégation différents (100, 500, 2000 et 6500 *CDCU*).

L'étude des courbes de charge moyennes foisonnées de la figure 4.19 est tout particulièrement intéressante. En effet, elle montre tout d'abord et ce il est vrai, plutôt de manière contre intuitive, que cette modification du cycle unitaire de puissance **pour la variante considérée**, ne semble avoir qu'un **effet de second ordre** sur les *CDCMF*. En effet, la forme et l'amplitude générales de ces courbes pour les deux scénarios peuvent être considérées comme étant identiques<sup>29</sup>. Concrètement cette modification du cycle unitaire entraîne deux conséquences qui se contre-balancent : le gain en matière de puissance électrique économisée à la pointe du cycle unitaire est compensé par l'augmentation du degré de simultanéité des appels de puissance à un niveau agrégé. Autrement dit, avec <sup>inst</sup>*CUP*<sub>2</sub> la diversité de consommation est moindre qu'avec <sup>inst</sup>*CUP*<sub>1</sub>. Nous pouvons également ici souligner l'aspect plus lisse des *CDCMF* de <sup>inst</sup>variante iso-consommation scénario principal ASL<sup>\*</sup> par rapport à celle de <sup>inst</sup>scénario principal ASL<sup>\*</sup>. Ces résultats étaient prévisibles du fait que la durée de fonctionnement de la machine à laver est plus importante dans la variante du scénario à iso-consommation (d'où une probabilité plus grande à chaque instant qu'un lave-linge soit en marche par rapport au <sup>inst</sup>scénario principal ASL<sup>\*</sup>).

Si nous menons une étude plus détaillée de ces courbes de charge moyennes foisonnées, nous pouvons observer un effet de décalage temporel des appels agrégés de puissance de la machine à laver dans la variante à iso-consommation du scénario. Concrètement, la puissance agrégée appelée par cet appareil apparaît systématiquement plus tardivement dans le temps dans la variante à iso-consommation par rapport à l'autre scénario. De plus, la pointe concomitante journalière de cet équipement est plus élevée dans le cas de <sup>inst</sup>scénario principal ASL<sup>\*</sup> que pour la variante de ce précédent scénario. Ce résultat est une conséquence triviale du choix de modification du cycle unitaire de puissance considéré.

Des études supplémentaires doivent être conduites afin d'améliorer nos connaissances en ce qui concerne l'évaluation, à un niveau agrégé, de l'impact en puissance suite à modification du cycle unitaire d'un appareil. Nous devrions notamment tester des scénarios qui mêlent à la fois l'influence du taux de substitution<sup>30</sup> d'un cycle par un autre (plusieurs taux doivent être investigués) et celles du type de modification<sup>31</sup> du *CUP* de référence de l'équipement sélectionné. De manière générale, ceci permettrait d'étoffer et de valider et/ou d'infirmier les conclusions que nous avons tirées précédemment.

Les travaux supplémentaires suggérés ci-dessus n'ont pas pu être réalisés dans le cadre de nos recherches. Ces études constituent donc potentiellement un point d'avancement tout à fait pertinent et riche d'enseignements en matière de connaissance de la courbe de charge domestique, aussi bien dans le contexte actuel que d'un point de vue prospectif.

Ce paragraphe clôture la partie relative à l'étude descriptive des résultats de notre méthode de reconstitution de courbe de charge pour les scénarios que nous avons testés. La section 4.3 expose la mise en application de la méthode du *DTW* sur des courbes d'appels de puissance, en particulier sur des résultats de nos simulations.

29. C'est d'ailleurs pour cette raison que nous avons jugé inutile de proposer ici une description heure par heure des résultats en puissance à l'image des développements du paragraphe 4.2.2 pages 175 et suivantes

30. Dans la variante à iso-consommation dont nous venons de décrire les résultats en puissance, nous avons choisi un taux de substitution de 1, c'est-à-dire que le cycle unitaire de référence du lave-linge (<sup>inst</sup>*CUP*<sub>1</sub>) a systématiquement été substitué par un autre cycle (<sup>inst</sup>*CUP*<sub>2</sub>)

31. Limitation en puissance, ajout ou suppression d'une ou plusieurs phases du cycle unitaire, modification point par point. . .

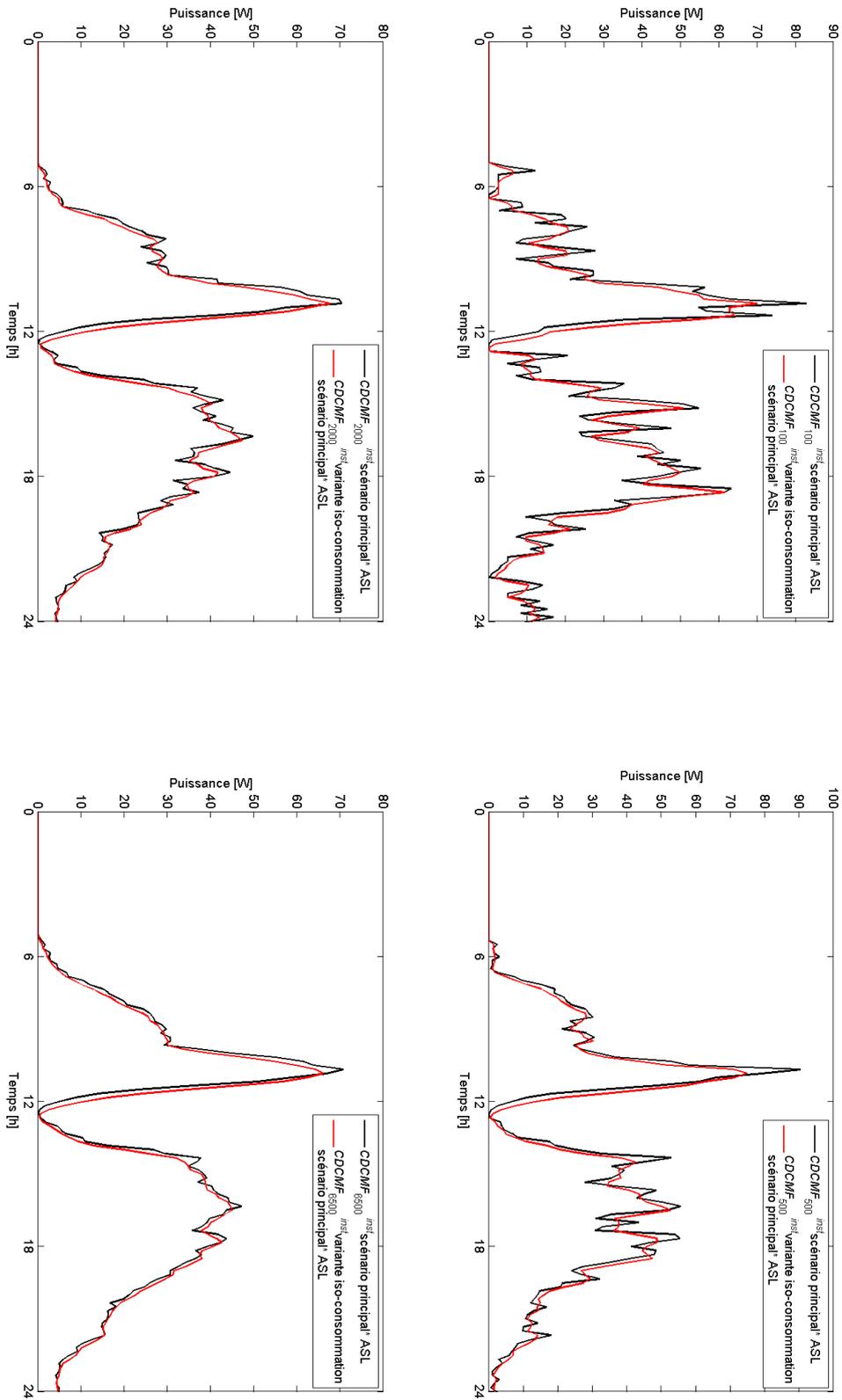


FIGURE 4.19 – Illustration de courbes de charge moyennes foisonnées de machine à laver issues du  $inst$ scénario principal ASL\* et de sa variante à iso-consommation pour quatre niveaux d’agrégation

## 4.3 Mise en œuvre du Dynamic Time Warping pour l'analyse de courbes de charge

Comme nous l'avons souligné dans la partie 4.1.3, le Dynamic Time Warping propose des caractéristiques particulièrement intéressantes dans le cadre de l'étude de courbes de charge. Nous avons donc cherché à tirer profit de cette méthode notamment pour quantifier l'obtention du foisonnement sur des courbes de charge moyennes foisonnées. Nous présentons dans les paragraphes ci-dessous les différentes utilisations du *DTW* au service de notre problématique.

### 4.3.1 Comparaison de deux courbes de charge

Dans un premier temps et afin de mieux nous familiariser avec la technique, nous avons calculé des mesures *DTW* entre différentes courbes de charge de machine à laver issues de la simulation du scénario principal ASL et nous avons observé l'aspect du chemin de déformation associé dans la matrice  $M$ .

Pour des raisons évidentes de pertinence et d'espace requis pour cette analyse, nous réduisons notre premier type d'étude *DTW* à la machine à laver et aux cinq courbes de charge moyennes foisonnées exposées dans la figure 4.17 c'est-à-dire aux *CDCMF* construites à partir des cinq niveaux d'agrégation suivants : 100, 500, 1000, 2000 et 6500 *CDCU*. Concrètement, il s'agit d'établir et de visualiser le chemin de déformation existant entre les courbes étudiées prises deux à deux. Étant donné qu'avec 5 courbes de charge, 10 chemins de déformation peuvent être construits, nous nous limiterons aux seuls chemins existants entre courbes de charge moyennes foisonnées « adjacentes » en matière de niveau d'agrégation c'est-à-dire entre *CDCMF*<sub>100</sub> et *CDCMF*<sub>500</sub> (ce cas est noté *CDCMF* 100-500), entre *CDCMF*<sub>500</sub> et *CDCMF*<sub>1000</sub> (*CDCMF* 500-1000)... ce qui aboutit à 4 chemins de déformation au total. Les croquis (a) à (d) de la figure 4.20 illustrent nos propos et du fait de la densité de ce type de représentation, nous nous attardons volontairement ci-dessous sur l'explicitation de cette figure en en faisant une description détaillée.

Chacune des illustrations de la figure 4.20 représente tout d'abord la matrice  $M$  (carré au centre de chacun des croquis) et le chemin de déformation  $W$ . Celui-ci est représenté par la courbe blanche qui traverse la matrice  $M$  suivant la diagonale principale orientée nord-ouest – sud-est. Dans la suite de cette partie, nous appelons diagonale **NOSE**\* la diagonale précédemment mentionnée. De plus, sur chacune de ces figures, les courbes de charge comparées sont représentées à gauche et en dessous<sup>32</sup> de la matrice  $M$ . Enfin à droite de  $M$  se trouve une échelle de couleurs qui traduit visuellement les valeurs des distances calculées<sup>33</sup> entre les points des courbes de charge comparées.

La propriété de  $W$  de s'orienter suivant la diagonale **NOSE** vient directement de son mode de construction : lorsque deux courbes sont comparées au moyen de la méthode *DTW*, le cas le plus favorable (distance *DTW* minimale et courbes les plus proches) correspond à l'alignement optimal entre les points de ces deux courbes. Dans ce cas, seul subsiste un éventuel écart sur l'axe des ordonnées, c'est-à-dire une différence d'amplitude exprimée en watt, pour chacun des points constitutifs des courbes comparées. Autrement dit et toujours dans le cas le plus favorable, il n'est pas nécessaire de déformer l'axe des abscisses pour rapprocher les courbes d'un point de vue temporel puisqu'il n'y a pas de déphasage entre elles. En définitive, le chemin optimal de déformation dans la méthode *DTW* est toujours la diagonale principale **NOSE** de la matrice  $M$ .

De (a) à (c), nous pouvons observer de plus que le chemin de déformation devient de moins en moins chahuté et donc de plus en plus lisse. Le cas (d), en particulier la dernière partie du chemin de déformation qui témoigne d'un déphasage notable entre les courbes de charge considérées, souligne que les courbes de charge moyennes foisonnées construites respectivement à partir de 2000 et de 6500 simulations semblent plus éloignées entre elles que ne le sont les *CDCMF* issues de 1000 et de 2000 itérations. **Ceci souligne donc que le foisonnement n'est pas atteint après 2000 simulations.** Cependant, nous pouvons

32. Pour reconnaître sur les figures les courbes comparées la règle est simple : si le couple *CDCMF* X-Y est observé, la courbe de charge moyenne foisonnée X sera située à gauche de la matrice  $M$  alors que la courbe de charge moyenne foisonnée Y sera sous la matrice de distance

33. Pour cette étude *DTW*, la métrique choisie est la distance euclidienne

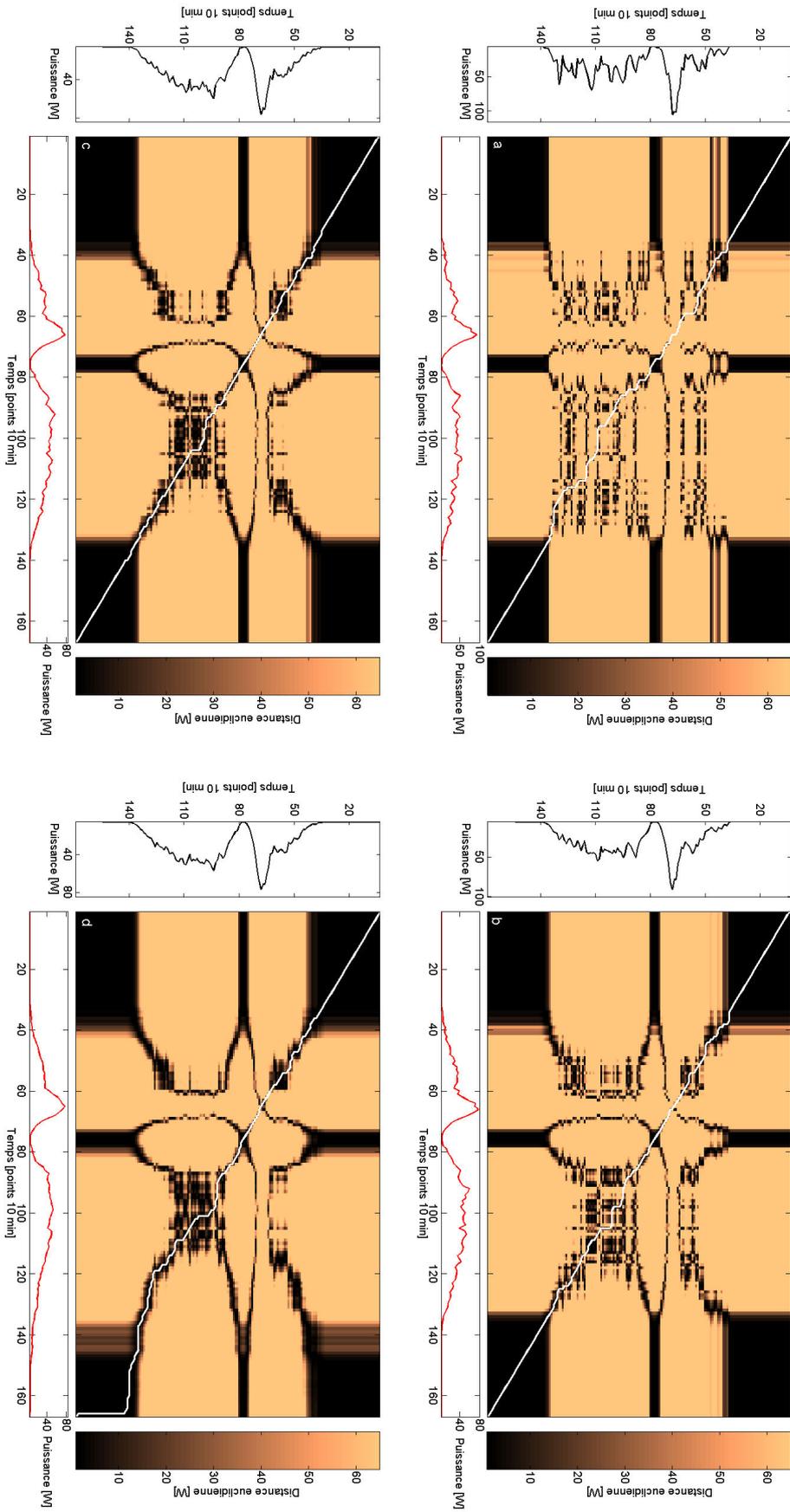


FIGURE 4.20 – Matrice  $M$  et chemin de déformation  $W$  pour les cas  $CDCMF$  100-500 (a),  $CDCMF$  500-1000 (b),  $CDCMF$  1000-2000 (c) et  $CDCMF$  2000-6500 (d)

constater à l'étude de la matrice  $M$  que les zones de divergence entre courbes de charge moyennes foisonnées semblent se préciser et s'atténuer (contours lissés des zones de différence). **Ceci montre que le bruit entourant les *CDCMF* se réduit avec le nombre de simulations.**

### 4.3.2 Comparaison de courbes de charge multiples

Une seconde manière de mettre en œuvre le Dynamic Time Warping au service de notre problématique est de calculer toutes les distances *DTW* possibles entre les courbes de charge à étudier. Il est ainsi constitué une **matrice symétrique de distances *DTW*** dont la diagonale ne comporte que des valeurs nulles. En associant aux valeurs de la distance *DTW* une échelle de couleurs à l'image de celle dont il est fait l'usage au sein des illustrations de la figure 4.20, nous pouvons visuellement estimer la proximité de courbes de charge, et ce, conformément à l'organisation de ces courbes dans la matrice de distances *DTW*. Plus particulièrement, nous avons fait l'usage de cette technique sur les *CDCMF* issues du scénario principal ASL c'est-à-dire à la fois sur les 6500 courbes de charge unitaires de machine à laver mais également sur un échantillon de courbes de même effectif pour le sèche-linge. De cette manière, cela nous permet d'observer les éventuelles ressemblances et les probables divergences portées d'une part par les courbes de charge moyennes foisonnées de ces appareils (lorsqu'elles sont comparées deux à deux) et d'autre part sur le résultat global en regard du foisonnement (lorsque toutes les courbes de charge moyennes foisonnées sont considérées simultanément).

Les figures 4.21 et 4.22 représentent les matrices de distances *DTW* respectivement pour la machine à laver et pour le sèche-linge. Dans les deux cas, il s'agit de matrices carrées  $65 \times 65$ , symétriques, uniquement composées d'éléments de  $\mathbb{R}^+$  où chaque ligne (colonne) correspond à une courbe de charge moyenne foisonnée qui a été construite avec un nombre de courbes de charge unitaires égal à 100 fois l'indice de la ligne (colonne). Par exemple, si l'on considère la 10<sup>ème</sup> ligne (colonne) d'une de ces matrices, cela correspond à la courbe de charge moyenne foisonnée construite à partir de 1000 résultats de simulation. À noter que sur les figures 4.21 et 4.22, les distances *DTW* affichées ont dû être multipliées par un coefficient suffisamment grand pour faire apparaître des différences visuellement décelables sur l'échelle de couleurs employée. En pratique, les distances *DTW* entre courbes de charge de machine à laver ont été multipliées par 750 alors que celles qui ont été calculées entre courbes de charge de sèche-linge ont été multipliées par un facteur 1000. Soulignons enfin que pour les figures 4.21 et 4.22, les courbes de charge moyennes foisonnées ont été obtenues en choisissant de manière séquentielle les courbes de charge unitaires (nous proposons dans la suite un autre exemple avec une construction différente des courbes de charge moyennes foisonnées).

La figure 4.21 expose une matrice majoritairement de couleur claire (signe de « grandes » distances *DTW*) hormis dans une zone qui entoure la diagonale NOISE de couleur noire, les points de cette diagonale correspondent à des distances nulles comme nous l'avons évoqué plus haut. De gauche à droite et de haut en bas, autrement dit dans le sens d'une augmentation du nombre de simulations, la diagonale foncée s'épaissit témoignant de la sorte, de faibles distances *DTW* dans cette zone de la matrice. Pratiquement pour observer l'obtention du foisonnement, il faut s'intéresser aux distances *DTW* entre courbes de charge moyennes foisonnées « proches » c'est-à-dire construites à partir d'un nombre comparable mais non identique<sup>34</sup> de courbes de charge unitaires. Comme nous l'avons souligné au préalable, le passage d'une ligne (colonne) à une autre au sein de la matrice de distances *DTW* correspond à un « saut » du nombre de courbes de charge unitaires considérées. Cet écart est égal à 100. Avec l'échelle de couleurs choisie, nous voyons que ce n'est qu'à partir de 800 simulations que les effets du foisonnement commencent à apparaître sur cette représentation et uniquement pour la courbe de charge moyenne foisonnée la plus proche, i. e. construite avec un écart de 100 résultats de simulations. **La proximité entre *CDCMF***

34. Calculer la distance *DTW* entre courbes de charge moyennes foisonnées construites à partir d'un même nombre de courbes de charge unitaires par sélection séquentielle de celles-ci dans la base de données revient de toute façon à reconstruire la diagonale principale de la matrice de distances *DTW*

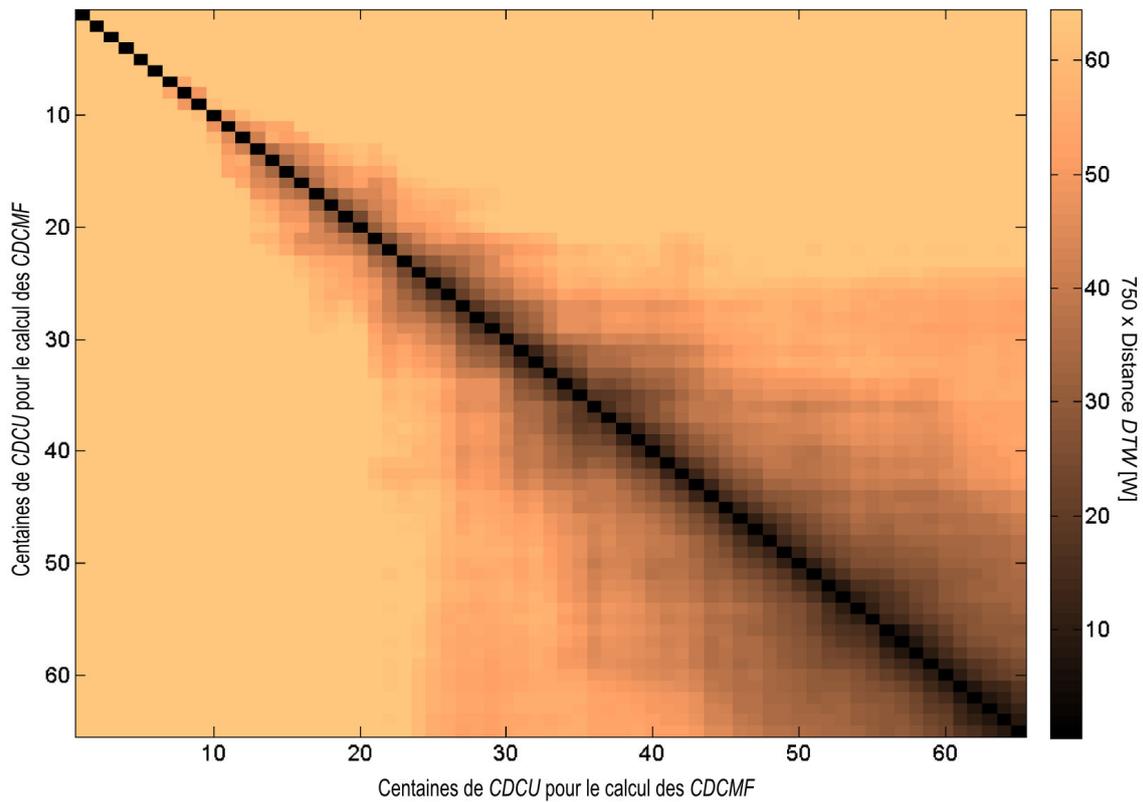


FIGURE 4.21 – Matrice de distances  $DTW$  entre courbes de charge moyennes foisonnées de machine à laver construites par sélection séquentielle des courbes de charge unitaires issues du scénario principal ASL

devient véritablement notable après environ 2000 simulations et ne cesse de croître jusqu'à 6500 résultats.

La figure 4.22 présente le même aspect général que l'illustration 4.21 (matrice d'apparence globale claire et diagonale principale foncée) mais elle propose trois particularités notables :

1. les premières traces visibles de foisonnement sont décelables plus tôt que pour la machine à laver (autour de 600 itérations pour cet appareil alors que le foisonnement du lave-linge n'était visible qu'à partir de 800 simulations) ;
2. l'obtention du foisonnement semble plus rapide et davantage marquée que pour la machine à laver : autour de 3800 résultats unitaires, il semble que les courbes de charge moyennes foisonnées convergent ;
3. un « défoisonnement <sup>35</sup> » survient après 3800 simulations mais les  $CDCMF$  « refoisonnent » ensuite suivant un rythme plus lent et très comparable à celui que nous avons identifié sur la figure 4.21.

Pour interpréter ce comportement atypique, nous proposons deux explications : la première concernant l'arrivée plus rapide et davantage marquée du foisonnement pour le sèche-linge par rapport au lave-linge s'explique par les contraintes du scénario simulé. En effet dans le scénario principal *stricto sensu*, les deux appareils devaient impérativement être déclenchés dans la journée simulée (comme nous l'avons souligné au préalable, ceci est équivalent à dire que  $ProbabiliteDeclenchement = 1$  pour le lave-linge

35. C'est d'ailleurs cette représentation qui nous a permis d'identifier l'existence d'une différence, en matière de scénario modélisé, entre les 6500  $CDCU$  regroupées (d'où la distinction par la suite entre scénario principal *stricto sensu*, scénario principal assimilé et scénario principal ASL)

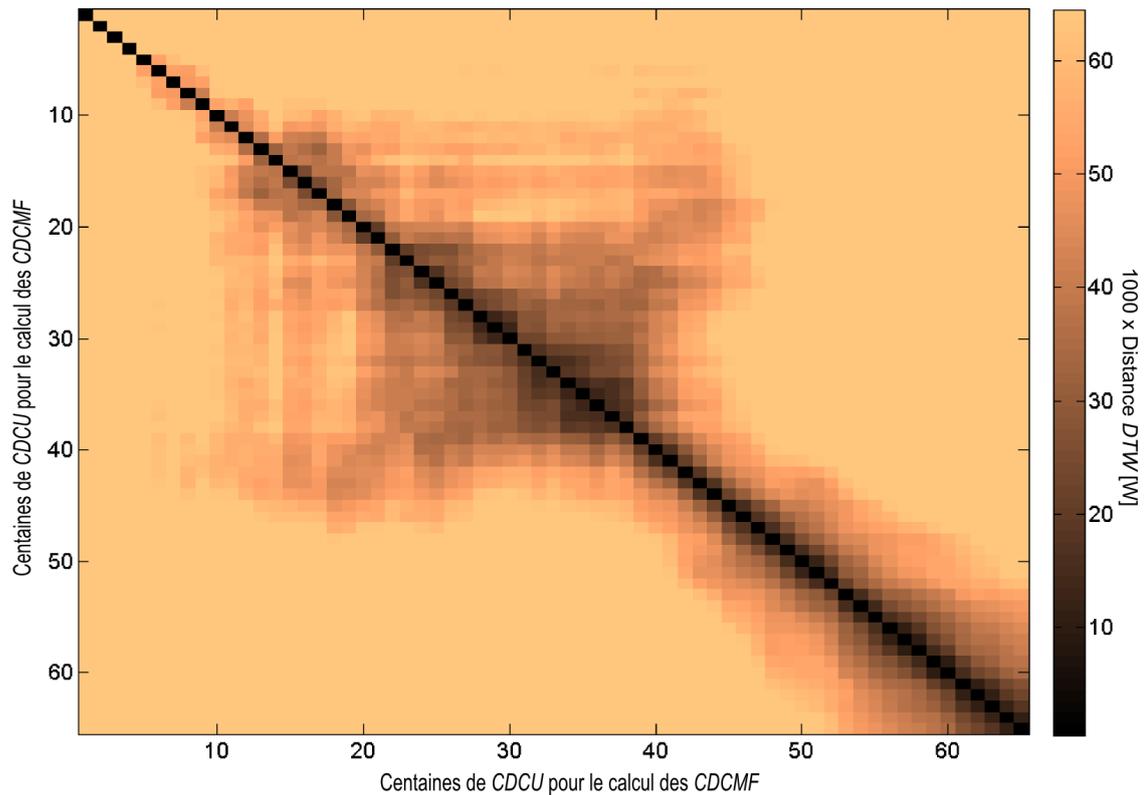


FIGURE 4.22 – Matrice de distances *DTW* entre courbes de charge moyennes foisonnées de sèche-linge construites par sélection séquentielle des courbes de charge unitaires issues du scénario principal ASL

et le sèche-linge). En outre, dans la modélisation (i. e. quel que soit le scénario considéré dans ce chapitre) nous avons introduit la hiérarchie « classique » suivante : le sèche-linge ne peut être déclenché qu'à partir du moment où le cycle de la machine à laver est terminé. **Cette contrainte laisse donc moins de libertés de placement au sèche-linge qu'au lave-linge d'où la convergence plus rapide en termes de courbes de charge moyennes foisonnées.** En second lieu, le phénomène de « défoisonnement/refoisonnement » s'explique par la nature de l'échantillon de courbes de charge unitaires choisies pour cette étude. En effet, comme nous l'avons souligné au début du paragraphe 4.2.2, les courbes de charge issues du scénario principal ASL correspondent à une compilation de résultats de différentes simulations du scénario principal *stricto sensu* et du scénario principal assimilé (cf. note 20 page 173). Une fois agrégés, ceux-ci aboutissent à 6500 courbes de charge unitaires. La modification du seul paramètre `ProbabiliteDeclenchement` pour le sèche-linge est à l'origine de la divergence observée sur la figure 4.22 au seuil de 3800 simulations. Ce précédent nombre coïncide avec une des « frontières » de la compilation des résultats de diverses simulations. Au delà de ce seuil, il s'agit de résultats issus de la simulation du seul scénario principal assimilé. Autrement dit, à partir de ce nombre de courbes de charge unitaires et à cause du changement affectant le paramètre `ProbabiliteDeclenchement` du sèche-linge, les simulations ont abouti à quelques courbes de charge unitaires n'engendrant qu'une consommation de veille de cet appareil étant donné que celui-ci ne se déclenche pas systématiquement au cours du jour simulé.

Au vu de ces observations, nous avons choisi de répéter le précédent travail mais en choisissant cette fois-ci de construire les courbes de charge moyennes foisonnées par sélection aléatoire avec remise au sein de l'échantillon de 6500 courbes de charge unitaires. Les résultats en termes de matrices de distances *DTW* pour la machine à laver et le sèche-linge sont respectivement présentés au travers des figures 4.23 et 4.24 (pour ces deux représentations, le facteur multiplicatif utilisé – cf. page 183 – est égal à 750).

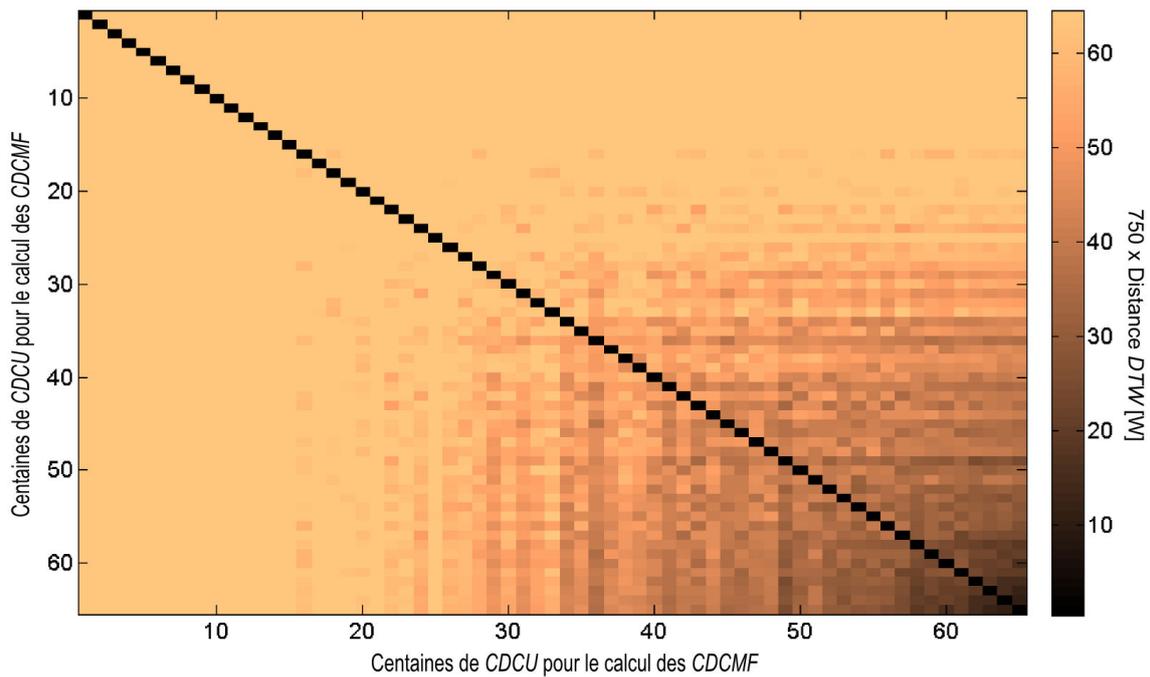


FIGURE 4.23 – Matrice de distances  $DTW$  entre courbes de charge moyennes foisonnées de machine à laver construites par sélection aléatoire de courbes de charge unitaires issues du scénario principal ASL

La figure 4.23 est relativement semblable à l’illustration 4.21. La différence majeure réside dans la vitesse d’apparition du foisonnement qui se traduit par l’élargissement de la zone foncée autour de la diagonale principale. En effet, les prémices de son apparition (tâches très légèrement plus foncées que le fond clair de la matrice) se situent peu après 2000 itérations mais il n’est véritablement notable qu’après 3000 courbes de charge unitaires. À noter également que l’épaississement de la diagonale se fait très progressivement, presque continûment, dans la direction des grands nombres de simulations et ce, jusqu’au nombre maximal d’itérations (6500).

Ces observations s’expliquent directement de la manière avec laquelle sont construites les courbes de charge moyennes foisonnées. En effet, dans le précédent cas – illustré par la figure 4.21 – la sélection dans l’échantillon des courbes de charge unitaires pour la construction des  $CDCMF$  était séquentielle. Autrement dit, la  $CDCMF_{100}$  était représentative des 100 premières simulations,  $CDCMF_{200}$  des 200 premières. . . ce qui revient finalement à élargir progressivement un « socle commun » de courbes de charge unitaires employées pour l’élaboration des courbes de charge moyennes foisonnées. Ce mode de sélection atténue donc l’impact en matière de variabilité des  $CDCMF$  suite à la prise en compte d’une centaine de courbes de charge unitaires pour passer de la  $CDCMF_n$  à la  $CDCMF_{n+100}$ . En revanche, dans le cas de la sélection aléatoire des courbes de charge unitaires pour la constitution des  $CDCMF$ , cet effet de « socle commun » n’existe plus d’où la plus grande variabilité entre les courbes de charge adjacentes et l’apparition plus tardive du foisonnement sur la figure 4.23 par rapport à ce que nous avons pu observer sur l’illustration 4.21. Nous reviendrons plus en détails à la fin de ce paragraphe sur l’impact du mode de construction des courbes de charge moyennes foisonnées.

La matrice de distances  $DTW$  pour le sèche-linge proposée par la figure 4.24 ne présente pas les mêmes particularités que celles de la figure 4.22. En effet, elle suggère une apparition très rapide et progressive du foisonnement sans le phénomène de divergence que nous avons observé sur la figure 4.22 autour de 3800 résultats. Sur la figure 4.24, la progressivité d’obtention du foisonnement est notable au travers de 6 seuils facilement repérables : premières « traces » de foisonnement autour de 500 simulations et accentuation graduelle à environ 1100, 3000, 5000, 6000 et 6200 simulations.

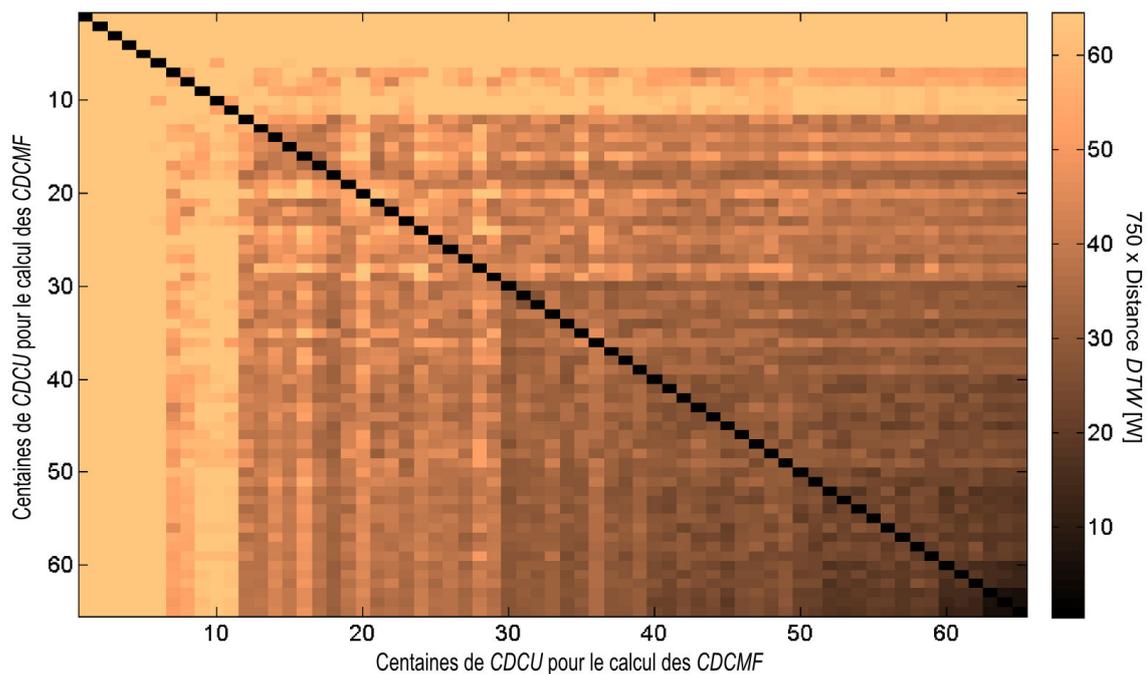


FIGURE 4.24 – Matrice de distances *DTW* entre courbes de charge moyennes foisonnées de sèche-linge construites par sélection aléatoire de courbes de charge unitaires issues du scénario principal ASL

Dans les exemples précédents de matrices de distances *DTW*, quel que soit le mode de sélection des courbes de charge unitaires, nous ne construisons qu'une seule courbe de charge moyenne foisonnée pour chaque centaine de courbes de charge unitaires disponibles. **Les distances *DTW* calculées ensuite entre ces courbes de charge moyennes foisonnées prises deux à deux ne peuvent donc témoigner que d'une partie de l'hétérogénéité existante au sein de l'échantillon de résultats de simulations.** Cependant, comme nous l'avons constaté ci-avant, cette procédure met bien en lumière l'obtention du foisonnement entre les courbes de charge.

Pour aller plus loin dans l'analyse, nous avons décidé d'opérer un travail complémentaire : plutôt que de calculer des distances *DTW* entre courbes de charge moyennes foisonnées construites à partir de différents effectifs de courbes de charge unitaires sélectionnées aléatoirement<sup>36</sup>, nous nous plaçons à un effectif donné, nous construisons plusieurs courbes de charge moyennes foisonnées pour cet effectif et nous déterminons la distribution des distances *DTW* existantes entre ces courbes de charge moyennes foisonnées prises deux à deux.

Les figures 4.25 et 4.26 représentent respectivement ces distributions pour la machine à laver et le sèche-linge. Dans les deux cas, nous avons retenu 17 effectifs<sup>37</sup> (niveaux d'agrégation) pour construire à chaque fois 30 courbes de charge moyennes foisonnées. À chaque effectif, la distribution de distances *DTW* qui y est associée se compose de 435 points (ceci correspond à la somme des 29 premiers termes de la suite arithmétique de raison 1 et de premier terme égal à 1).

Les figures 4.25 et 4.26 témoignent d'une évolution baissière des distances *DTW* calculées entre courbes de charge moyennes foisonnées. **Cette distance *DTW* semble atteindre une asymptote au niveau**

36. Bien évidemment, il n'y a aucun intérêt à effectuer ce travail si les courbes de charge unitaires sont choisies séquentiellement dans l'échantillon des 6500 résultats de simulations

37. Pour ce travail, les effectifs de courbes de charge unitaires retenus pour le calcul de courbes de charge moyennes foisonnées sont les suivants : 100, 200, 300, 400, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000 et 6500

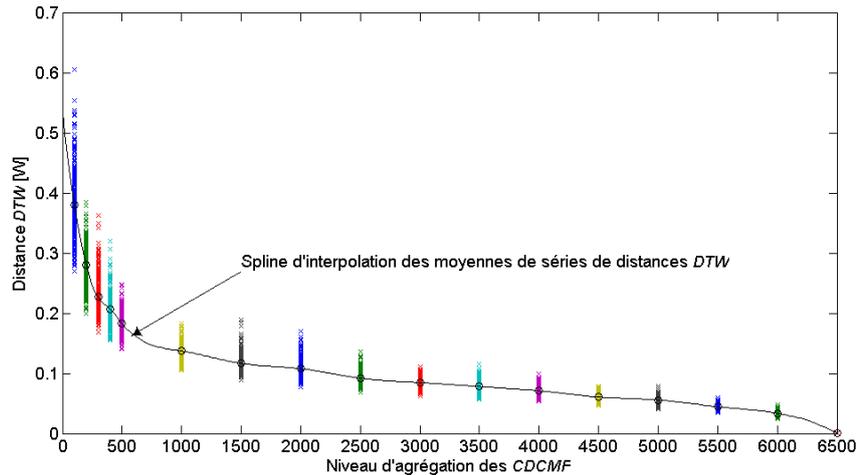


FIGURE 4.25 – Distribution des distances  $DTW$  entre courbes de charge moyennes foisonnées de machine à laver construites à partir de plusieurs effectifs de courbes de charge unitaires

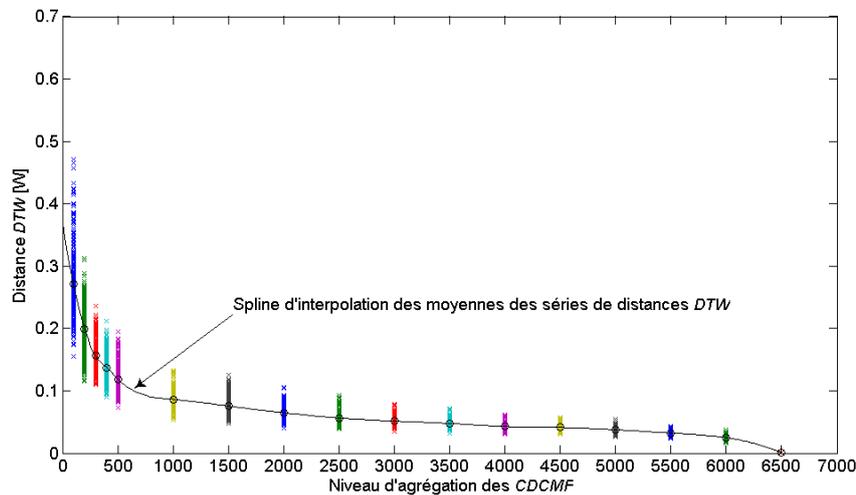


FIGURE 4.26 – Distribution des distances  $DTW$  entre courbes de charge moyennes foisonnées de sèche-linge construites à partir de plusieurs effectifs de courbes de charge unitaires

des **grands nombres de simulations** comme en témoigne dans chacun des cas la spline passant par les valeurs moyennes – repérées par des cercles noirs sur les figures 4.25 et 4.26 – des distributions de distances  $DTW$ . D'autre part au sujet de ces distributions, nous remarquons que **leur amplitude respective se réduit notablement avec l'augmentation du niveau d'agrégation des  $CDCMF$** . En réalité, ce dernier résultat est tout à fait trivial d'un point de vue statistique puisqu'en effet la variabilité entre  $CDCMF$  se réduit nécessairement avec l'augmentation de la fraction

$$\frac{\text{effectif de } CDCU}{\text{taille totale de l'échantillon de } CDCU}$$

En outre, nous avons volontairement choisi la même amplitude sur l'axe des ordonnées des figures 4.25 et 4.26 de sorte de pouvoir comparer plus facilement les tracés entre eux. **L'évolution des distances  $DTW$  pour la machine à laver et le sèche-linge sont de même type** (branche décroissante d'hyperbole). **La différence essentielle entre ces deux tracés réside dans l'amplitude des valeurs prises par la distance  $DTW$** . Pour illustrer nos propos nous pouvons souligner que pour le lave-linge, il faut plus de 2000 simulations pour atteindre une valeur de  $DTW$  inférieure à 0,1 W alors qu'un peu plus de

500 suffisent dans le cas du sèche-linge. Ceci corrobore les conclusions que nous avons pu déduire des illustrations 4.21 à 4.24.

### 4.3.3 Relations entre la distance *DTW* et des indicateurs statistiques usuels

Nous avons souhaité approfondir l'analyse des résultats de distances *DTW* calculées sur des courbes d'appels de puissance. Pour cela, nous avons recherché des éventuelles relations entre cette mesure et des indicateurs habituellement calculés sur des courbes de charge. Plus précisément, nous avons retenu 6 grandeurs d'intérêt qui sont : la puissance maximale journalière  $P_{max,jr}$ , la puissance moyenne journalière  $\overline{P_{jr}}$ , le *NVF*, le *MAE*, le *MAPE* et enfin le *RMSE*. Pour chacune d'elles, nous avons tracé la relation  $grandeur = f(DTW)$  en considérant des courbes de charge moyennes foisonnées de plusieurs appareils issues de nos simulations d'une part, des résultats de la campagne REMODECE d'autre part.

Plus concrètement, quel que soit l'appareil considéré et le nombre  $n$  de courbes de charge unitaires disponibles pour chacun d'eux, nous avons construit  $(\text{div}(n,100)^{38} + 1)$  courbes de charge moyennes foisonnées c'est-à-dire autant de *CDCMF* que le nombre de centaines de courbes de charge unitaires disponibles pour un équipement augmenté de 1. Cette dernière correspond à la courbe de charge moyenne foisonnée calculée en tenant compte de l'effectif total de l'échantillon de courbes de charge unitaires (s'il n'est pas un multiple de 100) à disposition pour un appareil. Ceci explique donc les différents nombres de points visibles sur les graphes que nous avons élaborés et que nous présentons dans la suite de cette partie et en annexes. D'autre part, en ce qui concerne les calculs à proprement parler, nous avons choisi d'estimer « de proche en proche » les valeurs des distances *DTW* et des indicateurs entre les *CDCMF* classées par ordre croissant de leur niveau d'agrégation<sup>39</sup>. Le tableau 4.7 illustre nos précédents propos en explicitant les grandeurs et les calculs effectués pour cette analyse.

Symbole	Grandeur calculée	Calcul opéré
$DTW^{40}$	$DTW$	$DTW(CDCMF_{n+100}, CDCMF_n)$
$P_{max,jr}$	$\Delta P_{max,jr}$	$\frac{ P_{max,jr}(CDCMF_{n+100}) - P_{max,jr}(CDCMF_n) }{P_{max,jr}(CDCMF_n)}$
$\overline{P_{jr}}$	$\Delta \overline{P_{jr}}$	$\frac{ \overline{P_{jr}}(CDCMF_{n+100}) - \overline{P_{jr}}(CDCMF_n) }{\overline{P_{jr}}(CDCMF_n)}$
<i>NVF</i>	<i>NVF</i>	$NVF(CDCMF_{n+100}, CDCMF_n)$
<i>MAE</i>	<i>MAE</i>	$MAE(CDCMF_{n+100}, CDCMF_n)$
<i>MAPE</i>	<i>MAPE</i>	$MAPE(CDCMF_{n+100}, CDCMF_n)$
<i>RMSE</i>	<i>RMSE</i>	$RMSE(CDCMF_{n+100}, CDCMF_n)$

TABLEAU 4.7 – Explicitation des calculs réalisés pour la recherche des relations entre la mesure *DTW* et les indicateurs d'intérêt entre courbes de charge moyennes foisonnées

La figure 4.27 présente les relations existantes entre la distance *DTW* et les indicateurs précédemment mentionnés dans le cas de courbes de charge de machine à laver obtenues par simulation (6500 courbes de charge unitaires). Des représentations similaires pour les autres appareils que nous avons considérés sont consultables dans l'annexe D. **L'intérêt de ces figures est de pouvoir aisément constater quelles**

38. Cette notation représente le quotient de la division entière de  $n$  par 100

39. Dit autrement, nous calculons les distances et les indicateurs entre les *CDCMF* « adjacentes » cf. page 181

40. Pour cette étude, les distances *DTW* ont été calculées avec la métrique euclidienne. Dans la suite de ce paragraphe et sauf précision contraire de notre part, nous adopterons systématiquement cette hypothèse du choix de la métrique

valeurs des indicateurs peuvent être attendues si l'on se place à une certaine distance  $DTW$  entre courbes de charge.

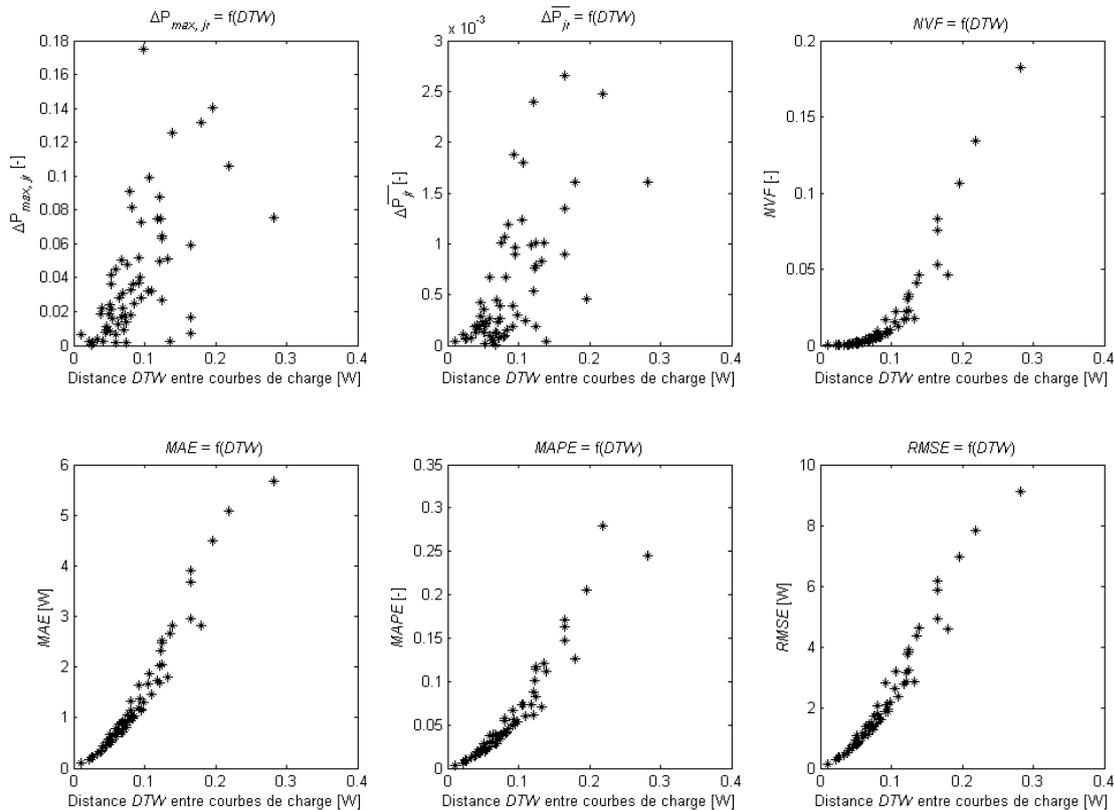


FIGURE 4.27 – Relations entre la distance  $DTW$  et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de machine à laver obtenues par simulation (6500  $CDCU$ )

Sur la figure 4.27, nous pouvons constater que le lien entre la mesure  $DTW$  et  $\Delta P_{max,jr}$  (ou  $\Delta \overline{P}_{jr}$ ) apparaît moins prononcé que ceux existant entre cette même mesure et le reste des indicateurs ( $NVF$ ,  $MAE$ ,  $MAPE$  et  $RMSE$ ). Ces relations semblent être particulièrement notables pour les valeurs faibles de la distance  $DTW$ . Cette constatation était prévisible du fait de la proximité des relations mathématiques existantes entre la distance précédente (calculée avec la métrique euclidienne) et les indicateurs cités au préalable, en particulier lorsque le déphasage temporel entre les courbes de charge comparées est faible voire nul. En effet, dans ce cas, comme nous l'avons mentionné page 164, la distance  $DTW$  non normalisée<sup>41</sup> entre deux courbes de charge sans déphasage entre elles correspond à la valeur de la distance euclidienne calculée entre ces deux mêmes courbes. De plus, une rapide revue des relations 4.1, 4.2, 4.3, 4.7, et 4.12 montre très clairement les proximités existantes entre ces grandeurs mathématiques. Nous pouvons remarquer notamment que la distance euclidienne et le  $RMSE$  ne sont distincts que d'un facteur  $n^{-1/2}$  près.

Étant donné que des courbes de charge simulées et mesurées sont disponibles pour le lave-linge, il est intéressant de procéder à une comparaison des relations existantes entre distance  $DTW$  et les indicateurs sélectionnés. Pour cela, nous avons superposé les résultats obtenus via des mesures et par simulations sur un seul et même graphique (figure 4.28). Cependant, notons tout de suite que nous n'avons pas cherché à reproduire les courbes de charge mesurées de REMODECE, tout simplement puisque nous ne connaissons

41. Ceci signifie, rappelons-le, que le coût minimal de déformation n'est pas divisé en fin d'algorithme par la longueur du chemin du même nom

pas les caractéristiques précises de l'échantillon instrumenté. Il s'agit ici seulement d'établir des constats entre ces deux jeux de données superposés.

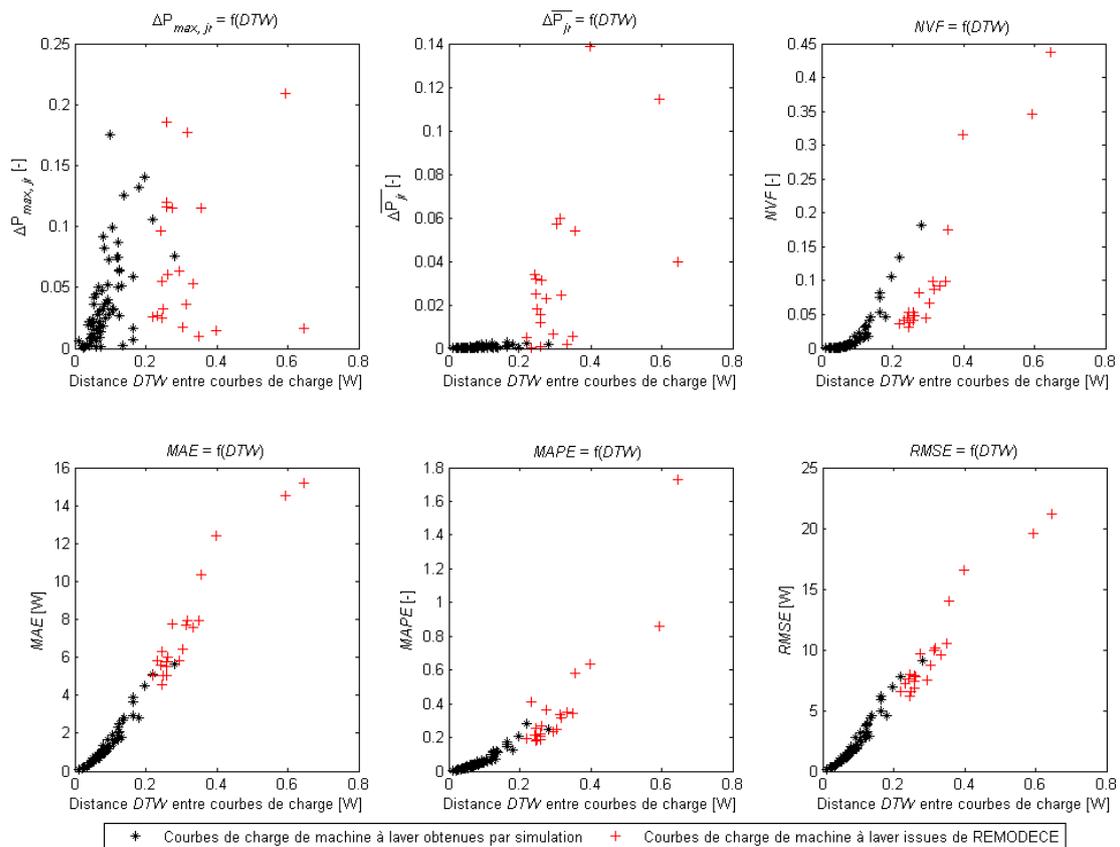


FIGURE 4.28 – Relations entre la distance *DTW* et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de machine à laver obtenues par simulation et par des mesures (respectivement 6500 et 2189 courbes de charge unitaires)

La figure 4.28 permet de se rendre compte que **le scénario modélisé impliquant les machines à laver (scénario principal ASL) est moins divers que la réalité mesurée dans REMODECE**. En effet, tout d'abord les distances *DTW* calculées entre courbes de charge simulées sont plus faibles que leurs homologues évaluées entre données réelles. Ceci a une conséquence directe : les relations entre la distance *DTW* et les indicateurs choisis semblent être également plus évidentes dans le cas de la simulation que dans le cas réel. Au delà d'une plus grande homogénéité temporelle des instants de déclenchement, c'est également l'absence de différenciation en puissance dans nos simulations des cycles unitaires de machines à laver qui est responsable de ce constat. Toutefois, nous pouvons remarquer sur la figure 4.28 que les valeurs obtenues dans le cas réel « prolongent » celles calculées pour les données simulées. **Ceci semble souligner l'existence d'une relation entre certains indicateurs et la distance *DTW*.**

#### 4.3.4 Décomposition des écarts entre courbes de charge moyennes foisonnées

Précédemment, nous avons indiqué que pour deux courbes comparées sans déphasage entre elles et à un facteur de normalisation près, les calculs de distance euclidienne et de distance *DTW* (estimée à partir de la métrique précédente) aboutissaient à la même valeur. En revanche, dès lors qu'une déformation est nécessaire sur l'axe des abscisses pour remettre en phase deux séquences temporelles étudiées, la distance *DTW* est systématiquement inférieure à la distance euclidienne de par son mode de calcul (recherche

du coût minimal de déformation). Ainsi si l'on compare la distance  $DTW$  et la distance euclidienne<sup>42</sup> calculées pour différentes  $CDCMF$  (a fortiori les tracés d'appels de puissance issus de simulations et de mesures), nous pouvons facilement visualiser et quantifier la part des écarts entre courbes de charge relative au déphasage temporel de celle liée à un décalage en puissance. La figure 4.29 présente les proportions des écarts entre  $CDCMF$  qui sont liées d'un côté au déphasage temporel et de l'autre à une différence d'amplitude en watt.

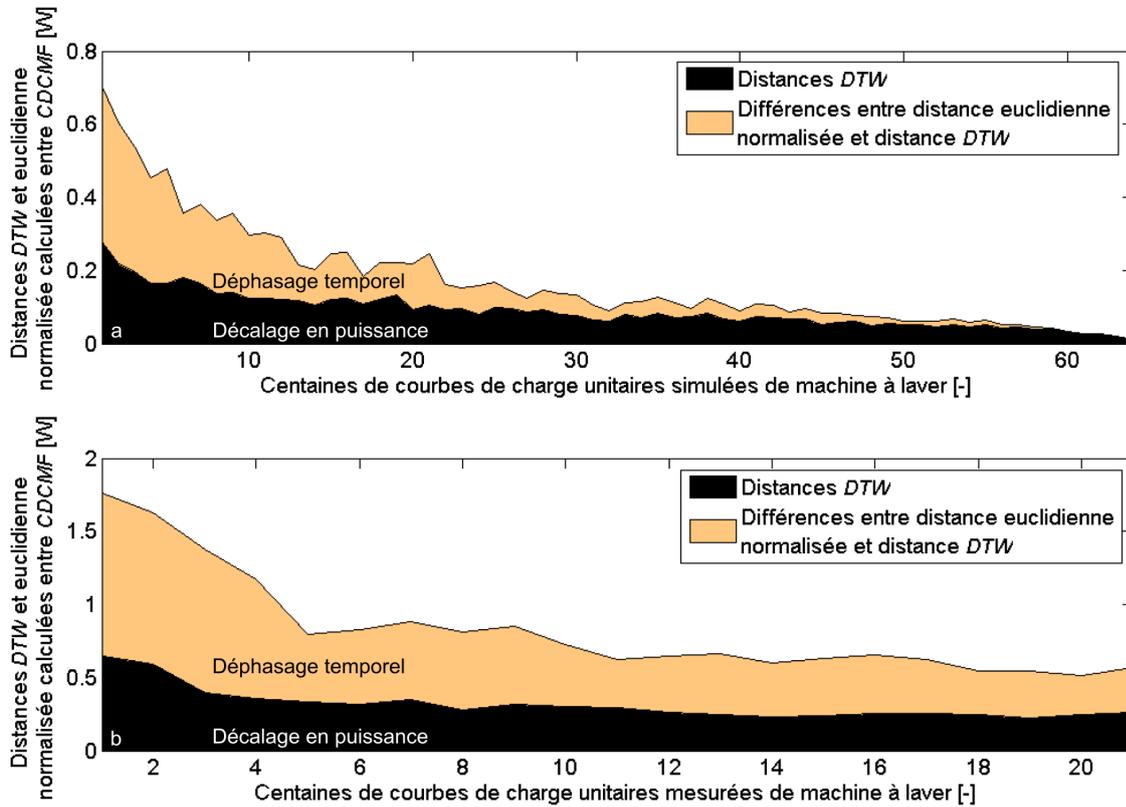


FIGURE 4.29 – Visualisation de la part des écarts d'origine temporelle et en puissance entre  $CDCMF$  de machine à laver issues de simulation (a) et de mesures (b)

L'illustration 4.29 est riche d'enseignements à plus d'un titre. Tout d'abord, elle fait apparaître clairement le rapport entre le déphasage temporel et le décalage en puissance, estimés tous deux entre courbes de charge moyennes foisonnées. Ainsi, nous pouvons remarquer que tant que l'on considère un nombre de courbes de charge unitaires inférieur à 2000 (qu'elles soient simulées ou mesurées), la part du déphasage dans l'écart total entre courbes prévaut sur celle du décalage en puissance. De plus, le déphasage temporel entre  $CDCMF$  décroît progressivement avec l'augmentation du nombre de  $CDCU$  prises en compte pour le calcul des  $CDCMF$ . Dans le cas de la simulation (a) et **pour le scénario modélisé<sup>43</sup>, le déphasage disparaît même totalement à partir d'environ 5900 courbes de charge unitaires** (autrement dit à ce stade, la distance euclidienne normalisée est égale à la distance  $DTW$ ).

Le décalage en puissance, quant à lui, évolue également à la baisse avec l'augmentation du nombre de  $CDCU$  considérées. Néanmoins cette réduction est bien moins prononcée que celle que connaît l'écart relatif au déphasage.

En ce qui concerne les données issues de mesures (b), nous observons les mêmes tendances que celles

42. Plus précisément, nous calculons dans un premier temps la distance  $DTW$  entre  $CDCMF$  avec la métrique euclidienne. Ensuite, nous évaluons entre ces mêmes  $CDCMF$  la différence entre la valeur de la distance euclidienne normalisée, i. e. divisée, par la longueur des courbes de charge – nombre de points temporels pour lesquels des données en puissance sont connues – et la distance  $DTW$ , qui elle rappelons-le, est divisée durant son calcul par la longueur du chemin de déformation

43. Il s'agit toujours du scénario principal ASL

évoquées ci-dessus au sujet des résultats issus de la modélisation/simulation. La première différence dans ce cas réside dans le fait que nous ne comparons pas le même domaine en matière de nombre de courbes de charge unitaires. En effet, pour la machine à laver, seules 2189 *CDCU* sont disponibles dans la base de données de REMODECE. Ensuite, nous pouvons remarquer que l'amplitude des valeurs de distances est supérieure dans le cas de la mesure que dans celui de la simulation. Là encore, nous pouvons attribuer cet écart au fait que le scénario que nous avons modélisé représente une bien moins grande diversité que la réalité mesurée lors de la campagne REMODECE. **En conséquence, la convergence des *CDCMF* est moins immédiate dans le cas réel que dans le cas modélisé.**

Les informations portées par la figure 4.29 peuvent naturellement être recherchées pour d'autres équipements. En plus de cette forme de variabilité, nous avons essayé d'estimer l'impact de la manière avec laquelle sont construites les courbes de charge moyennes foisonnées qui sont sujettes aux calculs des distances. En effet, plusieurs possibilités sont offertes pour les élaborer en partant d'une base de données de courbes de charge unitaires :

- **Tirage Séquentiel (*TS*)<sup>\*</sup>** : pour construire la  $CDCMF_{100}$ , on prend les 100 premières *CDCU* de la base de données, pour  $CDCMF_{200}$ , on choisit les 200 premières etc ;
- **Tirage Aléatoire Avec Remise (*TAAR*)<sup>\*</sup>** :  $CDCMF_{100}$  est obtenue en sélectionnant 100 *CDCU* de manière aléatoire (100 tirages aléatoires consécutifs sans remise) au sein de l'ensemble des *CDCU* disponibles. Après remise des 100 *CDCU* tirées ci-avant,  $CDCMF_{200}$  est construite en extrayant 200 *CDCU* de la base de données (200 tirages aléatoires consécutifs sans remise) et ainsi de suite ;
- **Tirage Aléatoire Sans Remise (*TASR*)<sup>\*</sup>** : la construction de  $CDCMF_{100}$  est identique au mode de tirage précédent. En revanche, les 100 courbes de charge unitaires sélectionnées sont conservées en mémoire pour l'élaboration des autres *CDCMF*. Ainsi pour calculer  $CDCMF_{200}$ , il ne suffit plus que de choisir 100 autres *CDCU* dans la base de données. Pour  $CDCMF_{300}$ , nous gardons les 200 *CDCU* préalablement tirées et nous n'en choisissons que 100 supplémentaires, ainsi de suite pour les autres *CDCMF*.

À ces trois variantes de tirage, il peut s'y ajouter, une opération préalable sur les *CDCU* qui double les cas. En effet, du fait des spécificités de la base REMODECE<sup>44</sup>, il est parfois souhaitable de mélanger les courbes de charge unitaires disponibles avant de les utiliser dans des calculs. En conséquence, les *CDCMF* peuvent être construites avec ou sans permutation aléatoire<sup>45</sup> préalable des *CDCU* de la base de données.

Sur les mêmes courbes de charge simulées de machine à laver que celles sur lesquelles nous avons déjà opéré des traitements ci-avant, nous avons construit des illustrations identiques à la figure 4.29 pour les six cas possibles de [type de tirage aléatoire] × [option choisie de permutation]. Nous présentons ces représentations au travers de la figure 4.30 et nous proposons au sein de l'annexe E des représentations identiques pour les autres équipements simulés et mesurés que nous avons considérés dans cette étude.

La figure 4.30 révèle l'influence du processus de construction des *CDCMF* sur les écarts mesurés existants entre elles. En matière d'amplitude, nous pouvons souligner que les processus *TS* et *TASR* conduisent à des résultats comparables alors que *TAAR*<sup>46</sup> aboutit à des écarts plus conséquents du fait de l'absence de continuité – y compris pour des *CDCMF* adjacentes – entre les tirages impliqués dans la construction des courbes de charge moyennes foisonnées. De plus, la méthode *TAAR* conduit à un biais en matière d'utilisation de l'information disponible sous forme de *CDCU* : deux courbes de charge unitaires peuvent en effet avoir été sélectionnées un nombre très différent de fois dans le cadre de la construction des *CDCMF*. **La méthode *TAAR* est donc à proscrire définitivement.**

En outre, il semble que la permutation préalable des courbes de charge unitaires ne soit d'aucune importance **dans le cas de la simulation de la machine à laver pour le scénario principal ASL**

44. Dans cette campagne pour un équipement donné, les *CDCU* sont classées par individu comme le montre le tableau 4.9

45. Ces options sont respectivement identifiées par *perm* ≡ avec permutation et *nonperm* ≡ sans permutation

46. Le cas *TAAR nonperm* de la figure 4.30 est identique au cas (a) de la figure 4.29

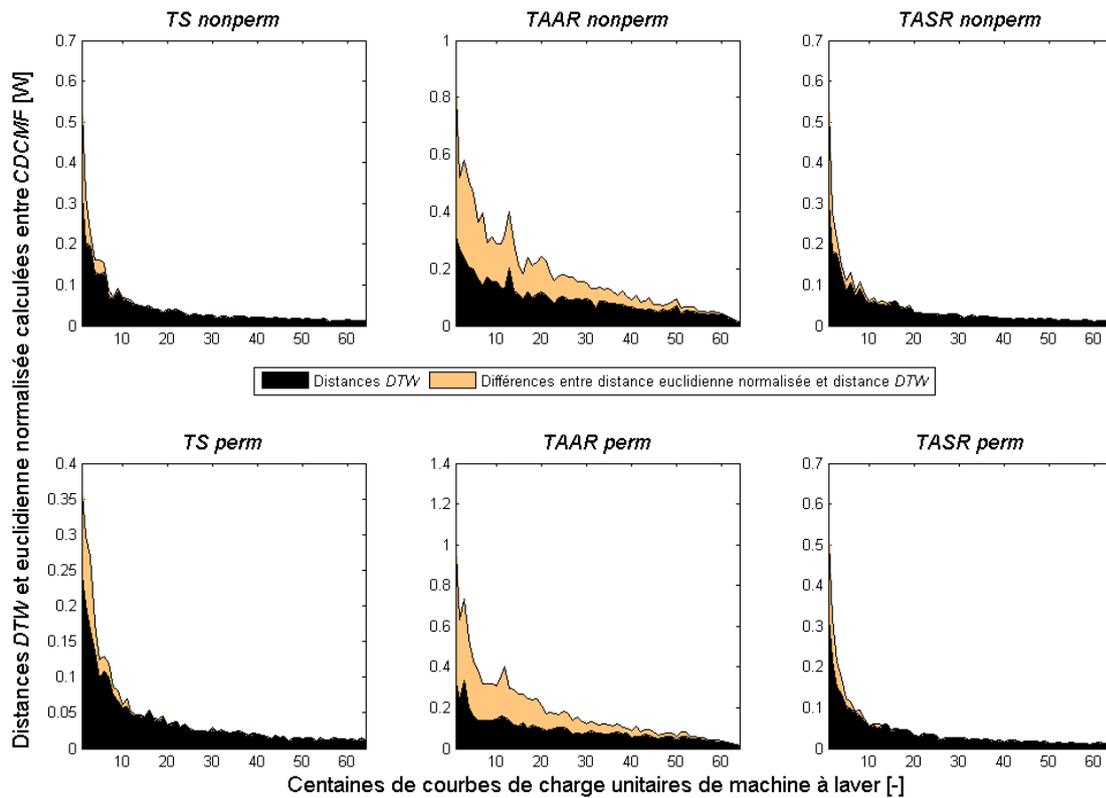


FIGURE 4.30 – Visualisation de la part des écarts d’origine temporelle et en puissance entre *CDCMF* de machine à laver construites par différents processus à partir de *CDCU* simulées.

puisque les graphiques représentant les écarts avec ou sans permutation préalable des *CDCU* sont très similaires (aussi bien en amplitude qu’en forme). Ceci était attendu du fait que nous traitons ici des courbes de charge issues de la modélisation. Contrairement aux données de REMODECE, celles-ci ne sont pas classées selon un ordre précis dans la base de données de résultats<sup>47</sup> issus de la simulation d’un **unique** scénario au sens strict du terme. Dans le cas des données mesurées en revanche (cf. graphes E.2 à E.4), la permutation des *CDCU* dans la base de données est absolument indispensable du fait de la disposition très particulière des courbes de charge unitaires. Cette permutation a un net impact sur les écarts entre *CDCMF* en particulier pour le processus *TS*.

Si maintenant, on s’intéresse aux parts respectives du déphasage et du décalage en puissance des *CDCMF* issues de données simulées, on peut souligner – la méthode *TAAR* étant exclue de notre analyse – que le déphasage entre courbes de charge moyennes foisonnées est pour ainsi dire inexistant dès que l’on considère plus de 1500 courbes de charge unitaires. **Pour le scénario simulé, la variabilité temporelle est donc atteinte après un nombre réduit de simulations.**

La proximité des résultats entre les méthodes *TS* et *TASR*, en particulier lorsqu’un brassage préalable des *CDCU* a été conduit, s’explique mathématiquement. Bien que les *CDCMF* aient été obtenues par deux processus différents, la variabilité des résultats est la même : dans le cas *TS perm*, on tire séquentiellement des *CDCU* dans une base de données de courbes de charge unitaires mélangées ; dans celui de *TASR perm*, on tire au hasard des *CDCU* dans une base de données analogue ce qui au final revient au

47. En revanche pour le sèche-linge, du fait de la différence existante entre les scénarios regroupés (scénario principal *stricto sensu* et scénario principal assimilé), il existe en effet une certaine structure dans la base de données regroupant les courbes de charge unitaires simulées (cette structure transparaît d’ailleurs dans la figure 4.22). Ainsi, la permutation préalable des courbes de charge n’est pas sans importance dans ce cas

même. La seule divergence existante entre les deux procédés est liée au choix des premières centaines de courbes de charge unitaires. Ceci conditionne l'allure plus ou moins chahutée de l'évolution des distances sur l'intervalle  $[0; 1000]$  *CDCU* des graphes de la figure 4.30.

De plus, on peut se demander si l'élaboration des différentes séries de *CDCMF* via les quatre processus précédemment évoqués (*TAAR* étant exclu) conduisent à des résultats stables si plusieurs passes de calculs sont menées. En effet, les résultats du processus *TASR* dépendent du point de départ (courbe de charge) sélectionné par le tirage aléatoire. Ceux de la méthode *TS* dépendent de l'état de mélange de la base de données de *CDCU*.

Pour se donner une idée de la robustesse des conclusions que nous avons formulées au sujet de la figure 4.30 (et des figures analogues sur les autres équipements disponibles dans l'annexe E), nous avons répété les calculs de distances *DTW* et euclidienne normalisée sur une dizaine de séries de courbes de charge moyennes foisonnées construites pour chaque appareil. Il ressort de ce travail que **les conclusions relatives à l'évolution des écarts entre *CDCMF* (déphasage et décalage en puissance) restent justes sur les différents groupes de *CDCMF* que nous avons construits.**

Enfin, nous pouvons préciser que les illustrations permettant la visualisation des parts des écarts de type déphasage et décalage en puissance (figures analogues à l'illustration 4.29) sont complémentaires de celles de type nuage de points montrant les relations entre les indicateurs calculés sur des courbes de charge et la distance *DTW* (illustrations similaires à la figure 4.27). Une fois que le déphasage temporel est gommé, on est dans le domaine des petites distances *DTW* et on peut visualiser, sur les projections sous forme de nuage de points, quelles sont les valeurs des indicateurs attendus puisqu'à ce moment là on  $\delta_{DTW} = \delta_{eucl}$ .

#### 4.3.5 Bilan de la mise en œuvre de l'algorithme *DTW* pour l'analyse de courbes de charge

L'identification d'une insuffisance en ce qui concerne une méthode simple et performante pour mesurer l'écart entre courbes de charge nous a amené à nous intéresser à l'algorithme du Dynamic Time Warping. Nos recherches ont été plus particulièrement axées dans un but d'analyse et de quantification des écarts entre courbes de charge. La technique de visualisation que nous avons établie permet d'observer l'apparition progressive du foisonnement au fur et à mesure que l'on agrège des données unitaires d'appels de puissance. La méthodologie que nous avons mise en place permet en outre d'établir, pour chaque niveau d'agrégation, la composition des écarts (en amplitude et en temporalité) pouvant exister entre des courbes de charge que l'on souhaite comparer. Soulignons ici que **la procédure de visualisation est particulièrement pertinente** lorsque doivent être étudiées simultanément un **grand nombre de courbes de charge**.

Dans le paragraphe 4.3.4, nous avons détaillé un exemple relatif à la comparaison de courbes de charge moyennes foisonnées de machine à laver. Dans cet exemple, nous avons constaté que la part temporelle des écarts entre les courbes de charge s'annulait au delà du seuil de 5900 *CDCU* prises en compte. Ce nombre est à rapprocher d'une notion statistique à savoir le **taux de sondage\*** de la population (ou d'une strate de population). Celui-ci désigne la fraction de la population (de la strate de population) à sonder de manière à obtenir une estimation satisfaisante, c'est-à-dire dans un certain intervalle de confiance fixé au préalable, de la grandeur d'intérêt pour l'ensemble de la population (de la strate de population). L'exemple résumé ci-dessus indique donc qu'avec 5900 courbes de charge unitaires correspondantes au scénario modélisé, on parvient à une représentation convenable en puissance de la diversité de comportement de la population étudiée. Dans notre cas, **le critère d'arrêt des simulations pourrait donc être la suppression de la part temporelle de l'écart entre courbes de charge moyennes foisonnées.**

Dans le cadre de la modélisation de la courbe de charge résidentielle, a fortiori pour l'échelle nationale, identifier le taux de sondage de la population toute entière ou bien pour chacune des strates qui la constitue (par exemple pour chaque ménage-type) est un enjeu majeur puisqu'il coïncide avec le nombre de simulations à effectuer par notre modèle pour avoir une représentation satisfaisante en puissance, d'un point de vue de la diversité, de la population ou de l'une de ses strates.

Dans cette partie, nos études et analyses ont avant tout porté sur des courbes de charge moyennes foisonnées construites à partir d'une sélection de courbes de charge unitaires issues de nos simulations. Nous avons cherché à établir la règle d'évolution de la distance  $DTW$  en fonction du nombre de résultats considérés, et ce, en tenant compte de la variabilité issue de la construction des courbes de charge moyennes foisonnées.

Nous avons ensuite voulu rapprocher la distance  $DTW$  des indicateurs usuels calculés sur les courbes de charge. Dans ce cadre, nous avons superposé les relations obtenues pour nos résultats de simulation avec des données issues d'une campagne de mesures. Nous avons ainsi pu observer la formation de faisceaux de points plus ou moins continus.

Enfin, nous avons valorisé la distance  $DTW$  pour décomposer les écarts pouvant exister entre courbes de charge. En effet, ceux-ci sont formés d'une composante temporelle et d'une composante en puissance. Pour ce qui concerne nos résultats de simulation, nous avons pu observer que les écarts relatifs au déphasage temporel entre courbes est réduit progressivement pour finalement être supprimé au fur et à mesure de l'augmentation du niveau d'agrégation des  $CDCMF$ . Parallèlement, le décalage en puissance connaît également une évolution décroissante mais cet écart s'amortit beaucoup plus doucement. Pour les résultats de nos simulations, celui-ci n'est d'ailleurs pas totalement supprimé. Néanmoins, il faut garder en tête que cet écart sera toujours présent tant que des courbes de charge comparées se différencient en au moins un point.

Dans la suite du chapitre, nous abordons une toute autre thématique expérimentale à savoir la classification de courbes de charge. Cette partie du travail vise à identifier des comportements réels d'utilisation des appareils domestiques à partir de mesures de leurs courbes d'appels de puissance électrique.

Malgré la différence apparente avec les sujets précédemment abordés, nous pourrions constater l'importance d'une distance adaptée à la courbe de charge pour ce travail de classification des appels de puissance.

## 4.4 Identification de comportements d'usage par classification de courbes de charge – Étude d'un cas réel

### 4.4.1 Positionnement du problème et buts de l'étude

L'objectif de reconstituer les courbes de charge électriques résidentielles implique, comme nous l'avons souligné précédemment, de s'intéresser à la diversité présente au sein du parc de logements simulé. En effet, **pour assurer un certain réalisme et une pluralité en matière de consommation d'électricité, nous sommes amenés à définir et à simuler des comportements variés d'usage des appareils** qui sont une réalité effective dans la vie quotidienne. Nous voyons deux manières de définir ces comportements d'usage. Ceux-ci peuvent en effet se traduire par :

- des périodes de fonctionnement, d'arrêt et de veille<sup>48</sup> dans la journée et/ou
- par des probabilités d'instant de déclenchement définies au pas horaire ou infra-horaire ou bien encore dans chacun des intervalles infra-journaliers que nous avons définis au chapitre 3. À chacun de ces instants de déclenchement est ensuite attribuée une distribution de durée de fonctionnement des appareils, distribution qui est à la fois relative à l'équipement et à l'instant de déclenchement considérés.

48. Le cas échéant, suivant les caractéristiques de l'appareil

Afin de mettre en lumière différents comportements de consommateurs, nous avons donc procédé à des analyses de courbes de charge mesurées. Avant d'entrer dans le détail de ce travail, soulignons que des études similaires ont d'ores et déjà été menées par Abreu et al. et reportées notamment dans [ACPF12]. Le but de leurs recherches est d'obtenir des motifs caractéristiques de consommation électrique à partir de données de 15 foyers instrumentés. En plus des valeurs de consommation électrique relevées au pas 15 min, les ménages ont été auditionnés de manière à recueillir des informations qualitatives sur leurs modes de consommation d'électricité. Une analyse des données est menée, foyer par foyer, grâce à une *Analyse en Composantes Principales (ACP)*\* suivie d'une classification hiérarchique et d'une méthode de partitionnement de type k-means<sup>49</sup>. Les résultats montrent que les algorithmes arrivent à distinguer d'une part la consommation de base présente toute la journée et d'autre part, des consommations concentrées sur certaines périodes de la journée et/ou qui sont influencées par la saisonnalité. Les travaux d'Abreu et al. trouvent eux-mêmes leur source dans ceux de Calabrese et al. [CRR10] qui pour leur part cherchaient à identifier et catégoriser des comportements humains à partir de l'analyse de l'activité des points d'accès au réseau Wi-Fi du MIT. Pour cela, Calabrese et al. font l'usage d'une analyse en séries de Fourier, d'une factorisation de la matrice de données et là encore d'une classification k-means.

Dans cette partie, nous avons considéré la télévision principale comme cas d'étude puisque cet équipement est largement utilisé au quotidien par les ménages français comme le suggère l'étude Médiamétrie d'avril 2012 [MED12]. En outre ce choix a été motivé par le fait que nous avons à disposition une base de données de courbes de charge journalières mesurées pour cet appareil. Nous détaillons dans les prochains paragraphes les données employées et les résultats des travaux menés avec un collègue qui ont donné lieu à une présentation sous forme de poster lors d'une conférence internationale [GV11].

#### 4.4.2 Présentation des données analysées

Les mesures utilisées pour cette étude proviennent de l'instrumentation mise en place au sein d'un échantillon de consommateurs domestiques français retenus dans la campagne REMODECE. Ces données sont relativement récentes puisqu'elles ont été collectées en 2007, et avantageusement fournies dans la mesure où, pour le poste audiovisuel, 101 ménages de l'Hexagone ont été instrumentés pendant une période moyenne de 43 jours [ENE08]. Il est utile de souligner que les enregistrements de puissance n'ont pas été conduits tous simultanément : en considérant l'ensemble des ménages instrumentés et en moyenne les 43 jours de mesure, les enregistrements se répartissent donc sur toute une année civile. L'énergie consommée par chacun des appareils instrumentés fut mesurée au pas de temps 10 min par l'intermédiaire de wattmètres branchés en série ce qui a permis d'obtenir des courbes de charge électriques moyennes à cette même résolution temporelle. En plus de la courbe de charge, les télévisions étaient caractérisées par leur technologie (écran cathodique, LCD ou plasma<sup>50</sup>), leur taille (mesure de la diagonale en cm) et leur fonction dans chaque logement (téléviseur principal, secondaire...). Ainsi huit types de télévision ont été définis lors de l'instrumentation. Le codage correspondant à chacun d'eux est explicité dans le tableau 4.8.

Pour limiter le nombre de courbes à traiter, nous avons choisi de considérer exclusivement les appels de puissance de la télévision principale c'est-à-dire les courbes de charge de l'unique téléviseur disponible le cas échéant (pour les ménages ne disposant au maximum qu'une seule télévision) ou par celui de technologie la plus « avancée » (la hiérarchie admise étant plasma > LCD > cathodique) et de plus grande diagonale pour les ménages multi-équipés. Enfin, étant donné que nous souhaitions extraire des comportements d'usage de cet appareil, nous avons retiré de l'ensemble des courbes de charge celles aboutissant à une consommation journalière nulle ce qui correspond à la fois à la non utilisation de l'équipement le jour considéré et à l'absence de consommation de veille.

Suite à ce nettoyage des données, nous aboutissons à une base de 3949 courbes de charge journalières échantillonnées au pas 10 min. Celles-ci peuvent être regroupées sous la forme d'une matrice de 3949

49. Nous invitons le lecteur désireux de plus de précisions sur cette méthode à se reporter à la section 11.2 de [SAP06]

50. Bien qu'ils aient été également instrumentés, nous n'avons pas pris en compte les dispositifs audiovisuels de type projecteur (rétroprojecteur et vidéo-projecteur) qui correspondaient à 9 % des appareils de l'échantillon

Code de la télévision	Propriétés de l'appareil
34	TV CRT, $\varnothing \leq 55$ cm
35	TV CRT, $55 \text{ cm} < \varnothing \leq 72$ cm
36	TV CRT, $\varnothing > 72$ cm
37	TV LCD, $\varnothing \leq 52$ cm
38	TV LCD, $52 \text{ cm} < \varnothing \leq 81$ cm
39	TV LCD, $\varnothing > 81$ cm
40	TV Plasma, $\varnothing \leq 107$ cm
42	TV Plasma, $\varnothing > 107$ cm

TABLEAU 4.8 – Types de télévision principale dans notre analyse de courbes de charge

**observations**  $\times$  144 **variables** où une observation correspond à un jour mesuré et une variable coïncide avec un point 10 min. La structure de cette matrice de données est reproduite au sein du tableau 4.9.

				144 colonnes (points 10 min)							
				0 h 00	0 h 10	0 h 20	...	...	...	23 h 40	23 h 50
3949 lignes	Ménage 1	TV principale du ménage 1	jour 1 jour 2 ... jour j	Mesures de puissance électrique à une résolution temporelle de 10 min							
	Ménage 2	TV principale du ménage 2	jour 5 jour 6 ... jour l								
	...	...	...								
	Ménage k	TV principale du ménage k	jour ...								

TABLEAU 4.9 – Structure des données de courbes de charge analysées

Pour simplifier la lecture des résultats, nous avons choisi d’associer à chaque courbe de charge un identifiant porteur de sens. En effet, celui-ci comporte le numéro du ménage instrumenté, le code correspondant au type de télévision principale mais aussi le numéro du jour de la semaine durant lequel ont été collectées les mesures (1  $\equiv$  lundi, 2  $\equiv$  mardi, ..., 7  $\equiv$  dimanche). Pour rendre cette nomenclature univoque<sup>51</sup>, nous lui avons ajouté l’indice de la ligne (voir tableau 4.9) renfermant la courbe de charge en question.

L’analyse de ces données a été menée grâce à deux méthodes qui sont l’ACP et la [Classification Ascendante Hiérarchique \(CAH\)](#)\*. Au total, quatre études ont été conduites :

1. ACP sur les courbes de charge mesurées – que nous appellerons dans la suite **données brutes** – puis CAH sur les coordonnées des observations dans la nouvelle base, coordonnées issues de l’ACP ;
2. CAH seule sur les données brutes ;
3. ACP sur les courbes de charge mesurées et normées – que nous désignerons dans la suite par le terme **données normées** – en prenant comme référence la puissance électrique maximale de la télévision principale des ménages pour un jour donné. De cette manière, le « biais » causé par les différents niveaux de puissance appelée en fonction des types de télévision est supprimé. Suite à cette ACP, nous avons réalisé une CAH sur les coordonnées des observations dans la nouvelle base ;
4. CAH seule sur les données normées. Deux passes ont été menées : dans la première, nous n’avons pas fixé au préalable le nombre de classes souhaitées ; dans la seconde au contraire, nous avons contraint l’algorithme à former 7 groupes de manière à identifier l’influence éventuelle de chaque jour de la semaine sur l’usage de la télévision.

51. Chaque ménage ayant été instrumenté pendant six semaines en moyenne, à ce stade, chacun des identifiants dont la structure a été précédemment exposée peut potentiellement renvoyer à six courbes de charge différentes

Pour des raisons de place, nous ne détaillerons dans cette section que les résultats des études menées sur les données brutes (cf. parties 4.4.4 et 4.4.5). L'analyse des deux autres études est proposée au sein de l'annexe F. Toutefois nous résumerons dans cette partie les conclusions principales de ces analyses supplémentaires.

Avant d'entrer dans les détails des études sur les courbes de charge de télévision, nous rappelons brièvement dans la partie 4.4.3 les principes de base de l'ACP et de la CAH.

### 4.4.3 Principes des méthodes choisies d'analyse des données

La notion d'analyse des données (*exploratory data analysis*) naît avec les théories de Jean-Paul Benzécri et de John Wilder Tukey qui furent développées dans les années 1960 [TUF10]. Selon Benzécri [BEN82] p. 75, l'analyse des données est « la recherche inductive de dimensions cachées définies par combinaisons de mesures primaires ». Autrement dit, il s'agit d'identifier la structure et les particularités d'un nuage de points qui peut être potentiellement très complexe.

Ces méthodes ont étendu les possibilités d'analyse permises par la statistique classique. Grâce notamment aux progrès de l'informatique, en particulier en matière de puissance de calcul et de capacité de stockage de données, les études statistiques qui initialement s'intéressaient aux valeurs prises par quelques variables pour un nombre réduit d'individus ont vite évolué vers des recherches sur des données à traiter beaucoup plus nombreuses. Ce furent alors des dizaines de variables pour des milliers d'individus qui ont pu être alors considérées dans les travaux statistiques. Bien souvent, les techniques d'analyse s'appuient sur le tableau de données croisant les individus et les variables observées (individus  $\times$  variables). Dans le même temps, le besoin d'une représentation visuelle des résultats a gardé toute son importance.

Dans les années 1990, de nouvelles méthodes ont vu le jour et ont donné naissance à la fouille de données (*data mining*). Ces techniques étendent encore davantage les possibilités d'analyse aussi bien en matière de quantité de données traitées (des millions d'individus et des milliers de variables sur une population suivie qui de plus peut être en constante évolution) qu'en termes de qualité (prise en charge de données lacunaires ou non numériques notamment). Actuellement, ces méthodes sont largement employées dans le monde de l'entreprise et ce, dans différents domaines : agronomie, commerce, loisirs, banque, assurance, médecine... L'étude de ces méthodes dépasse amplement notre travail expérimental et ne sera donc pas conduite dans ce manuscrit. En revanche, nous donnerons quelques éléments de caractérisation des techniques d'analyse des données (analyse en composantes principales et classification ascendante hiérarchique) que nous emploierons par la suite, tout en laissant le soin au lecteur désireux d'en connaître davantage de consulter les ouvrages dédiés dont nous donnerons les références.

#### 4.4.3.1 L'Analyse en Composantes Principales (ACP)

Pour construire cette partie, nous nous sommes principalement appuyés sur le chapitre 7 de [SAP06] et la section 7.1 de [TUF10]; nous convions le lecteur soucieux de plus de détails de s'y référer. En outre, des éléments sur l'histoire de cette technique sont proposés dans [BEN82].

Dans le cas où l'on a à étudier les valeurs prises par un nombre conséquent de variables (en pratique supérieur à 3) pour un ensemble important d'individus (également appelés observations), les techniques statistiques classiques sont inadéquates et notamment l'analyse visuelle du nuage de points qui est tout simplement impraticable. Il faut donc avoir recours à des méthodes plus puissantes qui tiennent compte du caractère multi-dimensionnel du problème. Dans de pareils cas, la première étape est d'organiser les données sous forme d'un tableau (ou d'une matrice) individus  $\times$  variables.

Afin de faciliter la compréhension de cette partie, nous employons une notation « classique » des grandeurs manipulées. Les éléments de cette notation doivent être considérés à part de ceux de la liste présentée en début de manuscrit (p. xiii). Pour cette raison, ils n'y seront d'ailleurs pas répertoriés. En outre, la définition de chaque grandeur sera donnée au moment de l'introduction de celle-ci dans le texte. Enfin, cette définition ne vaut que pour ce chapitre et l'annexe F.

Notons  $n$  le nombre d'individus observés et  $p$  le nombre de variables considérées. Les données à étudier seront donc organisées suivant la structure du tableau 4.10, celui-ci pouvant être identifié à la matrice  $X$ , où chaque  $x_i^j$  représente la valeur de la variable  $j$  pour l'individu  $i$ .

		Variables					
		1	2	...	$j$	...	$p$
Individus	1	$x_1^1$	$x_1^2$	...	$x_1^j$	...	$x_1^p$
	2	$x_2^1$	$x_2^2$	...	$x_2^j$	...	$x_2^p$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	$i$	$x_i^1$	$x_i^2$	...	$x_i^j$	...	$x_i^p$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	$n$	$x_n^1$	$x_n^2$	...	$x_n^j$	...	$x_n^p$

TABLEAU 4.10 – Tableau individus  $\times$  variables

Les observations  $i$  sont chacune décrites par un vecteur de taille  $p$  et elles peuvent être affectées d'une pondération  $p_i$ , appelée également **poïds**, et définie telle que

$$\sum_i p_i = 1.$$

Par défaut, tous les individus ont la même importance et dans ce cas  $p_i = 1/n, \forall i$ . Lorsque l'ensemble des individus est considéré, le nuage de points à étudier, également appelé **nuage des individus**, s'inscrit ainsi dans un espace vectoriel de dimension  $p$  appelé **espace des individus**. Le **centre de gravité** du nuage est le point « moyen » de celui-ci : il est représenté par le vecteur  $\mathbf{g}$  de taille  $p$  qui contient les moyennes arithmétiques  $\overline{x^j}$  de chaque variable de la matrice  $X$  ([SAP06] p. 156) :

$$\mathbf{g}^T = (\overline{x^1}, \overline{x^2}, \dots, \overline{x^p})^{52}.$$

L'idée directrice de l'analyse en composantes principales est de réduire, par projection des observations, l'espace des individus à un plan  $\mathbb{R}^2$  ou à l'espace à trois dimensions  $\mathbb{R}^3$ , de manière générale à un sous-espace de dimension inférieure à  $p$ . Bien évidemment, le choix des variables sur lesquelles les individus sont projetés ne peut pas se faire arbitrairement. D'autre part, la projection d'un nuage de points à  $p$  dimensions réduit nécessairement l'information disponible, mais plus précisément, il diminue les distances entre les objets observés. Le processus de sélection des variables pour la projection doit donc faire en sorte que les distances entre les individus soient le moins diminuées possible de manière à identifier plus facilement ce qui distingue les individus entre eux : c'est justement le processus de l'ACP.

Avant toute analyse en composantes principales les variables sont « homogénéisées » pour simplifier les représentations graphiques et les calculs. En clair, elles sont d'abord centrées par leur moyenne respective (calcul de  $\overline{x_i^j - x^j}$ ) de manière à obtenir des variables de moyenne nulle et ensuite elles peuvent être réduites par leur écart-type  $\sigma^j$  (calcul de  $\frac{x_i^j - \overline{x^j}}{\sigma^j}$ ) de sorte d'avoir des variables du même ordre de grandeur. Il s'agit alors d'une ACP normée. De cette manière, le centre de gravité du nuage de points se situe à l'origine de chaque axe formant l'espace  $\mathbb{R}^p$ .

La recherche des sous-espaces sur lesquels sont projetées les observations se fait axe par axe, et ce, en fonction de la part d'inertie du nuage obtenue suite à projection sur cet axe. Rappelons que l'inertie totale d'un nuage de points est la moyenne pondérée des carrés des distances des points au centre de gravité du nuage ([SAP06] p. 160). Concrètement, l'ACP choisit en premier lieu l'axe qui maximise la part d'inertie du nuage projeté sur celui-ci. Simultanément, cela revient à minimiser la perte d'informations à chaque projection, autrement dit à faire en sorte que la projection des individus sur un axe soit la plus proche

52. Dans cette équation le signe  $^T$  désigne l'opérateur transposé alors la virgule (,) est la représentation conventionnelle du saut de colonne dans une matrice

des individus eux-mêmes (ceci revient à « étirer » le nuage sur les axes). Après cette première projection, l'ACP sélectionne un deuxième axe, orthogonal au premier c'est-à-dire non corrélé à lui, axe qui maximise à son tour la part d'inertie du nuage projeté sur celui-ci et ainsi de suite jusqu'à l'axe  $p$ . Soulignons que l'inertie portée par l'axe  $j$  ( $1 \leq j \leq p$ ) sera toujours supérieure ou égale à celle portée par l'axe  $j + 1$ . En outre, de par la sélection d'axes non corrélés entre eux, l'inertie totale du nuage est la somme de l'inertie portée sur chacun des axes.

Dans l'ACP, il est aussi question de projeter les variables sur des sous-espaces de dimension inférieure à  $n$  en cherchant à expliciter l'inertie du **nuage des variables**. L'ACP normée offre alors plusieurs propriétés facilitant la compréhension de la projection des variables sur les axes. Tout d'abord, la longueur de toute variable  $x^j$  (ou sa norme  $\|.\|$ ) vaut 1. De plus, la coordonnée de la projection d'une variable sur un axe est égale au produit de la norme de la variable (1) par le cosinus de l'angle formé par la variable et l'axe. Ce produit équivaut également au coefficient de corrélation linéaire entre la variable et l'axe.

L'inertie totale du nuage des variables est définie comme la somme des carrés de la norme des variables. Dans le cas de l'ACP normée, cette somme est égale au nombre de variables,  $p$ , et si les poids des individus sont identiques ( $p_i = 1/n, \forall i$ ), l'inertie du nuage des variables est égale à celle du nuage des individus.

Projeter les variables sur des axes suit la même logique que celle employée pour les individus : à chaque étape, la projection de la part maximale d'inertie du nuage est ainsi recherchée. Concrètement à chaque phase du processus, il s'agit de faire en sorte que la somme des carrés des coordonnées des projections des variables sur l'axe soit maximale (dit autrement le processus maximise la somme des carrés des cosinus des angles formés par les variables et l'axe sur lequel elles sont projetées). L'axe cherché est celui qui donne la direction d'inertie maximale, il est appelé **axe factoriel**. Comme pour les individus, la sélection des axes de projection se fait conjointement à une baisse de la part d'inertie du nuage projeté sur les axes. Là encore, la somme de l'inertie portée par chaque axe factoriel est égale à l'inertie totale du nuage. De par la manière dont sont obtenues les inerties sur chaque axe, l'expression de l'inertie totale du nuage correspond à la trace d'une matrice (matrice de corrélation entre les variables dans l'ACP normée – voir [TUF10] pp. 167-168), c'est-à-dire dans ce cas particulier, à ses valeurs propres. Les vecteurs propres de cette matrice correspondent alors aux axes factoriels précédemment mentionnés : le point central de l'ACP consiste donc en la diagonalisation de cette matrice.

Les objectifs de l'ACP sur les variables et les individus font qu'en définitive, il s'agit de trouver, à chaque étape de projection, la variable appelée **composante principale** (ou facteur principal) qui soit la plus corrélée aux variables  $x^j$  analysées et de variance maximale (se reporter à la page 169 de [TUF10]). Les composantes principales sont des combinaisons linéaires des variables centrées réduites (s'il s'agit d'une ACP normée) qui sont au nombre de  $m$  ( $m \leq p$ ) et qui détiennent chacune une information **mesurable** (part d'inertie du nuage de points) et **décroissante** avec leur rang. La première composante principale véhicule une information plus importante que la seconde et ainsi de suite. En se rappelant que l'ACP a un objectif de simplification des représentations des individus dans un sous-espace vectoriel de dimension  $< p$ , il est logique de devoir choisir un nombre restreint de composantes principales, décision qui s'accompagne nécessairement d'une perte d'informations sur le nuage de points. Des critères existent (par exemple le pourcentage d'inertie totale expliquée par les axes, cf. [SAP06] p. 170) quant au nombre de facteurs principaux à conserver.

Dans les deux cas que nous aborderons aux paragraphes 4.4.4 et F.1, nous avons choisi de conserver cinq axes qui expliquent un peu plus de 50 % de l'inertie du nuage.

#### 4.4.3.2 La Classification Ascendante Hiérarchique (CAH)

Comme pour l'ACP, nous ne proposons dans cette partie qu'une description succincte de l'algorithme de classification ascendante hiérarchique et nous laissons à l'initiative du lecteur intéressé d'approfondir ce thème de consulter les ouvrages qui nous ont servi de référence (chapitre 11 de [SAP06], chapitre 9 de [TUF10] et [NC04]).

Le but d'une classification (*clustering*) est de regrouper un ensemble d'éléments, dont les caractéristiques sont connues, en un certain nombre de classes (ou groupes) non prédéfinies les plus **homogènes** possible, et ce, en fonction de la **distance** existante entre les objets considérés pris deux à deux. L'homogénéité interne des groupes et l'hétérogénéité entre ceux-ci peuvent se mesurer grâce à des indicateurs comme l'inertie (ou la variance) intra et inter-classe. Comme son nom l'indique, la classification ascendante hiérarchique fait partie des méthodes hiérarchiques de classification. Celles-ci opèrent par regroupements successifs des objets classés, formant ainsi des parties de l'espace de moins en moins fines (et donc de moins en moins homogènes) jusqu'au regroupement de l'ensemble des éléments considérés.

Le **dendrogramme** (ou arbre hiérarchique), illustré par la figure 4.31, est la représentation usuelle d'une classification hiérarchique où  $e_1$  à  $e_5$  seraient les éléments à classer.

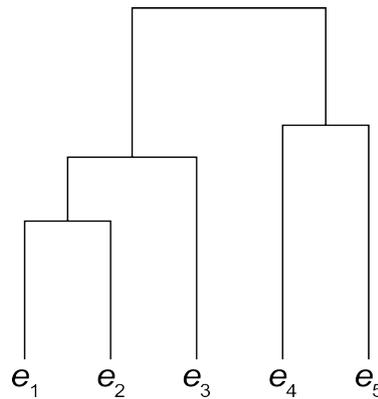


FIGURE 4.31 – Exemple de dendrogramme

Le principe de l'algorithme de la CAH est simple : en partant de  $n$  objets à classer, les distances entre les éléments pris deux à deux sont calculées. Une métrique *ad hoc* doit être choisie de façon à mettre en exergue les similarités et les différences qu'il existe entre les éléments à classer. L'algorithme regroupe ensuite les deux objets les plus proches sous forme d'une classe assimilée à un nœud. Différentes méthodes ou stratégies d'agrégation existent en fonction de la distance ou de la dissimilarité choisie (pour mémoire, les définitions formelles de ces notions ont été données à la section 4.1.2). Nous donnerons quelques éléments de caractérisation de la **méthode d'agrégation de Ward** que nous emploierons dès lors que nous effectuerons une CAH. Il ne reste alors plus que  $n - 1$  objets à classer et l'algorithme recommence les calculs de distances et les regroupements entre les objets les plus voisins pris deux à deux jusqu'à l'obtention d'une classe unique.

Le critère d'agrégation de Ward est celui qui minimise la perte d'inertie entre les classes à chaque étape d'agrégation. En effet, comme nous l'avons souligné au préalable, la classification rassemble des objets de sorte que les classes formées soient les plus différentes possibles entre elles mais d'une grande homogénéité interne. En d'autres termes, **l'algorithme cherche à maximiser l'inertie inter-classe et à minimiser l'inertie intra-classe**. De plus, la diminution du nombre de classes s'accompagne systématiquement d'une baisse de l'inertie inter-classe (simultanément d'une hausse de l'inertie intra-classe). La méthode de Ward rassemble donc les classes de sorte que la perte d'inertie inter-classe soit minimale (ou que le gain d'inertie intra-classe soit minimal). La dissimilarité entre deux classes se mesure alors par la réduction d'inertie inter-classe qui a accompagné leur formation. Il s'agit de la méthode d'agrégation la plus utilisée en CAH car elle s'applique bien aux données pour lesquelles la distance euclidienne est adaptée. C'est d'ailleurs cette distance que nous avons choisie « par défaut » pour les travaux qui suivent. Les différences entre objets s'expriment donc en watt.

#### 4.4.4 ACP sur données brutes puis CAH sur les résultats de l'ACP

Le premier travail d'analyse que nous avons mené a emprunté un processus en deux étapes. Tout d'abord, nous avons effectué une ACP sur les données brutes de sorte d'identifier les grandes tendances au sein du nuage de points. Ensuite, nous avons conduit une CAH sur les résultats de l'ACP, autrement dit sur les coordonnées des observations dans la nouvelle base.

L'ACP effectuée sur le tableau 4.9 a permis de déterminer 144 valeurs propres, chacune d'elles étant associée à un vecteur propre. Le tableau 4.11 présente les résultats de cette ACP sur les cinq premiers axes factoriels.

	F1	F2	F3	F4	F5
Valeur propre	34,380	16,018	12,735	10,458	7,720
Pourcentage d'inertie expliquée [%]	23,875	11,124	8,844	7,263	5,361
Pourcentage cumulé d'inertie expliquée [%]	23,875	34,999	43,842	51,105	56,466

TABEAU 4.11 – Résultats sur les cinq premiers axes de l'ACP conduite sur les données brutes

La figure 4.32 représente la répartition sur les 144 axes de l'ensemble de la variabilité du nuage ainsi que les valeurs propres obtenues par cette ACP.

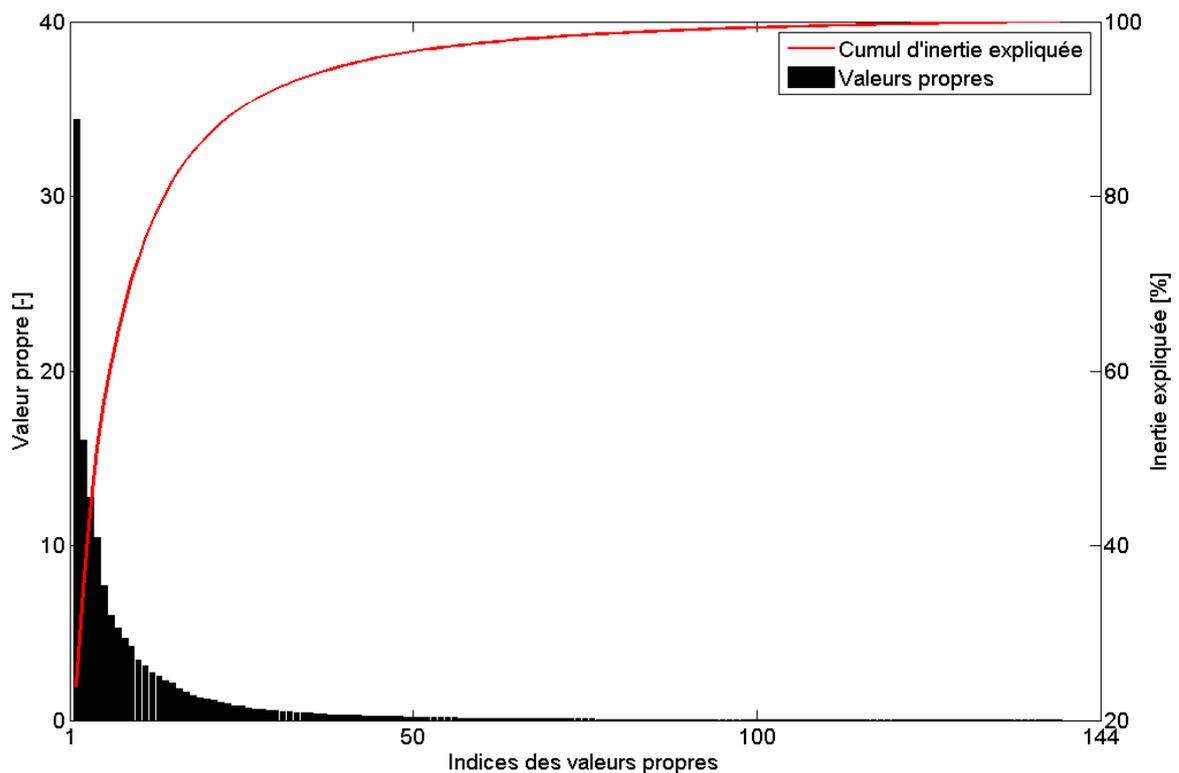


FIGURE 4.32 – Valeurs propres issues de l'ACP menée sur les données brutes et pourcentage cumulé d'explicitation de la variabilité du nuage de points

Les figures 4.33 et 4.34 illustrent les projections des valeurs de puissance sur les trois premiers axes de l'ACP<sup>53</sup>. Les motifs obtenus semblent souligner une certaine corrélation entre les variables puisqu'elles sont arrangées suivant une trajectoire facilement identifiable et orientée. Les projections se répartissent

53. Sur ces figures, les pourcentages désignent les valeurs d'inertie du nuage expliquée par les axes

dans une zone relativement limitée du cercle de corrélation pour la projection sur F1 et F2, plus précisément sur la partie positive du premier axe factoriel soulignant ainsi que toutes les variables sont corrélées positivement entre elles. Ce phénomène est appelé **effet taille** (voir [TUF10] p. 175) et souligne l'intensité d'un phénomène qui est construit à partir de plusieurs éléments sous-jacents. Ici par exemple il s'agirait de la durée journalière de visionnement de la télévision (ou de la consommation électrique du petit écran) qui est fonction du niveau de marche (respectivement du niveau de puissance) de l'appareil pour chaque point 10 min de la journée. En ce qui concerne, les facteurs F2 et F3, les projections forment une zone plus étendue.

Lorsque nous avons étudié les tableaux de résultats comprenant les valeurs de contribution et de cosinus carré des variables, nous nous sommes aperçus que :

- sur le premier axe (F1), les variables 63 à 124 (correspondant aux instants compris entre 10 h 20 et 20 h 30) participent principalement à la construction de F1 (hautes valeurs de contribution et de cosinus carré) et que les variables 51 à 135 (instants compris entre 8 h 20 et 22 h 20) sont celles qui sont les plus expliquées par cet axe ;
- sur le second axe (F2), les variables 1 à 34 (de 0 h 00 à 5 h 30) participent majoritairement à la construction de F2 et les variables 1 à 31 (de 0 h 00 à 5 h 00), 143 et 144 (23 h 40 et 23 h 50) sont les plus expliquées par F2. Les variables 1 à 31 sont responsables de la formation du second axe puisqu'elles présentent à la fois des hautes valeurs de contribution et de cosinus carré ;
- sur le troisième axe (F3), les variables 26 à 39 (instants compris entre 4 h 10 et 6 h 20), 52 à 62 (de 8 h 30 à 10 h 10) et 123 à 142 (de 20 h 20 à 23 h 30) participent de manière prioritaire à la construction de F3. Les variables 32 à 34 (5 h 10 à 5 h 30), 41 (6 h 40) et 136 à 140 (de 22 h 30 à 23 h 10) sont celles qui sont les plus explicitées par ce troisième axe. Ces éléments soulignent que les variables 32 à 34 et 136 à 142 sont concrètement responsables de la construction de F3.

Pour résumer, le premier axe est constitué par des instants où les personnes sont les plus enclines à regarder la télévision au cours de la journée. D'un autre côté, le deuxième axe est établi par une période temporelle durant laquelle la majorité des gens dorment. Finalement, le troisième axe peut être expliqué en tant que regroupement d'instant de transition. Durant cette période, les personnes observées sont à la fois susceptibles d'allumer ou d'éteindre leur poste audiovisuel.

Après cette première étape, nous avons conduit une CAH sur les coordonnées des observations exprimées dans la nouvelle base. Nous avons choisi la méthode d'agrégation de Ward (cf. partie 4.4.3.2) pour réaliser cette CAH et, pour celle-ci, nous n'avons pas prédéfini le nombre de classes à obtenir en fin d'algorithme. La CAH rassemble les courbes de charge en cinq classes comme le montre le dendrogramme de la figure 4.35.

Le tableau 4.12 résume les caractéristiques des classes ainsi que celles de leur objet central respectif. D'autre part, des propriétés de la classification y sont également proposées.

La première classe est la plus conséquente puisqu'elle rassemble plus de 65 % des éléments de la population étudiée. Ensuite, la taille des autres classes diminue avec leur rang et simultanément la variance intra-classe augmente<sup>54</sup> soulignant une plus grande hétérogénéité entre les éléments regroupés. Les objets centraux sont potentiellement remarquables puisqu'ils font office de représentant de chacune des classes. Cependant, du point de vue de la CAH, ces éléments sont exprimés dans une nouvelle base dont les vecteurs correspondent à des combinaisons des variables initiales (points 10 min). Grâce aux identifiants que nous avons mis en place, il est cependant possible de remonter à la courbe de charge correspondante de chaque objet central. Cette représentation mérite en effet notre attention. La figure 4.36 représente donc les cinq objets centraux identifiés par l'algorithme de classification pour le cas considéré.

L'étude de la figure 4.36 semble nous indiquer que la CAH a réalisé une classification d'abord en fonction de l'intensité de l'appel de puissance et ensuite relativement à la forme des courbes de charge traitées. Des

54. La diminution de cette variance entre les classes 4 et 5 ne modifiant pas la tendance générale

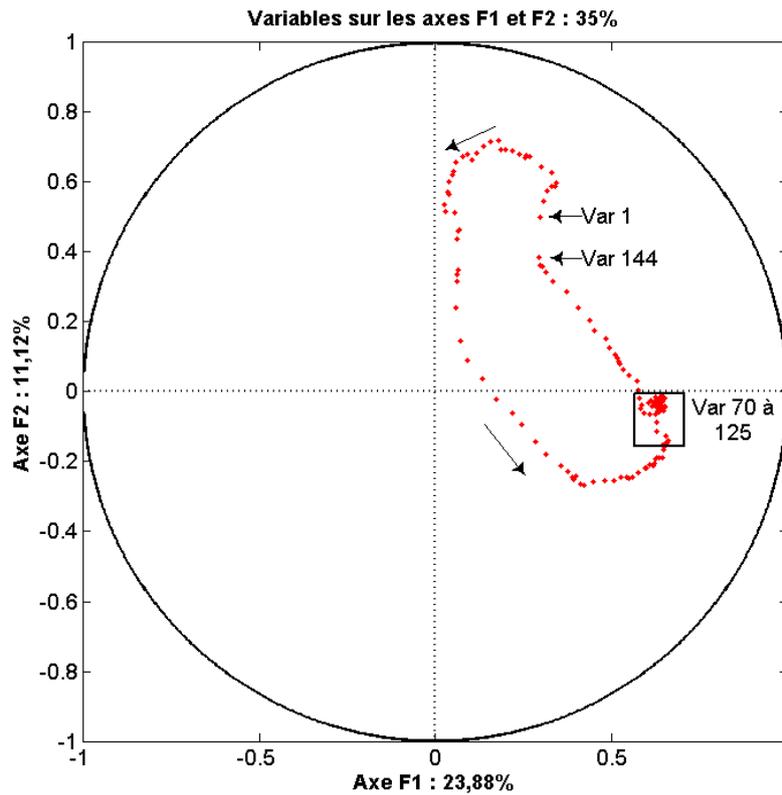


FIGURE 4.33 – Cercle de corrélation entre les variables et les facteurs principaux 1 et 2 de l'ACP sur données brutes

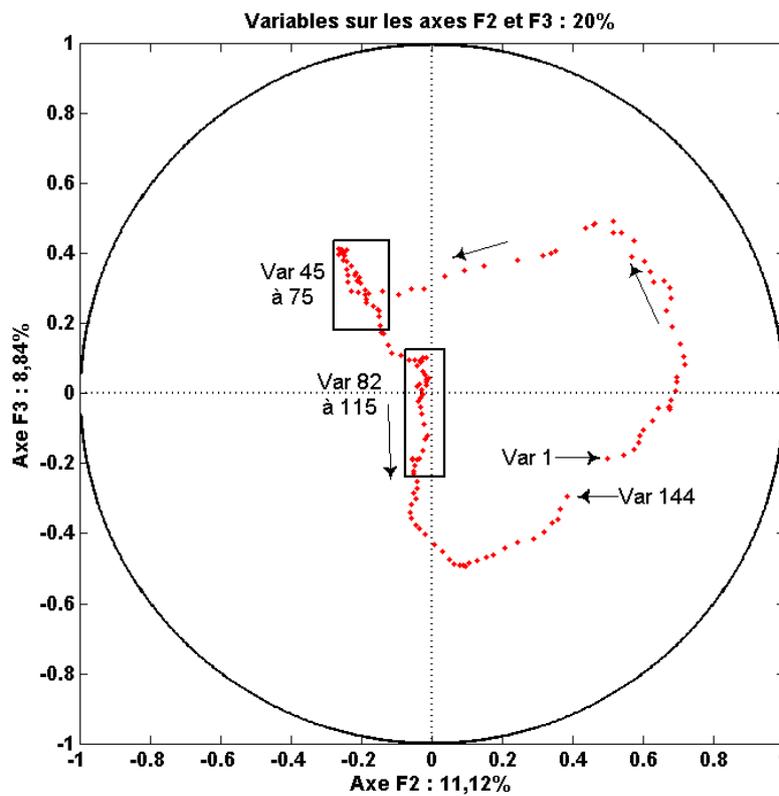


FIGURE 4.34 – Cercle de corrélation entre les variables et les facteurs principaux 2 et 3 de l'ACP sur données brutes

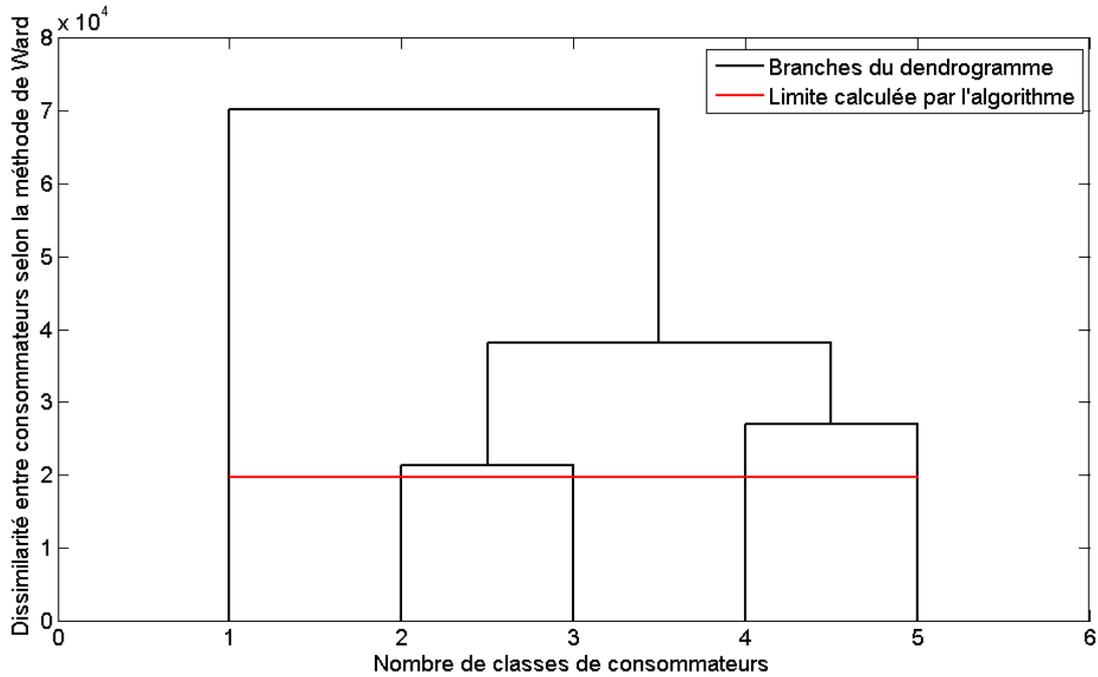


FIGURE 4.35 – Dendrogramme de la CAH conduite sur les résultats de l'ACP menée sur données brutes

Classe	Objet central	Taille de la classe	Variance intra-classe	Distance au barycentre de la classe minimale	moyenne	maximale
1	58_34_1_1119	2586	44,024	2,300	5,833	23,693
2	73_38_6_1774	785	154,501	5,409	11,633	33,829
3	13_40_1_276	264	302,378	7,417	16,543	34,637
4	68_39_2_1540	209	314,378	7,662	16,269	46,514
5	109_38_5_2914	105	310,209	4,495	12,891	68,461

TABLEAU 4.12 – Caractéristiques des classes de la CAH sur les résultats de l'ACP menée sur les données brutes

profils sensiblement différents de consommateurs ont ainsi été identifiés et nous les détaillons ci-dessous :

- 58\_34\_1\_1119 est une courbe de charge d'une petite télévision cathodique ( $\varnothing \leq 55$  cm) qui est utilisée en jour de semaine (dans ce cas, il s'agit d'un lundi) au moment du déjeuner (de 12 h 50 à 13 h 50 environ) et en soirée (de 17 h 20 à 23 h 20) avec deux interruptions d'une vingtaine de minutes peu après 18 h d'une part et autour de 19 h 30 d'autre part. L'appel de puissance de cet appareil est assez faible (de l'ordre de 50 W en mode marche) et il n'y a pas de consommation de veille. Durant la journée, le consommateur 58 adopte un visionnement fractionné de son poste de télévision. Nous pouvons dire qu'il s'agit d'un « téléspectateur actif » au sens où l'appareil semble n'être en marche que lorsqu'un individu est présent dans la maison (en mode éveillé) ;
- 73\_38\_6\_1774 est la courbe de charge d'une télévision LCD de taille moyenne ( $52 \text{ cm} < \varnothing \leq 81$  cm) pour un jour de week-end, ici le samedi. L'appareil est allumé en soirée uniquement (de 19 h 20 à 23 h 00). Durant la marche, le téléviseur appelle une puissance relativement importante de l'ordre de 150 W. De plus, lorsque l'appareil n'est pas allumé, il se trouve dans un état de veille ; ce mode appelle environ 10 W de puissance électrique. Là encore, il s'agit d'un « téléspectateur actif » qui ne semble allumer son poste qu'aux moments où effectivement il est face à son écran et y demeure ;
- 13\_40\_1\_276 correspond à la courbe de charge d'une télévision plasma de taille moyenne ( $\varnothing \leq 107$  cm) pour un jour de semaine (le lundi dans le cas présent). La télévision est allumée une grande

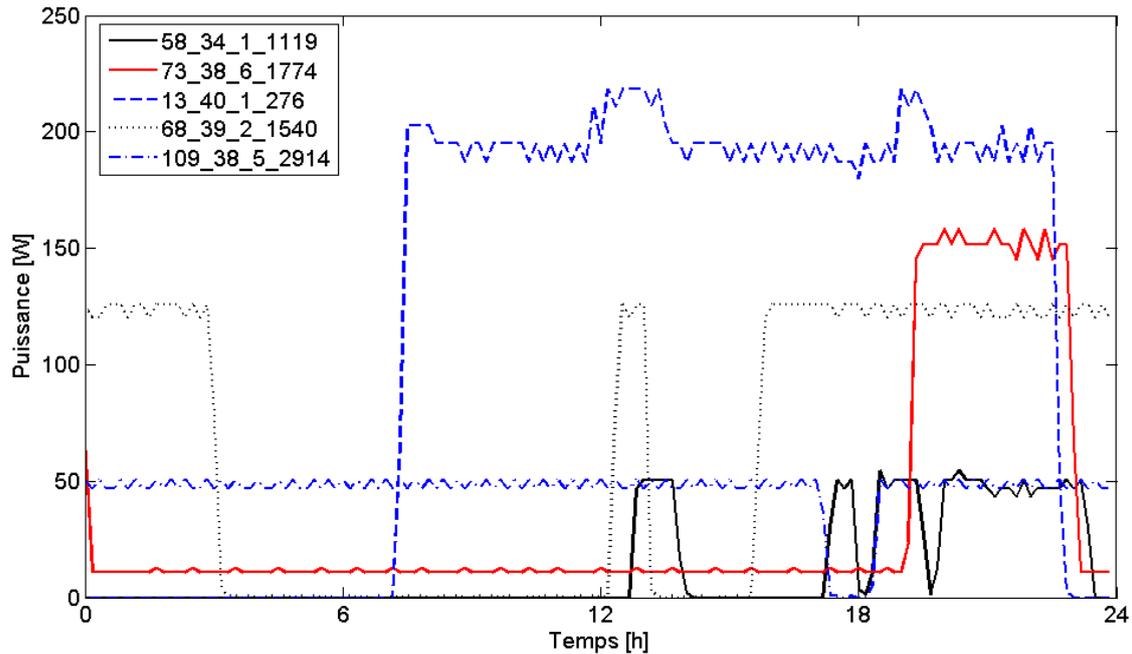


FIGURE 4.36 – Courbes de charge des objets centraux des classes de consommateurs issues de la CAH sur les résultats de l'ACP menée sur données brutes

partie de la journée (de 7 h 20 à 22 h 40). Durant la marche, l'appareil appelle une puissance élevée (de 180 à 210 W). Ces fluctuations semblent s'expliquer notamment par le niveau d'attention que porte le consommateur 13 à son petit écran. En effet, il apparaît qu'au cours des repas (et/ou lors de leur préparation), le niveau de puissance appelée est supérieur au reste de la journée. Ceci peut être lié par exemple au réglage du volume de l'appareil. En outre, cet équipement ne dispose pas de mode veille puisque lorsqu'il est éteint, il ne consomme plus d'électricité. Cet exemple diffère notablement des précédents car la télévision fait ici office de « bruit de fond » durant la majorité de la journée. Nous dirons de ce comportement qu'il est principalement de type « téléspectateur inactif » ;

- 68\_39\_2\_1540 correspond à la courbe de charge d'un téléviseur LCD de grande taille ( $\varnothing > 81$  cm) pour un jour de semaine (le mardi dans le cas présent). Lors du fonctionnement de l'appareil, une puissance fluctuante de l'ordre de 120 à 125 W est appelée par le poste. Celui-ci est utilisé une grande partie de la journée : tout d'abord durant les premières heures du jour observé (visionnement de 0 h 00 à 3 h 00<sup>55</sup>), puis durant la pause déjeuner (de 12 h 20 à 13 h 00) et enfin depuis le milieu de l'après-midi à la fin de la soirée du jour considéré (de 15 h 40 à minuit<sup>56</sup>). Comme précédemment, il semble que ce ménage laisse la télévision allumée dès lors que quelqu'un est présent au domicile. De plus, le rythme chronobiologique des membres de ce ménage paraît ne pas être le même qu'au sein des foyers précédents. En effet, il apparaît que l'habitude d'un couchage tardif ne soit pas un « écart de conduite ». Enfin, lorsque l'appareil est éteint, une très faible consommation de veille (inférieure à 0,5 W) peut être détectée ;
- 109\_38\_5\_2914 est la courbe de charge d'une télévision LCD de taille moyenne ( $52 \text{ cm} < \varnothing \leq 81$  cm) pour un jour de semaine (un vendredi le cas échéant). Cette courbe de charge est assez particulière puisqu'en effet, elle indique une consommation presque continue de la télévision pour le jour considéré : l'appareil appelle de l'ordre de 45 à 50 W tout au long de la journée hormis entre 17 h 20 et 18 h 20. Durant cette dernière période une faible consommation de veille peut

55. La personne concernée est sans doute passionnée de rediffusions et de documentaires traitant de la chasse au gibier...

56. Cette activité télévisuelle se prolonge sur le jour suivant

être observée (puissance appelée inférieure à 0,5 W). Clairement, dans ce ménage, la télévision fait office de présence et est laissée constamment allumée (y compris la nuit). Le comportement de visionnement des occupants est donc de type « téléspectateur passif<sup>57</sup> ».

L'étude des objets centraux de la CAH indique que des profils sensiblement différents d'utilisation de la télévision sont identifiés :

- téléspectateur régulier en journée et en soirée – pas de veille ;
- téléspectateur du soir – présence d'une veille ;
- consommateur « télévision bruit de fond » (continûment en journée et en soirée) – pas de veille ;
- téléspectateur très assidu (pause méridienne et soirée/nuit) – présence d'une veille ;
- consommateur « télévision compagnon » (appareil en marche continue à toute heure du jour et de la nuit) – présence d'une veille.

Malgré tout, la caractérisation des classes construites par la méthode hiérarchique doit être affinée.

De manière à aller plus loin dans l'analyse des classes de courbes de charge, nous avons décidé d'estimer les durées journalières des trois modes de fonctionnement possibles des télévisions à savoir : marche, arrêt et veille. Pour cette raison, nous avons choisi de mettre en place un codage des modes de fonctionnement (-1  $\equiv$  arrêt, 0  $\equiv$  veille et 1  $\equiv$  marche), codage qui doit se construire relativement aux caractéristiques de chaque poste de télévision considéré par l'analyse. En effet, au sein de l'échantillon de courbes de charge considérées dans cette étude, nous avons montré qu'il y avait une forte diversité en matière d'appels de puissance et ce, pour chacun des modes de fonctionnement précédemment cités.

Sans conduire une étude exhaustive (i. e. au cas par cas) des courbes de charge, il n'est donc pas possible d'identifier, de manière univoque et à chaque instant de chaque courbe de charge, le mode de fonctionnement correspondant de la télévision.

Par un jeu d'hypothèses, nous avons donc défini des intervalles de puissance qui correspondent à chacun des modes de fonctionnement de la télévision. Pour un instant donné, le mode marche est défini par un niveau de puissance supérieur ou égal à 10 % de la puissance maximale journalière demandée par l'appareil. Le mode arrêt correspond aux instants où les appels de puissance sont nuls. Enfin l'appareil est dit en mode veille dans tous les autres cas.

À ce codage simpliste, nous avons ajouté une règle supplémentaire pour éviter d'identifier à des veilles les intervalles 10 min durant lesquels la télévision a changé de mode (de marche à arrêt ou d'arrêt à marche). Du fait de l'échantillonnage, ces périodes 10 min peuvent présenter en effet une consommation intermédiaire entre marche et arrêt, consommation qui peut être assimilée à un niveau de veille (ce dernier mode de fonctionnement n'étant d'ailleurs pas forcément pertinent pour l'appareil en question). Nous appelons **pseudo-veille** de pareils états. Pour une courbe de charge donnée qui est codée avec les règles précédentes, si le nombre de 0 (veille) est inférieur ou égal à 6<sup>58</sup>, les points correspondant sont recodés à -1 (arrêt). Avec ce choix du critère de recodage, nous pouvons identifier trois cycles d'allumage/coupage de la télévision au sein d'une journée donnée. Sans la précédente correction, ces instants de transition auraient été considérés comme des veilles.

À l'aide du précédent codage, les durées journalières moyennes de chaque mode de fonctionnement de la télévision et pour chacune des classes sont présentées dans le tableau 4.13

Les durées de fonctionnement en mode veille et arrêt étant les plus dépendantes du codage choisi, nous ne considérerons que la durée en mode marche (valeurs les plus vraisemblables). Sur l'ensemble de l'échantillon (i. e. en considérant les 3949 courbes de charge), la durée moyenne de fonctionnement dans ce mode est de 6,39 h avec un écart-type important puisqu'égal à 4,71 h. Les classes 1 et 2 rassemblent

57. Pour segmenter les téléspectateurs de manière plus fine, une gradation qualitative pourrait être mise en place. Celle-ci serait susceptible par exemple de trouver son origine dans la durée journalière de visionnement du petit écran. Nous aboutirions par exemple à des téléspectateurs modérés, moyens, intensifs...

58. Ce qui correspond à une durée journalière cumulée en mode veille d'une heure au maximum

Classe	Durée quotidienne moyenne par mode de fonctionnement [h]		
	Mode veille	Mode arrêt	Mode marche
1	5,51	13,07	5,42
2	7,29	10,90	5,82
3	4,42	7,44	12,14
4	4,72	10,49	8,79
5	4,66	3,73	15,61
<b>Total</b>	<b>5,73</b>	<b>11,88</b>	<b>6,39</b>

TABLEAU 4.13 – Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH menée sur les résultats de l'ACP conduite sur données brutes

donc des courbes de charge correspondantes à des durées de fonctionnement en mode marche inférieures à la moyenne de l'ensemble de l'échantillon. En revanche, les trois autres classes et en particulier, les classes 3 et 5 (et la classe 4 dans une moindre mesure), semblent rassembler des courbes de charge de téléspectateurs assidus.

D'autre part, nous pouvons nous intéresser à la composition des classes en matière de type de télévision. Une telle analyse est proposée dans le tableau 4.14.

Classe	Composition des classes par type de télévision [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1	5,3	26,4	9,5	1,4	39,1	4,8	12,6	1,0
2	0,0	1,1	3,4	0,0	9,3	8,0	71,1	7,0
3	0,0	0,0	0,0	0,0	0,4	0,0	85,2	14,4
4	0,0	3,3	8,1	0,0	5,3	21,1	38,8	23,4
5	0,0	32,4	12,4	0,0	38,1	5,7	10,5	1,0
<b>Total</b>	<b>3,5</b>	<b>18,5</b>	<b>7,6</b>	<b>1,0</b>	<b>28,7</b>	<b>6,0</b>	<b>30,4</b>	<b>4,3</b>

TABLEAU 4.14 – Composition des classes par type de télévision de la CAH menée sur les résultats de l'ACP sur données brutes

La lecture du tableau 4.14 nous mène aux constats suivants :

- les types 34 et 37 (téléviseurs CRT et LCD de petite taille) sont uniquement présents dans la classe 1. Celle-ci semble d'ailleurs être la classe des téléviseurs de petite et de moyenne taille ;
- les classes 2, 3 et 4 semblent être celles des grands téléviseurs (en particulier la classe 3 qui est formée à 99,6 % par des téléviseurs plasma, types 40 et 42) ;
- la classe 5 semble être celle des téléviseurs de taille intermédiaire puisqu'elle est formée à 70,5 % par les types 35 et 38.

Les précédentes constatations nous ont poussé à mener une étude complémentaire, à savoir analyser la répartition des télévisions au sein des classes. Celle-ci est compilée dans le tableau 4.15.

Classe <sup>59</sup>	Répartition des types de télévision au sein des classes [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1 (65,4)	100,0	93,2	81,1	100,0	89,0	52,1	27,1	15,9
2 (19,9)	0,0	1,2	8,9	0,0	6,4	26,7	46,5	32,4
3 (6,7)	0,0	0,0	0,0	0,0	0,1	0,0	18,7	22,4
4 (5,3)	0,0	1,0	5,6	0,0	1,0	18,6	6,7	28,8
5 (2,7)	0,0	4,6	4,3	0,0	3,5	2,5	0,9	0,6

TABLEAU 4.15 – Répartition des types de télévision au sein des classes de la CAH menée sur les résultats de l'ACP sur données brutes

Le tableau 4.15 permet d'affiner les constatations que nous avons préalablement formulées :

- la classe 1 regroupe plus de 80 % des types 34 à 38 (télévisions dont l'amplitude des appels de puissance est faible à moyenne) ;
- les télévisions grosses consommatrices d'énergie (types 39, 40 et 42) sont pratiquement absentes de la classe 5 ;
- la classe 2 est celle qui « absorbe » le plus les téléviseurs plasma et LCD de grande taille.

Une autre dimension à analyser est la composition des classes en matière de type de jour de la semaine. Cette composition est visible au travers du tableau 4.16.

Classe	Composition des classes par type de jour [%]						
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
1	14,5	14,2	14,1	14,1	14,5	14,4	14,2
2	14,5	15,4	14,9	14,0	14,4	12,6	14,1
3	15,2	12,9	14,0	14,4	12,1	16,3	15,2
4	14,8	12,4	12,0	9,1	12,4	15,3	23,9
5	8,6	12,4	11,4	18,1	14,3	21,0	14,3
<b>Total</b>	<b>14,4</b>	<b>14,2</b>	<b>14,1</b>	<b>13,9</b>	<b>14,2</b>	<b>14,4</b>	<b>14,8</b>

TABLEAU 4.16 – Composition des classes par type de jour de la CAH menée sur les résultats de l'ACP sur données brutes

Le tableau 4.16 nous indique que la composition des classes 1, 2 et 3 par type de jour est relativement homogène<sup>60</sup>. La classe 4 attire davantage à elle les dimanches et « repousse » les jeudis. La classe 5 quant à elle favorise les jeudis et les samedis et laisse de côté les lundis.

L'analyse de la composition des classes en fonction de la durée journalière de visionnement de la télévision, du type d'appareil possédé et enfin de la répartition des types de jours nous conduit aux conclusions suivantes :

- la classe 1 est celle des téléspectateurs réguliers, sans distinction entre les jours de la semaine et qui possèdent une télévision peu consommatrice d'énergie ;
- la classe 2 semble représentative de téléspectateurs réguliers qui disposent d'un poste relativement énergivore et qui passent plus de temps devant leur écran en semaine que durant le week-end ;
- la classe 3 regroupe les téléspectateurs au comportement intensif, détenant un appareil très consommateur d'électricité, et qui conduisent cette activité en particulier en début de semaine et le week-end ;
- la classe 4 contient des consommateurs assidus de la télévision, disposant d'un équipement relativement énergivore et qui plébiscitent la fin de semaine pour s'adonner à ce passe-temps ;
- la dernière classe rassemble des téléspectateurs au comportement extrême<sup>61</sup> (en matière de durée journalière de visionnement), qui possèdent un poste de moyenne gamme et qui est davantage utilisé en milieu et en fin de semaine.

59. Les nombres entre parenthèses indiquent la proportion respective en pourcentage de chaque classe dans l'effectif total des courbes de charge : par exemple, la classe 1 comporte 65,4 % des courbes de charge de l'échantillon. . .

60. L'homogénéité stricte au sein d'une classe se serait traduite par une composition égalitaire de chacun des jours de la semaine. Autrement dit, chaque jour constituerait la classe à hauteur d'un septième ( $1/7 \approx 14,3$  %). Cependant, une condition *sine qua none* de l'homogénéité des classes réside dans la qualité de l'échantillonnage des jours de mesure : la dernière ligne du tableau 4.16 nous montre par exemple que les dimanches sont légèrement sur-représentés alors que les jeudis sont relativement minoritaires

61. Il faut néanmoins prendre du recul par rapport à la valeur de la durée de visionnement moyenne de cette classe (15,61 h). En effet, cette dernière est directement liée au codage des modes de fonctionnement que nous avons exposé page 208. Ce codage est relativement simpliste et il ne permet pas en outre d'identifier les courbes de charge journalières de télévision uniquement formées de puissance de veille. Il eut fallu justement filtrer ces courbes de charge qui faussent la formation des classes et qui deviennent parfois les objets centraux de celles-ci – cf. figures F.4, F.5 et F.7

En conclusion, les résultats du processus « CAH sur les résultats de l'ACP » ont permis d'établir une classification des courbes de charge qui rend possible une identification de cinq types d'utilisation de la télévision. Ceux-ci se différencient en matière de niveau d'appels de puissance (type de télévision), en forme (régularité de la demande) et en durée quotidienne de visionnement du petit écran. Le type de télévision, qui conditionne le niveau d'appel de puissance de l'appareil, semble être le facteur principal d'influence de la classification. En outre, le type de jour apparaît comme un facteur secondaire intervenant dans la formation des classes.

Toutefois, ces résultats doivent être vérifiés et validés par des méthodes statistiques plus élaborées que les simples éléments descriptifs dont nous avons fait l'usage. Il serait particulièrement intéressant de conduire un certain nombre de tests (par exemple sur la valeur moyenne d'un ou plusieurs critères descriptifs des classes) pour infirmer ou confirmer les conclusions de cette étude. De la sorte, les classes pourraient être dites différentes (ou non) au sens statistique du terme, pour la caractéristique sélectionnée, et à un certain taux d'erreur préalablement fixé.

#### 4.4.5 CAH sur les courbes de charge mesurées

Nous avons réalisé une CAH sur les données brutes de manière à déterminer la participation à la classification de l'ACP. Comme lors de la précédente CAH, nous n'avons pas contraint l'algorithme de classification à un nombre de classes prédéfini et nous avons conservé la méthode d'agrégation de Ward.

La CAH aboutit à trois classes dont les objets centraux sont ceux présentés dans le tableau 4.17 tout comme d'autres éléments de caractérisation des classes.

Classe	Objet central	Taille de la classe	Variance intra-classe	Distance au barycentre de la classe		
				minimale	moyenne	maximale
1	80_35_6_1954	2690	254580,97	140,42	447,25	1721,81
2	73_38_6_1767	698	855174,14	325,21	873,01	2254,61
3	68_39_7_1566	561	1313597,33	498,38	1088,47	2407,28

TABLEAU 4.17 – Caractéristiques des classes de la CAH sur données brutes

L'étude du tableau 4.17 nous permet de constater que la classe 1 est la plus importante en termes d'effectif : elle regroupe 68 % de l'ensemble des courbes de charge traitées. Suivent ensuite par ordre d'importance décroissante les classes 2 et 3. Ensuite, il est utile de souligner la différence en matière d'ordre de grandeur de la variance intra-classe pour cette CAH en comparaison avec la précédente (voir tableau 4.12). L'écart s'explique par le fait qu'ici, les données utilisées pour la classification sont les valeurs de puissance qui constituent les courbes de charge et non plus les coordonnées dans la base produite par l'ACP de ces mêmes données en puissance. Nous pouvons également souligner les points suivants :

- l'objet central de la classe 2 de cette CAH est presque identique à son homologue de la classification précédente : il s'agit d'un jeudi du consommateur 73 ;
- le consommateur 68 demeure le centre de la classe 3. Ici le jour de référence n'est pas le même que pour la première CAH (un dimanche à la place d'un mardi) ;
- l'objet central de la classe 1 correspond à un consommateur qui dispose d'une télévision peu gourmande en énergie, ce type n'étant pas représenté par les objets centraux des autres classes.

Les éléments du tableau 4.17 sont relativement peu commodes pour visualiser mentalement les objets centraux de la CAH. De ce fait, nous avons représenté au moyen de la figure 4.37 les courbes de charge qui constituent le centre des classes construites par cette classification.

Une rapide étude de la figure 4.37 nous montre que :

- 80\_35\_6\_1954 est la courbe de charge d'un samedi d'une télévision de type CRT et de taille moyenne ( $55 \text{ cm} < \varnothing \leq 72 \text{ cm}$ ). La puissance de veille fluctue<sup>62</sup> entre 6 et 12 W alors que la

62. Ces variations sont vraisemblablement liées à l'échantillonnage 10 min des courbes de charge

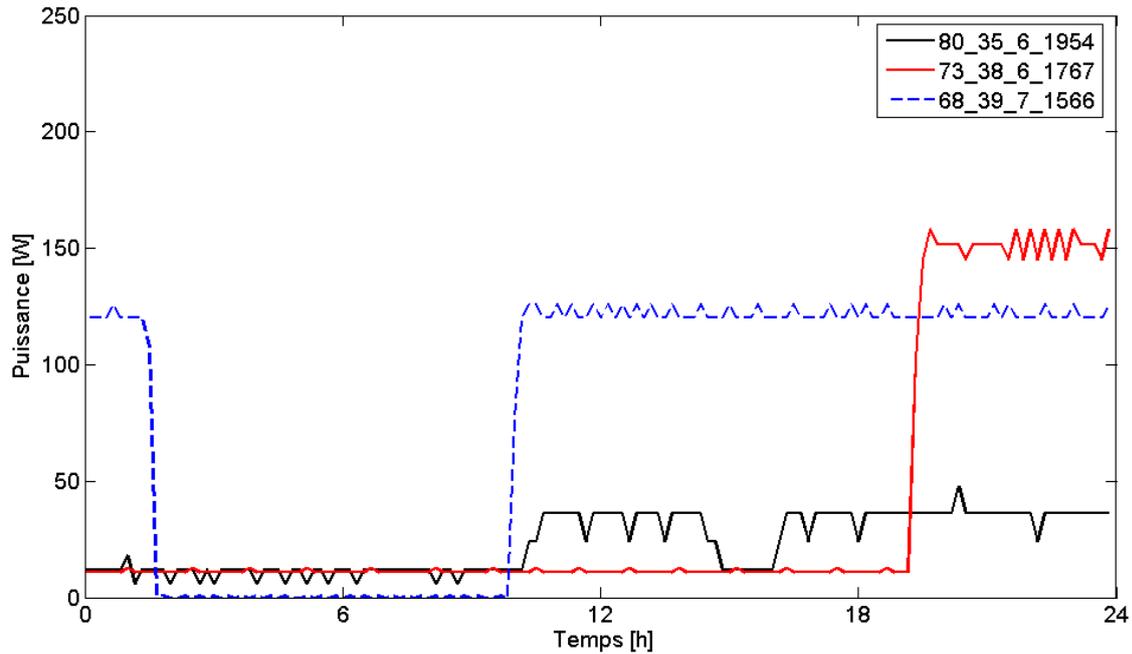


FIGURE 4.37 – Courbes de charge des objets centraux des classes de consommateurs issues de la CAH menée sur données brutes

puissance en fonctionnement oscille autour de 40 W. Deux plages de visionnement sont détectables : la première s'étend de 10 h 20 à 14 h 40 et la seconde débute à 16 h 10 et se poursuit au moins jusqu'à la fin du jour considéré ;

- 73\_38\_6\_1767 est la courbe de charge d'une télévision LCD ( $52 \text{ cm} < \varnothing \leq 81 \text{ cm}$ ) pour un samedi. La puissance de veille de ce téléviseur est de l'ordre de 10 W alors que l'appareil nécessite autour de 150 W de puissance électrique lorsqu'il est en fonctionnement. Ici, nous avons très probablement affaire à un « téléspectateur actif » puisque la période de visionnement est relativement courte durant la journée (de 19 h 20 jusqu'à la fin du jour considéré) ;
- 68\_39\_7\_1566 est la courbe de charge d'une télévision LCD de grande taille ( $\varnothing > 81 \text{ cm}$ ) pour un dimanche. Nous pouvons constater qu'en fonctionnement, cet appareil requiert de 120 à 125 W de puissance électrique alors que la veille est peu gourmande (inférieure à 0,5 W). Contrairement aux cas précédents, ce téléspectateur laisse sa télévision allumée durant toute la journée (hormis la nuit). La veille, un visionnement tardif avait eu lieu (le poste ayant été éteint à 1 h 30). L'allumage en journée se situe à 10 h du matin. Nous pouvons donc considérer qu'il s'agit d'un « téléspectateur inactif », puisqu'il semble peu probable qu'il ne fasse pas autre chose que de visionner le petit écran lorsque celui-ci est en marche.

Les objets centraux de cette classification mettent en lumière trois types de comportement :

- téléspectateur très régulier en journée et en soirée/nuit – présence d'une veille ;
- téléspectateur du soir/nuit – présence d'une veille ;
- consommateur « télévision bruit de fond » (continûment en journée et en soirée) — présence d'une veille ;

Une fois encore, l'investigation de la composition des classes doit être menée plus en avant pour savoir si le processus de classification identifie les influences que sont la durée quotidienne de visionnement du petit écran, le type de jour et le type de télévision.

En matière de durée quotidienne de fonctionnement et en utilisant le même codage que celui décrit au paragraphe 4.4.4 p. 208, nous aboutissons aux données du tableau 4.18.

Classe	Durée quotidienne moyenne par mode de fonctionnement [h]		
	Mode veille	Mode arrêt	Mode marche
1	5,66	12,84	5,50
2	7,30	10,75	5,95
3	4,08	8,67	11,25
<b>Total</b>	<b>5,73</b>	<b>11,88</b>	<b>6,39</b>

TABLEAU 4.18 – Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH menée sur les données brutes

Comme ce fut le cas de l'étude du tableau 4.13, nous ne prendrons ici en considération que la durée en mode marche de la télévision. Au travers du tableau 4.18, nous constatons donc que les classes 1 et 2 semblent regrouper des courbes de charge correspondantes à des durées de fonctionnement inférieures à la durée moyenne de l'échantillon (6,39 h) alors qu'il en est tout autrement pour la classe 3 qui représente des courbes de charge relatives à des durées quotidiennes de fonctionnement importantes (la moyenne journalière est de 11,25 h). En cela, la courbe de charge 68\_39\_7\_1566 semble donc bien être représentative du « contenu » de cette classe.

Intéressons nous maintenant à la composition des classes en matière de type de télévision. Comme précédemment, nous construisons deux tableaux : tout d'abord le tableau 4.19 indique la composition de chaque classe par type de télévision alors que le tableau 4.20 montre comment se répartissent les différentes technologies et tailles d'écran sur l'ensemble des classes.

Classe	Composition des classes par type de télévision [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1	5,1	27,2	8,2	1,3	38,2	4,1	14,3	1,6
2	0,0	0,0	1,6	0,0	8,3	13,8	65,3	11,0
3	0,0	0,2	12,5	0,0	8,7	5,2	64,3	9,1
<b>Total</b>	<b>3,5</b>	<b>18,5</b>	<b>7,6</b>	<b>1,0</b>	<b>28,7</b>	<b>6,0</b>	<b>30,4</b>	<b>4,3</b>

TABLEAU 4.19 – Composition des classes par type de télévision de la CAH menée sur les données brutes

Le tableau 4.19 nous montre que la première classe est principalement constituée des télévisions de petite et de moyenne taille (les types 34 et 37 étant notamment exclus de la composition des autres classes). Les deux autres classes sont construites principalement à partir de télévisions gourmandes en électricité, en particulier la classe 2 qui est composée à 98,4 % de téléviseurs des types 38 à 42. La classe 3 est davantage éclectique dans la mesure où six types de télévision sont représentés avec une part (12,5 %) non négligeable de téléviseurs CRT de grande taille (type 36).

En plus de ces constatations, le tableau 4.20 nous confirme tout d'abord qu'au minimum 90 % des télévisions de petite et moyenne taille se concentrent dans la classe 1. D'autre part et c'est une surprise, les téléviseurs de type 40 se répartissent presque homogènement entre les trois classes. En revanche, la répartition des appareils de type 39 et 42 est plus sensiblement orientée vers une classe précise (respectivement les classes 1 et 2).

Nous avons ensuite examiné les résultats de la classification en matière de type de jour. Cette répartition est transcrite au travers du tableau 4.21.

L'étude du tableau 4.21 nous indique que la répartition des types de jour au sein de la classe 1 est relativement uniforme. La seconde catégorie semble privilégier les jours du début de la semaine (lundi,

63. Les nombres entre parenthèses indiquent la proportion respective en pourcentage de chaque classe dans l'effectif total des courbes de charge

Classe <sup>63</sup>	Répartition des types de télévision au sein des classes [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1 (68,1)	100,0	99,9	73,2	100,0	90,6	47,0	32,0	24,7
2 (17,7)	0,0	0,0	3,6	0,0	5,1	40,7	38,0	45,3
3 (14,2)	0,0	0,1	23,2	0,0	4,3	12,3	30,1	30,0

TABEAU 4.20 – Répartition des types de télévision au sein des classes de la CAH menée sur les données brutes

Classe	Composition des classes par type de jour [%]						
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
1	14,3	14,3	13,9	14,0	14,8	14,2	14,4
2	15,0	15,3	15,2	14,2	14,0	12,9	13,3
3	13,7	12,5	13,5	13,2	11,2	17,5	18,4
<b>Total</b>	<b>14,4</b>	<b>14,2</b>	<b>14,1</b>	<b>13,9</b>	<b>14,2</b>	<b>14,4</b>	<b>14,8</b>

TABEAU 4.21 – Composition des classes par type de jour de la CAH menée sur les données brutes

mardi et mercredi) et délaisse légèrement les samedis et dimanches. À l'inverse, ces deux derniers jours se retrouvent prioritairement dans la dernière classe.

En plus de ces éléments, l'algorithme de classification calcule le barycentre de chacune des classes formées. Ce barycentre peut être considéré comme une courbe de charge « moyenne » représentative des éléments d'une classe. **Cette courbe moyenne calculée ne doit cependant pas être confondue avec les objets centraux des classes** qui eux correspondent à de véritables courbes de charge de l'échantillon initial. Les barycentres des trois classes de la CAH menée sur les données brutes sont représentés sur la figure 4.38.

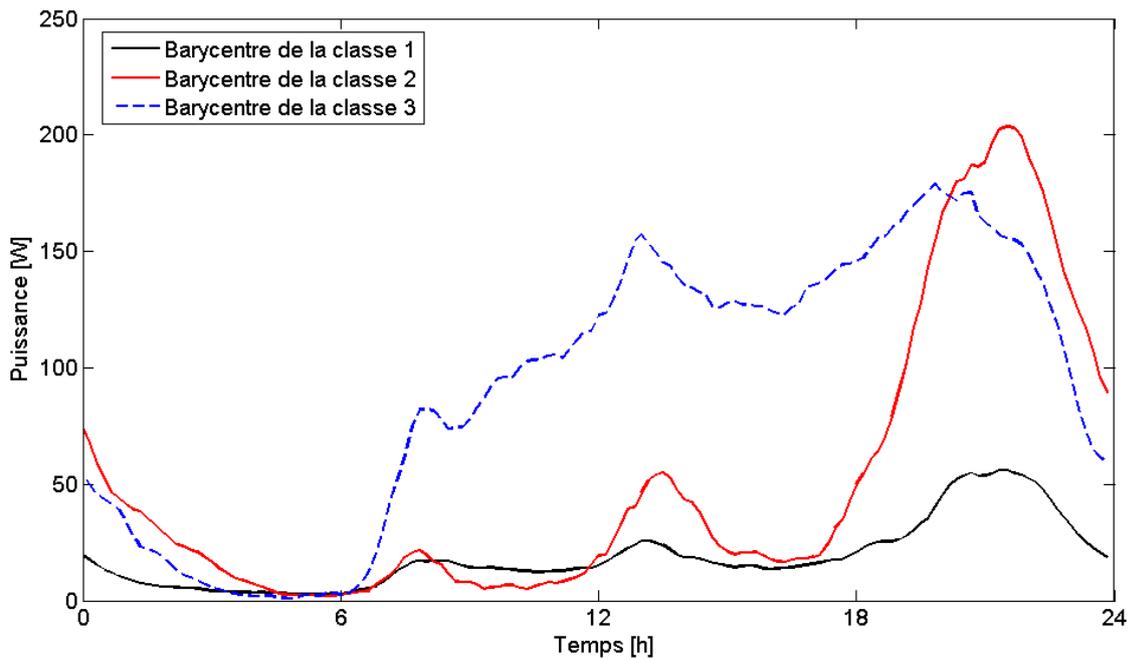


FIGURE 4.38 – Barycentres des classes de la CAH menée sur les données brutes

Bien qu'il ne s'agisse pas de courbes de charge réelles constituant l'échantillon, l'étude des barycentres des classes nous semble essentielle<sup>64</sup>. En effet, le barycentre d'une classe se construit par une somme pondérée de l'ensemble des courbes de charge qui la constitue. En matière de représentativité du comportement moyen d'une classe, cette construction est de la sorte plus pertinente que ne l'est l'objet central de la classe. Enfin, le barycentre s'apparente à l'intermédiaire de la distribution de probabilité de l'usage de la télévision au sein d'une classe.

La figure 4.38 indique que les courbes de charge barycentriques des classes 1 et 2 sont assez similaires en dehors de leur différence de niveau de puissance. La courbe de charge barycentrique de la classe 3 quant à elle suggère un usage diurne intensif de la télévision. Ces différences sont cohérentes avec nos conclusions précédentes en matière de composition des classes (type de jour et type de télévision).

Dans tous les cas, les barycentres témoignent des variations d'utilisation de la télévision à l'échelle journalière. Les habitudes d'usage de cet appareil sont ici clairement identifiées (visionnement du petit écran autour de l'heure des repas et en soirée).

Pour résumer, la CAH conduite sur les données brutes donne lieu à trois classes qui peuvent être résumées comme suit :

- la première classe est représentative des télévisions petites et moyennes de l'échantillon. Celles-ci sont utilisées de manière modérée sans que le jour de la semaine soit une influence notable ;
- la seconde classe comporte des télévisions de grande taille et de technologie récente qui fonctionnent plus longuement dans la journée que la moyenne de l'échantillon et en particulier au cours de la première moitié de la semaine ;
- la troisième et dernière classe est représentative de l'usage intensif de télévisions de taille moyenne à grande, tous types de technologies confondus. Ces constats s'expliquent notamment du fait que les jours de week-end sont sur-représentés dans cette classe.

Malgré la différence en termes de nombre de classes obtenues, l'ACP préalable ne semble pas indispensable ou déterminante pour la constitution de classes de courbes de charge. D'autres méthodes d'agrégation existent et auraient pu être utilisées à la place de la méthode de Ward. Une meilleure connaissance de chacune d'elles aurait pu conduire à un choix éventuellement plus judicieux. Une autre possibilité pour essayer d'améliorer ces résultats est d'imposer un nombre prédéfini de classes à obtenir (la connaissance du secteur peut orienter ou faciliter son choix). Par exemple, nous pourrions fixer à sept le nombre de classes souhaitées de manière à identifier le caractère discriminant ou non sur la classification du type de jour.

Jusqu'à présent, nous avons pu constater que l'algorithme hiérarchique de classification distribue les courbes de charge dans des classes principalement en fonction du niveau des appels de puissance occasionnés par les appareils (il semble d'ailleurs que ce soit le paramètre le plus discriminant) et par la durée d'utilisation journalière de la télévision.

De manière à supprimer de façon théorique l'influence sur la classification du type de télévision, nous avons décidé de conduire les deux mêmes analyses (ACP puis CAH et CAH seule) sur des courbes de charge normées, i. e. les courbes de charge originales réduites par la puissance maximale journalière appelée par le téléviseur. De la sorte, ce sont des **profils de visionnement\*** (évolution journalière de la puissance réduite appelée par la télévision) exprimés sans unité qui seront traités par le processus de classification. Comme nous l'avons mentionné au début de cette partie, les études menées sur les profils

64. Celle-ci n'a toutefois de sens qu'à la condition d'opérer la classification directement sur des données en puissance, plus généralement sur un tableau observations  $\times$  individus parfaitement connu. C'est d'ailleurs pour cette raison que nous ne considérons les barycentres qu'à partir de cette seconde classification : la première que nous avons menée traite en effet des coordonnées des observations restituées dans la nouvelle base induite par l'ACP – cf. paragraphe 4.4.4 page 204

de visionnement sont consultables dans l'annexe F. Nous proposons dans le paragraphe 4.4.6 d'en résumer les conclusions générales.

#### 4.4.6 Résumé des études supplémentaires de classification sur des courbes de charge normées de télévision

Dans ce paragraphe, nous avons extrait les principaux résultats des études détaillées dans l'annexe F sans les expliciter davantage. Nous invitons donc le lecteur curieux d'en savoir plus à se référer à cet annexe. Toutefois, nous mettons ici en perspective les résultats de ces études supplémentaires par rapport aux conclusions des travaux de classification analysés dans les sections précédentes.

##### ACP sur données normées et CAH sur les résultats de l'ACP

Comme pour l'ACP sur les données brutes, les cinq premiers axes issus de l'ACP sur les profils de visionnement n'explicitent que 56 % de l'inertie du nuage. De plus, les axes factoriels F1, F2 et F3 de cette seconde ACP sont chacun construits par une série continue de points 10 min (ce qui diffère de la première ACP). Chacune de ces séries correspond ainsi à une période infra-journalière parfaitement définie. Hasard de l'analyse ou non, les périodes identifiées s'enchaînent de manière chronologique conformément au rang des axes factoriels. En définitive, il semble que des probabilités différentes d'usage de la télévision au sein des trois périodes aient ainsi été révélées par l'ACP.

La CAH réalisée sur les coordonnées des résultats de l'ACP dans la nouvelle base a abouti à quatre classes. L'étude de la composition de ces dernières suivant la durée journalière de marche de la télévision, les types d'appareil et les jours de la semaine aboutit aux constatations ci-dessous :

- la première classe rassemble des téléspectateurs assez modérés en matière de durée journalière de visionnement. Ceux-ci disposent d'une télévision petite à moyenne et la regardent préférentiellement lors de jours de semaine « similaires » ;
- la seconde classe regroupe des téléspectateurs assidus, disposant d'un poste de grande taille et qui s'en servent principalement lors de journées potentielles de repos (mercredi et dimanche) ;
- la troisième classe semble être également représentative de téléspectateurs assidus, possédant un appareil de technologie récente et largement énergivore et qui l'utilisent particulièrement le week-end ;
- la dernière classe témoigne d'un usage intensif du petit écran, pour des télévisions de technologie plutôt ancienne et dont l'utilisation se concentre en fin de semaine.

Les résultats des deux ACP suivies de CAH se différencient sur plusieurs points :

- le nombre de classes identifiées : 5 pour l'analyse sur données brutes et seulement 4 pour celle menée sur données normées ;
- la capacité de regroupement des téléspectateurs au comportement journalier similaire en matière de durée de visionnement du petit écran : elle semble meilleure pour le processus mené sur données normées car il différencie nettement les comportements extrêmes et moyens ;
- le pouvoir d'identifier les consommations de veille : il est bien plus important dans le cas du processus conduit sur les profils de visionnement ;
- la faculté de distinguer le type de télévision : comme il était prévisible, cette capacité est supérieure pour le processus mené sur données brutes ;
- l'aptitude à différencier les jours de la semaine : une fois encore, celle-ci paraît légèrement plus prononcée pour le processus opéré sur les courbes de charge brutes.

### CAH libre sur données normées

La CAH libre (sans fixation préalable du nombre de classes souhaitées) menée sur les données normées met en exergue trois classes de comportement de l'utilisation de la télévision :

- la classe 1 regroupe des téléspectateurs sobres, possédant une télévision « moyenne » (en matière de taille de diagonale) et relativement indifférenciée (en type de technologie). Ils l'utilisent indistinctement quel que soit le jour de la semaine ;
- la classe 2 rassemble des individus assidus face à leur poste de télévision. Celui-ci est de taille moyenne à grande et son usage se concentre sur des jours potentiels de repos (mercredi, samedi et dimanche) ;
- enfin la classe 3 semble représentative des comportements extrêmes de consommateurs. Ceux-ci disposent d'une télévision plutôt récente et de taille moyenne à grande et privilégient la fin de semaine pour en faire l'utilisation.

De plus, les classes de cette CAH semblent être parfaitement bien résumées par leur barycentre respectif.

### CAH contrainte sur données normées

Les 7 classes d'utilisation du petit écran construites par la CAH contrainte (i. e. avec un nombre prédéfini de classes à obtenir) se caractérisent de la manière suivante :

- la première classe regroupe des téléspectateurs très modérés, qui disposent d'une télévision petite à moyenne utilisée préférentiellement lors des quatre premiers jours de la semaine ;
- la seconde classe rassemble des utilisateurs raisonnables du petit écran qui possèdent soit une petite télévision de technologie ancienne ou bien une télévision récente et de taille moyenne à grande. Celle-ci est regardée plutôt en début de semaine ;
- la classe 3 semble représentative de téléspectateurs très modérés. Ils sont équipés d'une télévision petite à moyenne et ils l'utilisent davantage au cours de la seconde moitié de la semaine ;
- la quatrième classe réunit des téléspectateurs assidus, ayant une télévision plutôt récente et de taille moyenne à grande et ils la regardent principalement le week-end ;
- la classe 5 regroupe des téléspectateurs faisant un usage intensif de la télévision. Celle-ci est plutôt énergivore (technologie récente et de grande taille) et elle est utilisée prioritairement lors de journées potentielles de repos (mercredi, samedi et dimanche) ;
- la sixième classe rassemble des utilisateurs assidus du petit écran. Dans cette catégorie, il s'agit principalement de télévisions de grande taille et gourmandes en énergie qui sont regardées de manière prépondérante en tout début et en fin de semaine (lundi, samedi et dimanche) ;
- enfin il n'est pas nécessaire de décrire la dernière classe puisqu'elle correspond exactement à la classe 3 de la CAH libre.

Comme pour la précédente classification, le tracé des barycentres des classes révèle une correspondance notable avec la description que nous venons d'en faire. Toutefois cette concordance est plus relative que pour la précédente classification du fait du plus grand nombre de classes.

Comparer les résultats de la CAH sur données brutes, avec ceux de la CAH libre et de la CAH contrainte (toutes deux opérées sur données normées) nous amène aux constatations suivantes :

- la comparaison des résultats est directe et semble plus pertinente<sup>65</sup> entre CAH sur données brutes et CAH libre sur données normées (plutôt qu'entre l'étude sur données brutes et l'analyse contrainte sur données normées) du fait du même nombre de classes auquel ces deux analyses aboutissent ;
- la segmentation du comportement d'utilisation de la télévision (en fonction de la durée journalière de visionnement) paraît mieux réalisée par les classifications sur données normées ;
- les deux études sur données normées identifient une même classe représentative des profils de veille. Celle-ci n'émerge pas de l'analyse sur données brutes ;
- l'étude sur données brutes a naturellement la plus grande propension à identifier l'influence du type de télévision et à l'intégrer dans la composition des classes ;
- dans chacun des cas, plusieurs classes semblent privilégier un à plusieurs type(s) de jour au détriment d'un ou plusieurs autre(s). Néanmoins, aucune véritable segmentation entre les classes par type de jour ne semble avoir abouti.

#### 4.4.7 Bilan du travail de classification et perspectives

À partir d'une base de données de courbes de charge de télévisions, nous avons cherché à sélectionner un petit nombre d'entre elles au moyen de méthodes d'analyse factorielle. Les courbes de charge extraites sont voulues représentatives et explicatives de l'ensemble de l'échantillon dont elles sont issues. Pour ce travail, nous avons fait l'usage de l'ACP et de la CAH avec méthode d'agrégation de Ward et métrique euclidienne. Les conclusions que nous pouvons tirer de l'ensemble des études conduites sont les suivantes :

- à notre avis, l'ACP apporte peu d'informations. Les axes principaux sont formés par une (ou des) série(s) de points 10 min représentant des périodes infra-journalières. Dans le cas de l'ACP sur les données brutes, le jour et la nuit sont ainsi distingués alors que dans l'analyse des données normées, l'ACP identifie trois périodes successives qui semblent correspondre à des intervalles dans lesquels diffère la probabilité d'usage de la télévision. Toutefois, les cercles de corrélation montrent dans les deux cas une structure « forte » entre les points 10 min, structure que nous semblons peiner à valoriser dans la classification ;
- pour les deux ACP, les cinq premiers axes factoriels retenus n'expliquent qu'environ 56 % de l'inertie du nuage ce qui est peu par rapport à d'autres études d'analyse factorielle (en effet, dans certains cas<sup>66</sup> trois axes suffisent pour expliquer 80-90 % de l'inertie du nuage de points étudié – cf. [SAP06] p. 210). Pour améliorer ce pouvoir explicatif, il est nécessaire d'introduire un maximum de variables dans l'analyse, Autrement dit, il s'agirait de reconduire les ACP en incluant toutes les données descriptives des ménages de la base REMODECE. Il se peut que les variables socio-démographiques (taille des foyers, types de ménages : famille, couple actif ou retraité. . .) aient un pouvoir explicatif sur le comportement d'usage de la télévision qui soit supérieur à celui des 144 points 10 min constitutifs des courbes de charge ;
- d'un autre côté, il est envisageable d'approfondir la caractérisation des classes. Par exemple, il serait sans doute intéressant de construire la variable donnant la consommation moyenne journalière correspondante à chaque courbe de charge. Une autre variable pourrait être celle d'une consommation moyenne journalière relative à la pointe de puissance de l'appareil pour le jour donné. Cette dernière construction témoignerait de la durée journalière de visionnement d'un poste de télévision, ceci de façon probablement plus vraisemblable que la valeur du mode marche issue du codage que nous avons décrit page 208 et mis en œuvre dans notre travail ;

65. La CAH contrainte sur données normées mériterait son équivalente sur données brutes. En effet, les résultats de la dernière classification que nous avons opérée « héritent » du caractère non abouti du processus qui les a engendrés. Du fait de la contrainte fixée et de l'importance relative du nombre de classes induites, les conclusions de cette analyse sont notamment moins tranchées que pour les autres études

66. Déclaration de M. Pascal Lemasson, professeur de statistiques à Mines ParisTech, lors d'un entretien à Paris le 16 octobre 2012

- sur les données brutes, ACP puis CAH ou CAH seule classent les courbes de charge prioritairement selon l'amplitude de la puissance appelée et donc suivant le type de télévision. Les influences de la durée journalière d'utilisation des téléviseurs et du type de jour ne semblent pas être identifiées et valorisées pour l'élaboration des classes. Les résultats relatifs à la statistique de visionnement journalier sont intimement liés au choix du codage des modes de fonctionnement (cf. p. 208). Ils doivent donc être corroborés ou infirmés par une analyse plus fine notamment à l'aide de tests statistiques ;
- le codage des modes de fonctionnement dont il a été question précédemment pourrait être amélioré de sorte de filtrer les courbes de charge journalières qui ne correspondent, le cas échéant, qu'à de l'unique consommation de veille des appareils. Ces courbes remarquables d'appels de puissance paraissent entacher les résultats des analyses sur données brutes mais pas ceux des études sur données normées ;
- sur les données normées en effet, les procédures d'analyse factorielle ont tout d'abord l'air de discerner très correctement les courbes de charge journalières correspondantes uniquement à de la consommation de veille. De plus, elles paraissent mieux identifier les comportements d'utilisation de la télévision – vis-à-vis des études sur données brutes – par l'intermédiaire de la durée de visionnement journalier du petit écran. En revanche, elles sont moins aptes que ces dernières à tenir compte de l'influence du type de télévision pour la formation des groupes. De plus et comme pour les analyses sur données brutes, le type de jour paraît être un élément assez peu déterminant pour la constitution des classes (y compris lorsque 7 classes sont exigées) ;
- nous n'obtenons pas une réelle caractérisation du comportement d'utilisation de la télévision tel que décrit au début du paragraphe 4.4.1 ;
- l'échantillon des foyers instrumentés n'étant pas représentatif de la population française totale, cette étude n'est donc relative qu'à l'échantillon considéré. Plus particulièrement dans la campagne REMODECE, ce dernier sur-représente les personnes âgées, susceptibles donc d'avoir des comportements atypiques de téléspectateurs (la très faible différenciation des courbes de charge par type de jour trouve vraisemblablement son origine dans cette propriété de l'échantillon). Les conclusions que nous avons pu tirer nous semblent donc très difficilement extrapolables à une plus large échelle.

Ces différents points nous poussent également à croire que la métrique euclidienne n'est pas nécessairement la plus adaptée pour la stratégie d'agrégation du processus de classification. Selon nous et comme nous l'avons suggéré au préalable, la métrique « idéale » pour l'étude de courbes de charge doit à la fois tenir compte de la différence d'amplitude et de déphasage pouvant exister entre elles. C'est dans ce contexte que nous avons mené des recherches et que nous avons envisagé le Dynamic Time Warping (voir partie 4.1.3) comme solution possible.

Nous avons donc adapté l'algorithme de classification de sorte que la métrique utilisée soit la distance *DTW*. Par la suite, nous avons conduit des premiers travaux de classification sur un échantillon réduit de courbes de charge de télévision (40 profils d'appels de puissance) de manière à tester la faisabilité de l'opération. En définitive, il s'avère que cette classification avec *DTW* est beaucoup plus chronophage que ne l'est celle avec distance euclidienne. Trois raisons principales expliquent ce constat : tout d'abord, calculer la distance euclidienne entre deux courbes de charge formées de  $n$  points chacune est une opération d'une complexité d'ordre  $n$  ( $O(n)$ ) alors qu'estimer la distance *DTW* entre ces deux mêmes courbes est d'ordre  $n^2$  ( $O(n^2)$ ). Ensuite, l'intégration de chacune de ces distances dans la procédure de classification amplifie encore cette différence d'ordre de grandeur. En effet, il s'agit alors de calculer pour chaque itération toutes les distances entre les objets à classer – des courbes de charge dans notre cas – pris deux à deux. Enfin et de manière moins déterminante, notre adaptation de l'algorithme classique n'est pas optimisée pour la rapidité des calculs.

Nous ne détaillerons pas dans le présent manuscrit les résultats de cette classification de courbes de charge avec métrique *DTW*. Du fait de la taille de l'échantillon de courbes de charge sélectionné pour ce test (et a fortiori des effectifs des classes obtenues), les conclusions ne sont pas pertinentes.

## Conclusion du chapitre

Ce chapitre synthétise la partie expérimentale de nos recherches. Celle-ci a eu pour but principal d'améliorer nos connaissances sur la courbe de charge. Des études ont été menées aussi bien sur des données réelles issues d'une campagne de mesures que sur des courbes de charge simulées par notre modèle.

Pour parvenir à cet objectif de connaissance, nous avons tout d'abord fait une synthèse des éléments de mesure de la proximité entre courbes de charge. Passer en revue les indicateurs usuels employés dans ce domaine nous a indiqué que l'information qu'ils délivrent individuellement est souvent trop synthétique. De ce fait, nous nous sommes intéressés à la recherche d'une distance adéquate pour la comparaison de courbes de charge. Nous avons ainsi identifié l'algorithme du Dynamic Time Warping comme méthode potentiellement valorisable pour l'étude des courbes d'appels de puissance.

En vue d'observer le foisonnement existant entre les courbes de charge obtenues à partir de simulations de notre modèle, nous avons adapté et mis en œuvre la méthode *DTW* sur des courbes de charge moyennes foisonnées. Nous aboutissons ainsi à des représentations visuelles de la proximité entre courbes d'appels de puissance et nous sommes capables grâce à elles, de décomposer les différences observées entre les courbes. À notre connaissance, ces méthodes sont originales. Toutefois, elles doivent être renforcées par un travail approfondi d'analyse statistique, en particulier en ce qui concerne la détermination du taux de sondage de courbes de charge. Ce dernier est d'une importance de tout premier plan puisqu'il correspond à la suppression du déphasage d'origine temporelle pouvant exister entre courbes de charge.

Le reste de ce chapitre a traité de l'étude de courbes de charge issues d'une campagne de mesures. À l'aide de méthodes d'analyse factorielle, nous avons cherché à identifier des comportements types d'usage d'un appareil (la télévision dans notre cas) à partir d'appels de puissance journaliers. Pour cela, nous avons effectué différents travaux mettant en jeu l'analyse en composantes principales et la classification ascendante hiérarchique. Nous avons analysé la composition des groupes formés en regard de l'ensemble des informations disponibles.

Cette étude révèle des comportements très distincts pour cet appareil mais elle n'a pas pu aboutir à une caractérisation précise de ceux-ci. Néanmoins, des voies d'amélioration ont été identifiées : d'une part, il s'agirait d'améliorer la méthodologie statistique employée dans ce travail et d'autre part, il serait intéressant de poursuivre les travaux de classification que nous avons initiés avec la distance *DTW* ou bien avec l'une de ses variantes.

# Conclusion générale et perspectives

*« Rien n'est fait tant qu'il reste quelque chose à faire. »*

Romain Rolland

Arrivés au terme de ces trois années de recherches, nous dressons un bilan synthétique de leurs enseignements. L'issue de la thèse constitue également l'occasion de mettre en exergue les limites identifiées de notre travail et nous suggérons des pistes de recherche pour les repousser. Enfin, nous élargissons le spectre en pensant quels peuvent être les prolongements à donner à cette thèse.

## Synthèse des enseignements de nos recherches

La reconstitution de la courbe de charge électrique résidentielle fut l'objectif ambitieux de ce doctorat. Ambitieux il l'est à plusieurs titres puisque ce travail soulève des problèmes multi-disciplinaires complexes. Parmi eux, nous avons focalisé nos recherches sur quelques points :

- l'élaboration d'une méthode conceptuelle inédite pour la synthèse des appels domestiques de puissance ;
- la définition et la mise en œuvre d'une technique de génération de diversité d'origine comportementale qui explicite le foisonnement impactant les courbes de charge ;
- l'approfondissement des connaissances relatives à la courbe de charge des usages domestiques aussi bien du point de vue unitaire qu'agrégé.

L'étude bibliographique que nous avons entreprise au commencement de la thèse nous a permis de **capitaliser des connaissances issues des méthodes de la littérature** capables de générer des courbes de charge des usages domestiques de l'électricité. Nous avons cherché à comparer les modèles identifiés et pour cela, nous avons sélectionné trois critères d'intérêt : le champ d'application du modèle, sa résolution temporelle et la technique de génération de diversité. Positionner les méthodes nous a permis ainsi de constater un **manque de modèles capables de générer la courbe de tous types d'usage domestique, à un pas de temps de résolution fin et en explicitant le foisonnement**. Indirectement cet état de l'art a laissé transparaître les difficultés d'établir une méthode de reconstitution des appels de puissance conforme à notre besoin.

Partant d'une page blanche et tout en ayant à l'esprit les spécifications précédemment mentionnées, une phase conséquente de réflexion, conduite sous forme d'analyse fonctionnelle, a donné lieu à ce qui apparaît pour nous comme la principale originalité et l'avancée majeure de notre travail. Effectivement, **la méthode de reconstitution de courbe de charge que nous proposons est une base conceptuelle a priori universelle** : d'une part, elle est utilisable quel que soit le périmètre étudié (de l'échelle unitaire au niveau national sans être spécifiquement conçue pour la situation d'un parc habité déterminé). D'autre part, la méthode que nous avons construite peut donner lieu à différentes adaptations c'est-à-dire à des modèles aux paradigmes et aux finalités potentiellement divers. Ainsi, nous pourrions imaginer des mises en œuvre valorisant les **Systèmes Multi-Agents (SMA)\*** ou bien encore des adaptations de la méthode pour des régions très spécifiques du globe (les îles isolées par exemple).

À la suite de ces développements conceptuels, nous avons cherché à mettre en œuvre notre méthode. Plus particulièrement, nous nous sommes concentrés sur la partie de la chaîne de simulation relative à la génération de diversité comportementale. Pour élaborer cette dernière, nous avons conçu un algorithme

adapté à la méthode et nous avons modélisé un certain nombre des concepts définis. L'explicitation du foisonnement, qui synthétise entre autres la pluralité des comportements des utilisateurs domestiques d'électricité, fut au cœur de notre travail de recherche. **L'algorithme que nous avons élaboré se place dans cette perspective en cherchant à reproduire de manière réaliste des scénarios journaliers d'usage des appareils domestiques.**

Enfin, le volet expérimental n'a pas été exclu de cette thèse. Pratiquement, une meilleure connaissance de la courbe de charge des usages domestiques et des comportements qui y sont liés est une condition *sine qua none* pour mieux penser les appels de puissance du secteur résidentiel dans une démarche de reconstitution. C'est également à partir de cette connaissance que l'étude des appels de puissance, qu'ils soient issus de mesures ou de simulations de notre modèle, s'avère encore plus riche d'enseignements. En outre, ce travail nous a conduit à une méthode d'observation et de caractérisation du foisonnement affectant les appels de puissance à un niveau agrégé. **Cette méthode s'avère indispensable pour atteindre l'objectif de reconstitution bottom-up de la courbe de charge résidentielle.**

### Critiques et perspectives de travail

*Quid de l'évaluation et de la validation du modèle de courbe de charge ?*

Deux étapes majeures, vis-à-vis du modèle élaboré, n'ont pas pu être conduites au cours de nos travaux de recherche. Plus particulièrement, il s'agit de son **évaluation** et de sa **validation**.

La réalisation du premier point implique des **études de sensibilité** des paramètres du modèle. De cette manière, l'influence de chacun d'eux pourrait être mieux appréciée. En particulier, il s'agirait d'établir une hiérarchie des déterminants de la forme, de l'amplitude et du niveau de foisonnement des courbes de charge générées par notre modèle. À l'issue de la thèse, seule la variation d'un paramètre (passage d'un cycle unitaire de puissance à un autre) a pu être testée.

La validation du modèle de reconstitution de courbe de charge est un travail qui va bien au delà de la « simple » comparaison, pour quelques cas d'application, entre données réelles et données simulées. Il s'agit là en effet de déterminer la **pertinence** et la **vraisemblance** des résultats du modèle et simultanément de **calibrer** les valeurs des paramètres utilisées dans les simulations.

Valider le modèle de courbe de charge implique la nécessaire disponibilité de données de consommation d'électricité qui soient à la fois diversifiées, riches et si possible représentatives de la population du territoire étudié. En cela, la capitalisation d'informations issues de campagnes de mesures, de panels ou bien directement de relevés d'électricité à grande échelle au niveau du compteur dans un contexte « Smart Grid <sup>67</sup> » constitue une étape préalable indispensable.

*Est-il possible de prouver formellement notre algorithme ?*

Suite à la modélisation des concepts, leur mise en œuvre sous forme algorithmique a suscité différents questionnements du type :

- pour un cas donné de placement d'un équipement, les règles de l'algorithme mises en jeu peuvent-elles toutes s'appliquer ou y a-t-il un possible risque de contradictions entre elles ?
- est ce que l'ajout de règles et de contraintes rend l'algorithme plus efficace ou au contraire cela induit-il des problèmes de solvabilité ?
- l'algorithme aboutit-il systématiquement à une solution cohérente ?

Les différentes interrogations ci-dessus nous ont orienté vers la thématique de la preuve formelle. N'étant pas familiers de cette discipline, nous avons tout d'abord consulté quelques ouvrages qui y sont dédiés [DUM79, AECL90, CGP99] et nous avons ensuite pris contact avec M. Pierre Jouvelot (Maître de recherche au Centre de Recherche en Informatique de Mines ParisTech) à qui nous avons exposé notre problématique. Il est ressorti de cet entretien <sup>68</sup> que **prouver formellement l'algorithme n'a pas**

67. Réseau intelligent d'électricité

68. Celui-ci a eu lieu le 9 mai 2012 à Paris

**de sens.** En effet, il n'y a pas de théorèmes ou de propositions pour lesquels une preuve pourrait être cherchée puisque l'algorithme est hautement non déterministe. Cette dernière caractéristique provient de la finalité même de l'algorithme qui est de reproduire des comportements humains qui ne peuvent pas être formalisés de par leur complexité.

En revanche, cet entretien nous a confirmé le **besoin d'une validation « métier »** de notre modèle de reconstitution de courbe de charge, celle-ci pouvant à la fois être conduite sur des données mesurées en puissance ainsi que sur des indicateurs du comportement à l'échelle unitaire comme macroscopique. Par exemple, nous pourrions vérifier la cohérence des scénarios d'usage des appareils dans chaque foyer modélisé. En outre, il serait possible de tester la conformité des résultats de l'algorithme à un niveau agrégé par rapport à une fréquence hebdomadaire réelle d'utilisation d'un appareil qui aurait été modélisé.

*Qu'en est-il des limites identifiées ou présumées de notre méthode de reconstitution des appels de puissance et de sa mise en œuvre ? Quelles sont les solutions envisageables pour les repousser ?*

Dans la méthode conceptuelle que nous avons élaborée, nous avons cherché à réduire autant que possible la complexité d'un parc à simuler. Néanmoins une représentation réaliste en puissance d'un ménage à l'échelle unitaire implique un grand nombre de variables ce qui peut **rendre la combinatoire, in fine le nombre de simulations nécessaires, assez conséquente** dès lors que l'on se place à une échelle agrégée d'un parc à modéliser. Vu le temps dont nous avons eu besoin pour réaliser la simulation d'un scénario simple, **il nous semble nécessaire d'améliorer notre implémentation dans un but d'optimiser le temps de calcul voire de repenser au choix de l'environnement de modélisation** qui a indéniablement réduit la quantité de résultats à laquelle nous aurions pu potentiellement aboutir.

L'implémentation d'une partie des concepts de la méthode de reconstitution de courbe de charge a soulevé un certain nombre de problématiques plus ou moins complexes. Certaines d'entre elles ont pu être solutionnées dans le cadre de la thèse, d'autres au contraire posent encore potentiellement ou effectivement des difficultés. Nous nous proposons ici de discuter de quelques-unes d'entre elles.

Tout d'abord comme nous l'avons fait remarquer dans le manuscrit, une des difficultés de la modélisation des appels de puissance est d'en assurer la continuité lors de la simulation de plusieurs jours calendaires consécutifs. Conscients de cette problématique et parce que nous avons centré la reconstitution des appels de puissance sur l'individu, nous avons adopté une approche « décalée » aussi bien conceptuellement que sur le plan pratique. De par sa construction, la continuité de la courbe de charge est assurée au passage de deux journées successives du calendrier. Néanmoins, c'est la concordance en matière de comportement humain à l'échelle pluri-journalière (représentation du rythme hebdomadaire par exemple) qui demeure à assurer par un processus à mémoire venant se superposer à notre approche (il s'agit notamment de transmettre des informations entre les jours-types simulés qui se suivent). Ce dernier processus n'a pas été pensé dans la thèse et demeure à élaborer.

D'autre part, lorsque nous avons détaillé dans le troisième chapitre les éléments permettant aux équipements de tenir compte les uns des autres pour rendre plus réaliste la logique de déclenchement des appareils, nous avons introduit la notion de résumé d'intervalle. Cette construction capitalise pour chaque intervalle infra-journalier les événements qui ont eu lieu en son sein, autrement dit l'information nécessaire (aux paramètres de fonctionnement près) pour reconstituer la courbe de charge électrique journalière. Dans notre cas, les événements s'inscrivent correctement dans les résumés d'intervalles mais il n'a pas été possible d'obtenir leur capitalisation au fur et à mesure d'une simulation. Il s'agit ici d'un « pur » problème informatique et plus précisément de séquençement des ordres de lecture et d'écriture dans des fichiers de données associés aux modèles d'équipements domestiques.

*Quelles avancées pour la comparaison de courbes de charge ?*

Dans notre travail, l'algorithme du Dynamic Time Warping a été adapté à des courbes de charge agrégées aussi bien dans un but de visualisation que d'explicitation du foisonnement qui les affecte. Les résultats

que nous avons présentés dans le manuscrit constituent déjà une certaine avancée dans ce domaine. Néanmoins des recherches supplémentaires en ce sens et une analyse statistique plus approfondie permettraient à notre avis de renforcer et d'étoffer les éléments auxquels nous avons abouti à l'issue de la thèse. **Un indicateur de convergence entre courbes de charge, par atteinte du foisonnement, est un débouché possible de ces travaux annexes.** En outre, la distance *DTW* pourrait être valorisée de manière à identifier le(s) taux de sondage dans une population de sorte d'aboutir à une représentation satisfaisante en puissance de celle-ci tout en optimisant le nombre de simulations nécessaires.

*Quelles sont les suites à donner à cette thèse ?*

Poursuivre le travail de reconstitution de la courbe de charge résidentielle implique tout d'abord, l'intégration dans la modélisation d'un plus grand nombre et d'une plus large variété d'équipements domestiques d'électricité spécifique. Cela nécessite notamment une meilleure connaissance de chacun d'eux et l'élaboration d'une bibliothèque de cycles unitaires de puissance. À cela s'ajoute le besoin de représenter et de simuler une variété plus abondante de scénarios de vie des ménages.

Ensuite, il s'agirait de prendre en compte la dimension du bâtiment et des usages thermiques associés. Ceci passe nécessairement par le couplage des modèles et une réflexion sur la synchronisation des simulations doit être menée.

D'autre part, il faudrait enrichir l'algorithme de manière à ce qu'il représente les périodicités journalières, hebdomadaires et mensuelles en matière de comportement des ménages. Concrètement, c'est l'ensemble de la problématique de la distribution des tâches journalières en fonction du type de jour, du ménage et, le cas échéant, des données réalisées (issues de simulation de jours antérieurs) qui doit être traitée. Un processus à mémoire prendrait ici tout son sens.

Une fois ces améliorations apportées, le modèle pourrait être testé sur l'ensemble d'un parc habité pour une année donnée. Pour cela, plusieurs travaux sont à mener depuis la constitution de bases de données d'éléments-types à l'élaboration d'une méthode permettant de les associer (conformément à la structure du parc modélisé), en passant par la prise en compte de l'aspect météorologique. Ce parc obtenu pour une année donnée doit également pouvoir évoluer de façon dynamique de manière à permettre des études prospectives à différentes échelles temporelles. Représenter cette dynamique nécessite une méthode qui demeure à concevoir.

# Annexe A

## Aide mémoire d'électricité

Les développements proposés dans cet annexe s'appuient sur les ouvrages de Bastard et al. [BFL+99] et de Séguier et Notelet [SN06].

### Rappels sur l'énergie

L'énergie notée  $E$  est la capacité d'un système à interagir sur une autre entité – en modifiant son état initial – sous forme de mise en mouvement (travail mécanique), de transfert de chaleur ou d'émission de lumière. Cette grandeur physique s'exprime en joule [J]. Dans le domaine de l'électricité, on lui préfère le kilowattheure [kWh] qui correspond à l'énergie d'une source délivrant 1000 W pendant une heure soit  $E = 1000 \times 3600 = 3,6 \cdot 10^6$  J (1 J = 1 W.s). Par ailleurs, les multiples du kWh sont également utilisés pour des quantités significatives d'énergie, à savoir le MWh ( $10^3$  kWh) et le GWh ( $10^6$  kWh).

### Éléments historiques sur l'électricité : du paratonnerre à la centrale nucléaire

L'électricité est devenue une forme d'énergie « naturelle » qui va de paire au développement des pays industrialisés. Sa découverte est souvent associée à l'invention par Benjamin Franklin du paratonnerre en 1750 ([BFL+99] p. 131), mais des phénomènes électriques naturels – attraction de corps légers avec l'ambre frotté – étaient déjà connus et maîtrisés depuis l'antiquité. Mieux encore, les premières machines électriques ont vu le jour au milieu du XVII<sup>ème</sup> siècle.

Nous ne souhaitons pas être exhaustifs dans ces quelques rappels, néanmoins certaines grandes découvertes historiques dans ce domaine ne doivent pas être passées sous silence. C'est le cas notamment de l'établissement par Coulomb en 1785 d'une loi éponyme décrivant les interactions électrostatiques existantes entre deux particules chargées électriquement. Au début du XIX<sup>ème</sup> siècle, l'invention par Volta de la pile électrique et l'électrolyse de l'eau ouvrent la voie à l'utilisation mobile de l'électricité notamment dans les procédés chimiques. De 1820 à 1850, les grandes théories de l'électricité ont été établies par d'illustres scientifiques tels que Ampère, Faraday, Lenz et Ohm. Au même moment, les applications électriques se développent : pile thermoélectrique, roue de Barlow (premier moteur électrique), électroaimant, self-induction, pile à hydrogène, bobine d'induction.

La seconde moitié du siècle est marquée par les prémices de l'élaboration d'un réseau électrique dont la finalité est de relier les lieux de production aux zones de consommation [HUG83]. Parallèlement, les applications de l'électricité s'élargissent encore. En plus d'un vecteur énergétique, elle devient un moyen de transmission d'informations, propriété qui s'avère d'une importance majeure de nos jours. À la fin du XIX<sup>ème</sup> siècle, la découverte de la radioactivité par Becquerel en 1896 et les travaux ultérieurs de la première moitié du siècle passé vont permettre l'avènement d'une nouvelle source d'énergie : le nucléaire. Celle-ci occupe aujourd'hui un rôle central au sein du parc français de production d'électricité.

Pour conclure ce bref historique, il faut souligner l'invention en 1948 du transistor par Schockley, Brattain et Bardeen. Ce composant est notamment indispensable dans toutes les applications électroniques contemporaines ([BFL+99] p. 131).

Comme le précédent paragraphe le laisse supposer, l'électricité est probablement la forme d'énergie dont les applications sont les plus nombreuses et les plus diverses. Celles-ci peuvent être regroupées en deux catégories selon leur finalité ([BFL+99] p. 129) : gestion des signaux et gestion de la puissance. Nous proposons ci-après une liste non exhaustive des façons dont l'électricité peut être valorisée ([BFL+99] pp. 129 à 133) :

- gestion des signaux :
  - transmissions filaires (télégraphe puis téléphone et réseaux modernes de communication, commande électrique de systèmes) ;
  - transmissions sans fil (utilisation des ondes électromagnétiques : radio, télévision, satellites...);
  - traitement de l'information (informatique, domaine de l'image, domaine du livre...);
- gestion de la puissance :
  - production de matières premières (industrie métallurgique, industrie chimique...);
  - conversion en énergie mécanique (moteurs, transport, appareils domestiques);
  - conversion en énergie thermique (éclairage, chauffage domestique, soudure électrique...).

### Composition de la puissance électrique

La puissance électrique notée  $P$  est la dérivée par rapport au temps de l'énergie :

$$P = \frac{dE}{dt} \quad (\text{A.1})$$

Elle est exprimée en joule par seconde [ $\frac{\text{J}}{\text{s}}$ ] ou en watt [W] dans le système international. Là encore, on fait usuellement appel aux multiples et sous-multiples de l'unité de référence (kW, MW, GW...).

Les grandeurs instantanées, la tension  $\underline{v}$  et l'intensité  $\underline{i}$ , d'un courant électrique sinusoïdal circulant au sein d'une charge (élément consommant ou produisant de l'énergie électrique) sont deux signaux harmoniques qui varient au cours du temps  $t$ . Elles sont définies de manière classique (forme réelle) par les expressions suivantes [SN06] :

$$\underline{v}(t) = \underline{v}_{max} \cos(\omega t) = \sqrt{2}\underline{V} \cdot \cos(\omega t) \quad (\text{A.2})$$

et

$$\underline{i}(t) = \underline{i}_{max} \cos(\omega t + \phi) = \sqrt{2}\underline{I} \cdot \cos(\omega t + \phi) \quad (\text{A.3})$$

où

- $\underline{V}$  représente la **valeur efficace**\* de la tension (ou tension efficace) du courant électrique exprimée en volt [V] ;
- $\underline{I}$  est le courant efficace mesuré en ampère [A] ;
- $\underline{\omega}$  est la pulsation du signal ( $\omega = 2\pi f$  avec  $f$  la fréquence du signal) exprimée en  $\text{rad}\cdot\text{s}^{-1}$  ;
- $\underline{\phi}$  est le déphasage (ou « décalage » géométrique) existant entre la tension et le courant. Il est mesuré en radian.

Pour une fréquence donnée (en Europe 50 Hz, aux États-Unis 60 Hz), tension et courant électriques sont entièrement caractérisés par les trois seules grandeurs suivantes : tension efficace, courant efficace et déphasage.

À partir des grandeurs instantanées et des valeurs efficaces introduites par les équations A.2 et A.3 découlent les définitions suivantes :

- **puissance instantanée**\*  $\underline{P}$  exprimée en watt [W] :

$$\underline{P}(t) = \underline{v}(t) \cdot \underline{i}(t) \quad (\text{A.4})$$

- **puissance apparente**\*  $\underline{S}$  qui s'exprime en volt-ampère [VA] :

$$\underline{S} = \underline{V} \cdot \underline{I} \quad (\text{A.5})$$

En intégrant les expressions A.2 et A.3 dans l'équation A.4, la puissance instantanée devient :

$$\mathbf{P}(t) = 2\mathbf{VI} \cdot \cos(\omega t) \cdot \cos(\omega t + \phi) \quad (\text{A.6})$$

Si de plus, on utilise la définition de la puissance apparente (formule A.5), l'équation A.6 peut être simplifiée en :

$$\mathbf{P}(t) = 2\mathbf{S} \cdot \cos(\omega t) \cdot \cos(\omega t + \phi) \quad (\text{A.7})$$

Nous utilisons ensuite les formules de Simpson de transformation d'un produit de fonctions cosinus en somme :

$$\cos(\hat{a}) \cos(\hat{b}) = \frac{\cos(\hat{a} - \hat{b}) + \cos(\hat{a} + \hat{b})}{2} \quad (\text{A.8})$$

pour réarranger l'expression A.7 :

$$\mathbf{P}(t) = 2\mathbf{S} \frac{\cos(\omega t - (\omega t + \phi)) + \cos(\omega t + (\omega t + \phi))}{2} \quad (\text{A.9})$$

Enfin nous simplifions le terme commun au numérateur et au dénominateur, nous distribuons  $\mathbf{S}$  et, en nous rappelant que  $\cos \phi = \cos(-\phi)$ , la puissance instantanée  $\mathbf{P}(t)$  s'exprime plus simplement par :

$$\mathbf{P}(t) = \mathbf{S} \cdot \cos(\phi) + \mathbf{S} \cdot \cos(2\omega t + \phi) \quad (\text{A.10})$$

L'expression précédente met bien en valeur les deux composantes de la puissance instantanée :

– le terme

$$\mathbf{S} \cdot \cos \phi = \mathbf{VI} \cdot \cos \phi \quad (\text{A.11})$$

est **constant** et s'exprime en watt. Il s'agit de la puissance active notée  $\underline{P}$ .

– le second membre,  $\mathbf{S} \cdot \cos(2\omega t + \phi)$ , est quant à lui fluctuant à une fréquence de 100 Hz et de valeur moyenne nulle.

Par analogie avec la définition de la puissance active, les électriciens ont introduit une grandeur supplémentaire appelée **puissance réactive**<sup>\*</sup>, exprimée en volt-ampère réactif [VAR] (homogène au watt), notée  $\underline{Q}$  et définie par [SN06] :

$$Q = \mathbf{S} \cdot \sin \phi = \mathbf{VI} \cdot \sin \phi \quad (\text{A.12})$$

La figure A.1 propose la représentation graphique de la décomposition d'une puissance apparente en ses composantes active et réactive.

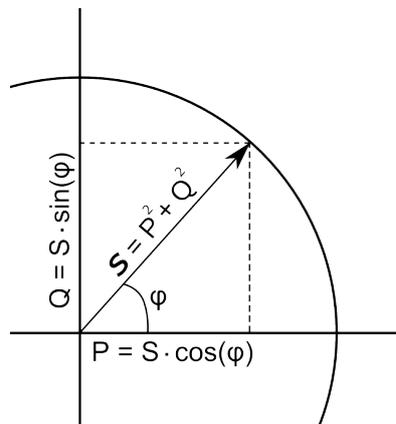


FIGURE A.1 – Décomposition d'une puissance apparente en puissance active et réactive

La puissance active est celle qui permet le fonctionnement des appareils électriques (chauffage d'une résistance, conversion mécanique...). En effet, de par son caractère constant, un transfert permanent d'énergie entre une source et une charge est rendu possible.

La puissance réactive<sup>1</sup> quant à elle, est un phénomène moins facilement « palpable » en matière d'interprétation physique. Selon Bastard et al. [BFL<sup>+</sup>99] (p. 305), il s'agit de

« la valeur crête de la puissance instantanée sinusoïdale échangée entre les éléments [...] inductifs<sup>2</sup> et capacitifs<sup>3</sup> de la charge et les sources ».

Séguier et Notelet [SN06] proposent une interprétation analogue à la précédente mais sans doute plus didactique en stipulant que la puissance réactive s'apparente aux échanges d'énergie et de valeur moyenne nulle s'effectuant entre une source et un récepteur de type inductance ou capacité.

Un bilan énergétique sur la puissance réactive échangée entre une (des) source(s) et une (des) charge(s) montre en effet que ce type d'échange est nul. Cependant, négliger cette quantité conduirait en réalité à sous-dimensionner tout réseau transitant un courant alternatif.

Ainsi, sur les lignes d'un réseau circulent simultanément de la puissance active et réactive. Le gestionnaire du réseau a tout intérêt à limiter le transit de puissance réactive. En effet, la circulation de puissance réactive entraîne des pertes d'énergie supplémentaires au simple acheminement de puissance active (utile). Pour illustrer nos propos, étudions le montage simple présenté sur la figure A.2. Il s'agit d'un générateur de courant alternatif associé à deux charges disposées en parallèle absorbant la même puissance active mais des puissances réactives distinctes ([BFL<sup>+</sup>99] p. 307). La figure A.3 présente quant à elle le diagramme de Fresnel correspondant au circuit électrique de l'illustration A.2.

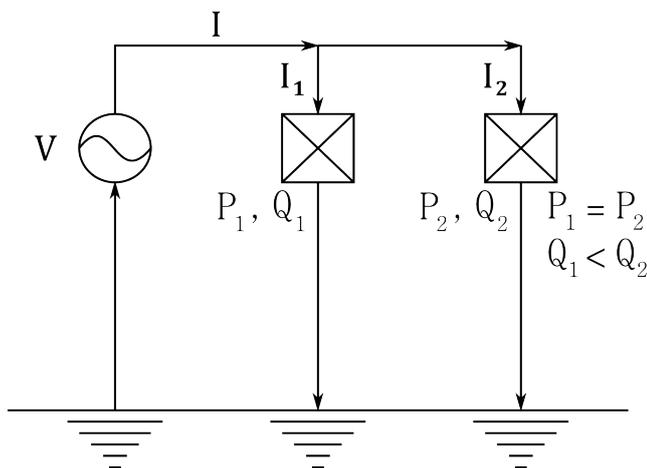


FIGURE A.2 – Schéma électrique de deux charges ayant une puissance active égale mais des puissances réactives distinctes

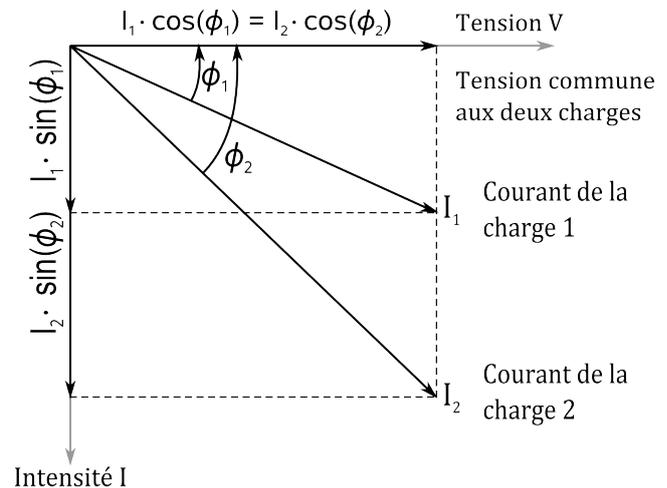


FIGURE A.3 – Diagramme de Fresnel associé au circuit électrique de la figure A.2

L'exemple précédent permet de bien appréhender l'impact du besoin en puissance réactive d'une charge électrique. En effet, pour une même puissance active requise, une puissance réactive plus importante nécessite la fourniture d'un courant électrique d'intensité plus élevée. Or, l'un des phénomènes occasionnant les pertes sur un réseau est l'échauffement des lignes conductrices par effet Joule. Celui-ci étant proportionnel au carré de l'intensité du courant transité, une plus grande puissance réactive conduit à des échauffements accrus et donc à des pertes supplémentaires au cours du transport de l'électricité.

La différence de besoin en puissance réactive n'est pas remarquée d'un point de vue consommateur. En revanche, côté producteur et distributeur d'électricité, la limitation du transit de la puissance réactive est d'un intérêt capital en ce qui concerne les pertes sur le réseau, de dimensionnement des installations et de mise en place de matériels de compensation de cette puissance réactive (typiquement intégrer des éléments capacitifs – bancs de condensateurs – à un réseau trop fortement inductif).

1. « gangrène des réseaux » selon l'ingénieur français Paul Boucherot [BFL<sup>+</sup>99] p. 305

2. Comportement type bobine, consommation de puissance réactive

3. Comportement type condensateur, fourniture de puissance réactive

Dans le jargon du réseau, on cherche souvent à « relever le  $\cos \phi$  », ce qui revient à diminuer le déphasage entre courant et tension. Ce paramètre est communément appelé **facteur de puissance**<sup>\*</sup>, nous le noterons  $F$ . Il mesure la « qualité » du courant circulant dans un réseau puisqu'il compare la puissance active et la puissance apparente d'un courant électrique :

$$F = \frac{\text{Puissance active}}{\text{Puissance apparente}} = \frac{P}{S} = \cos \phi \quad (\text{A.13})$$

Le facteur de puissance peut être vu comme un indicateur signifiant l'efficacité de fourniture, par un système électrique, de puissance active. Un bon facteur de puissance ne doit pas descendre en dessous de 0,8.

La puissance active est la seule qui soit effectivement facturée au consommateur. En revanche, des sanctions financières peuvent être appliquées à des clients dépassant les limites contractuelles en matière de besoin en puissance réactive.

Comme définie précédemment par l'équation A.5 et comme cela est indiqué sur la figure A.1, la puissance apparente peut également s'exprimer en fonction de la puissance active  $P$  et de la puissance réactive  $Q$  :

$$S = \mathbf{VI} = \sqrt{P^2 + Q^2} \quad (\text{A.14})$$

La puissance apparente n'a qu'un intérêt de dimensionnement des applications (taille, calibre...) et n'indique jamais la consommation en puissance d'un appareil. De ce fait, cette grandeur est historiquement exprimée en volt-ampère (VA).

## Revue des puissances sur le réseau électrique

### Présentation du réseau électrique national

De manière générale, le réseau électrique français se structure selon les trois « couches » suivantes :

- le **réseau de transport** composé des lignes 400 et 225 kV : ces lignes partent des centrales de production (nucléaire, thermique, hydraulique) et constituent le réseau **Très Haute Tension (THT)**<sup>\*</sup>. Elles alimentent certains gros consommateurs d'électricité (clients industriels) ainsi que les postes de transformation Très Haute Tension → **Haute Tension (HT)**<sup>\*</sup> ;
- le **réseau de répartition** formé par les lignes 90 à 63 kV : elles démarrent au niveau des postes de transformation THT → HT et alimentent les postes sources (lieux de la transformation Haute Tension → Moyenne Tension) ;
- le **réseau de distribution** dont les lignes ne transitent que des tensions inférieures à 50 kV (principalement du 20 kV) : ces tensions sont ensuite abaissées jusqu'au seuil de 220 V, qui correspond à la **Basse Tension (BT)**<sup>\*</sup>. Ces faibles tensions alimentent les consommateurs domestiques et les petits professionnels.

Cette simple représentation omet de nombreuses particularités présentes sur le réseau. De ce fait, nous proposons dans le tableau A.1 issu de [EG91] une caractérisation normalisée des paliers de tension et la sémantique qui y est associée.

Domaine de tension	Valeur de la tension nominale $V_{nom}$ exprimée en volt	
	en courant alternatif	en courant continu lisse
Très Basse Tension (domaine TBT)	$V_{nom} \leq 50$	$V_{nom} \leq 120$
Basse Tension (domaine BT)	domaine BTA $50 < V_{nom} \leq 500$	$120 < V_{nom} \leq 750$
	domaine BTB $500 < V_{nom} \leq 1000$	$750 < V_{nom} \leq 1500$
Haute Tension (domaine HT)	domaine HTA $1000 < V_{nom} \leq 50000$	$1500 < V_{nom} \leq 75000$
	domaine HTB $V_{nom} > 50000$	$V_{nom} > 75000$

TABLEAU A.1 – Paliers de tension sur le réseau de transport national



## Annexe B

# Détails de l'algorithme de placement des équipements au sein des intervalles infra-journaliers

Le tableau [B.1](#) est la version complète de celui (tableau [3.9](#)) que nous avons décrit au sein du chapitre [3](#), section [3.2.4.2](#). Plus particulièrement, il propose, par sa dernière colonne, les explicitations des calculs de cas favorables en fonction des configurations de placement des équipements au sein des intervalles infra-journaliers.

Le tableau [B.2](#) recense les différentes configurations de recherche des instants éligibles au déclenchement d'un équipement au sein d'un intervalle. Dans ce tableau, en plus des conventions d'écriture précédemment utilisées, nous introduisons l'élément  $IC^*$  qui désigne l'Intervalle Choisi.

Code cas	Critère 1 $IT = IL$	Critère 2 $IT = IAC$	Critère 3 Autorisation dépassement intervalle	Critère 4 Accueil de la forme développée de l'usage de l'équipement dans l'intervalle	Nombre de cas favorables
I.A	oui	oui	non testée	non testé	$\max(0, l_{Intervalle} - d_{Usage, i} + 1 - r_{EAC} + 1 - d_{AC})$
I.B	oui	non	non testée	non testé	$\max(0, l_{Intervalle} - d_{Usage, i} + 1 - d_{AC})$
II.A.1	non	oui	non	intégralité de $d_{Usage, i}$	$l_{Intervalle} - d_{Usage, i} + 1 - r_{EAC} + 1$
II.A.2	non	oui	oui	intégralité de $d_{Usage, i}$	$l_{Intervalle} - d_{Usage, i} + 1 - r_{EAC} + 1 + d_{Tampon} + 1$
II.A.3.1	non	oui	oui	intégralité de $d_{PNDU} + d_{Tampon}$	$d_{Tampon} + 1$
II.A.3.2	non	oui	oui	intégralité de $d_{PNDU}$	$l_{Intervalle} - d_{PNDU} + 1 - r_{EAC} + 1$
II.B.1	non	non	non	intégralité de $d_{Usage, i}$	$l_{Intervalle} - d_{Usage, i} + 1$
II.B.2	non	non	oui	intégralité de $d_{Usage, i}$	$l_{Intervalle} - d_{Usage, i} + 1 + d_{Tampon} + 1$
II.B.3.1	non	non	oui	intégralité de $d_{PNDU} + d_{Tampon}$	$d_{Tampon} + 1$
II.B.3.2	non	non	oui	intégralité de $d_{PNDU}$	$l_{Intervalle} - d_{PNDU} + 1$
III	oui	non pertinent	non testée	non testé	$\max(0, l_{Intervalle} - d_{Usage, i} - d_{AC} + 1)$
IV.A	non	non pertinent	non	intégralité de $d_{Usage, i}$	$l_{Intervalle} - d_{Usage, i} + 1$
IV.B	non	non pertinent	oui	intégralité de $d_{Usage, i}$	$l_{Intervalle} - d_{Usage, i} + 1 + d_{Tampon} + 1$
IV.C.1	non	non pertinent	oui	intégralité de $d_{PNDU} + d_{Tampon}$	$d_{Tampon} + 1$
IV.C.2	non	non pertinent	oui	intégralité de $d_{PNDU}$	$l_{Intervalle} - d_{PNDU} + 1$

 $l_{IAC} > 1$  $l_{IAC} = 0$ 

TABLEAU B.1 – Expliciteations des calculs du nombre de cas favorables dans les intervalles infra-journaliers en fonction des configurations de placement des équipements

Code cas	$IC = IAC$	$IC = IL$	$d_{AC} ?$	Report nécessaire	$r_{EAC}$	Autorisation dépassement intervalle	Bornes du compteur $comp_i$ de balayage dans l'intervalle	Bornes du compteur $comp_{usage, i}$ de balayage dans la forme développée de l'usage d'un équipement
I	non testé	non testé	non testée	oui	non évalué	effective mais non testée	de $l_{Intervalle} - n_{CF}$ à $l_{Intervalle}$	de 1 à $n_{CF}$
II.1	non testé	non testé	non testée	non	0	non	de 1 à $n_{CF}$	de 1 à $d_{Usage, i}$
II.2.1	non testé	non testé	= 0	non	0	oui	de 1 à $l_{Intervalle} - d_{Usage, i} + 1$ et de $(l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1$ à $(l_{Intervalle} - d_{PNDU} + 1)$	de 1 à $d_{Usage, i}$ et de 1 à $l_{Intervalle} - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1) - [n_{CF} - ((l_{Intervalle} - d_{PNDU} + 1) - [comp_i - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1)])]$
II.2.2.1	non testé	oui	> 0	non	0	oui	de 1 à $n_{CF}$	de 1 à $d_{Usage, i}$
II.2.2.2	non testé	non	> 0	non	0	oui	de 1 à $l_{Intervalle} - d_{Usage, i} + 1$ et de $(l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1$ à $(l_{Intervalle} - d_{PNDU} + 1)$	de 1 à $d_{Usage, i}$ et de 1 à $l_{Intervalle} - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1) - [comp_i - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1)])]$
II.3.1	oui	vide	non testée	non	> 0	non	de $TEAC$ à $TEAC + n_{CF} - 1$	de 1 à $d_{Usage, i}$
II.3.2	non	vide	non testée	non	> 0	non	de 1 à $n_{CF}$	de 1 à $d_{Usage, i}$
II.4.1	oui	vide	non testée	non	> 0	oui	de $TEAC$ à $TEAC + (l_{Intervalle} - TEAC - d_{Usage, i} + 1)$ et de $(l_{Intervalle} - d_{PNDU} + 1) - (n_{CF} - [l_{Intervalle} - (d_{Usage, i} - 1) - (TEAC - 1)]) + 1$ à $l_{Intervalle} - d_{PNDU} + 1$	de 1 à $d_{Usage, i}$ et de 1 à $l_{Intervalle} - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1) - [comp_i - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1)])]$
II.4.2	non	vide	non testée	non	> 0	oui	de 1 à $(l_{Intervalle} - d_{Usage, i} + 1)$ et de $(l_{Intervalle} - d_{PNDU} + 1) - (n_{CF} - [l_{Intervalle} - (d_{Usage, i} - 1)]) + 1$ à $l_{Intervalle} - d_{PNDU} + 1$	de 1 à $d_{Usage, i}$ et de 1 à $l_{Intervalle} - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1) - [comp_i - ((l_{Intervalle} - d_{PNDU} + 1) - [n_{CF} - (l_{Intervalle} - d_{Usage, i} + 1)] + 1)])]$

TABLEAU B.2 – Configurations possibles pour le balayage des instants éligibles dans les intervalles infra-journaliers



## Annexe C

# Description des étapes de la simulation d'un scénario grâce à notre modèle de reconstitution de courbe de charge dans l'environnement DYMOLA

Cet annexe décrit l'organisation et le mode opératoire de l'implémentation informatique sous DYMOLA de la méthode de reconstitution de courbe de charge tels qu'ils sont à l'issue de la thèse.

Nous détaillons ci-après l'ensemble des étapes d'une simulation en suivant le fil conducteur temporel qui les lie. En outre, chacun des concepts de l'algorithme que nous avons abordé dans le chapitre 3 est implémenté sous forme de modèle dans Dymola (en quelque sorte il s'agit d'une « boîte » programmable). Un regroupement de modèles est appelé [assemblage](#)\*. Dans un assemblage, les modèles sont liés les uns aux autres par l'intermédiaire de connecteurs de manière à échanger entre eux de l'information et des données.

La figure [C.1](#) illustre schématiquement les concepts d'assemblage et de modèle. Notons également ici que le script mentionné page [168](#) assure la simulation successive des différents assemblages et qu'en outre il comporte diverses instructions de bas niveau de manière à assurer la jonction entre les assemblages.

Les premiers éléments intervenant dans le scénario simulé sont les modèles de jour-type, de ménage-type et d'intervalles infra-journaliers. Ceux-ci sont groupés dans un assemblage appelé `ConstructionLongueurIntervalles`.

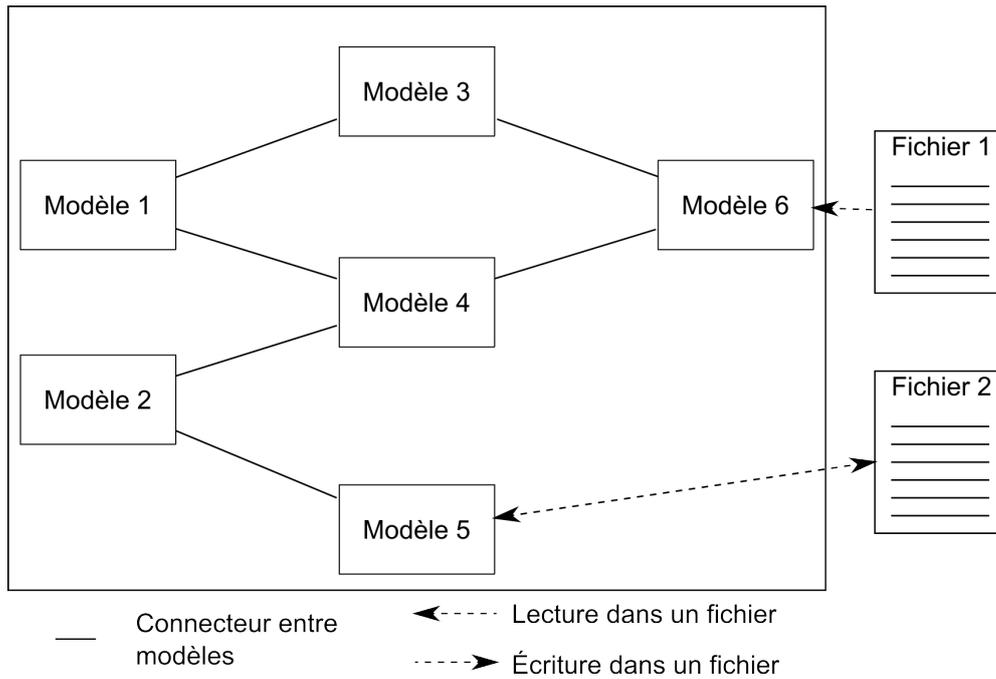


FIGURE C.1 – Représentation schématique d'un assemblage dans Dymola et des relations entre modèles et fichiers de données

Au niveau du jour-type doivent être paramétrées les limites inférieures et supérieures des instants caractéristiques ainsi que la résolution temporelle choisie pour mener les calculs. Pour le ménage-type, seul un paramètre relatif au type de foyer modélisé doit être renseigné. En outre, chacun des 6 intervalles infra-journaliers est connecté au jour-type et au ménage-type de façon à recevoir respectivement les limites temporelles inférieures et supérieures qui les encadrent et le type de ménage considéré.

La simulation de cet assemblage de modèles conduit à la définition des 7 instants caractéristiques  $t_\alpha$  et à l'obtention, dans chacun des intervalles infra-journaliers, du niveau de disponibilité du ménage considéré de manière agrégée. En outre, cette étape permet de mettre en mémoire les éléments permettant de définir la forme développée de chacun des intervalles infra-journaliers<sup>1</sup>. Le tableau C.1 indique les valeurs<sup>2</sup> des paramètres du scénario principal ASL sélectionnées au niveau de l'assemblage `ConstructionLongueurIntervalles`.

Paramètre	Valeur	Unité
$inf t_{Start}$	5	h
$supt t_{Start}$	8	h
$inf \Delta t_{Départ}$	0,25	h
$supt \Delta t_{Départ}$	2	h
$inf \Delta t_{Abs. Jour.}$	9	h
$supt \Delta t_{Abs. Jour.}$	14	h
$inf t_3$	11,75	h
$supt t_3$	12,75	h
$inf \Delta t_{Prés. Mér.}$	0,5	h
$supt \Delta t_{Prés. Mér.}$	1,5	h
$inf t_{End}$	20	h
$supt t_{End}$	24,5	h
$inf t_{StartBis}$	30	h
$supt t_{StartBis}$	32	h
TypeMenage	Inactif	-
ResolutionCDC	1	min

TABEAU C.1 – Valeurs des paramètres à renseigner au niveau de l'assemblage `ConstructionLongueurIntervalles` pour le scénario principal ASL

Comme nous l'avons souligné, au sortir du premier assemblage, nous ne disposons pas encore des intervalles infra-journaliers sous leur forme développée. En effet, puisqu'ils sont représentés par des tableaux dont les longueurs ne sont connues qu'à l'issue de la simulation du premier assemblage, nous avons dû

1. À ce niveau, le résumé de chacun des intervalles dans son état initial est également construit de manière standard, c'est-à-dire de façon identique quel que soit l'intervalle infra-journalier considéré, i. e. avec un nombre maximal – paramétrable – d'événements par défaut

2. Soulignons ici que l'ensemble des valeurs des paramètres choisies pour le scénario modélisé (présentées entre autres dans le tableau C.1), n'ont pas d'autre objectif que de constituer un cas d'étude. Ainsi il ne s'agissait pas de modéliser une réalité mesurée donnée

---

définir un second assemblage très basique qui ne fait que construire ces différents tableaux conformément à leur longueur qui a été calculée lors de la première étape.

Un troisième assemblage intervient ensuite pour réinitialiser un certain nombre de fichiers contenant des informations nécessaires au placement des équipements. Ces fichiers sont lus par l'assemblage suivant appelé `CalculCompatibiliteMicro`. Ce quatrième assemblage, réalise à lui seul presque la totalité des opérations restantes de l'algorithme. Il contient tout d'abord les modèles des macro-tâches et des micro-tâches. Ces dernières sont reliées aux modèles des équipements simulés au sein du scénario, modèles qui figurent également au sein de l'assemblage `CalculCompatibiliteMicro`.

Les différentes procédures de placement des équipements domestiques sont opérées directement dans les modèles de ceux-ci. La simulation de l'assemblage `CalculCompatibiliteMicro` aboutit ainsi aux instants de déclenchement de chaque équipement simulé. De plus, les éléments du paramétrage sont conservés de manière à construire les formes développées de l'usage des équipements domestiques dans un dernier assemblage. Enfin, les courbes de charge des appareils sont construites à partir d'une commande du script qui à la fois met en mémoire les appels de puissance mais aussi les instants de déclenchement des appareils domestiques, et ce, pour chaque simulation.



## Annexe D

# Relations entre la mesure $DTW$ et des indicateurs d'intérêt sur des courbes de charge moyennes foisonnées

Les figures D.1 à D.4 proposent les représentations des relations entre la mesure  $DTW$  et des indicateurs classiquement calculés sur des courbes de charge moyennes foisonnées.

Plus précisément les cas d'études présentés ci-dessous concernent le sèche-linge (65  $CDCMF$  construites à partir de résultats de simulation du scénario principal ASL – cf. figure D.1), la télévision (40  $CDCMF$  construites à partir des résultats de la campagne REMODECE – cf. figure D.2), la machine à laver (22  $CDCMF$  issues de mesures de la campagne REMODECE – cf. figure D.3) et le lave-vaisselle (20  $CDCMF$  issues de mesures de la campagne REMODECE – cf. figure D.4).

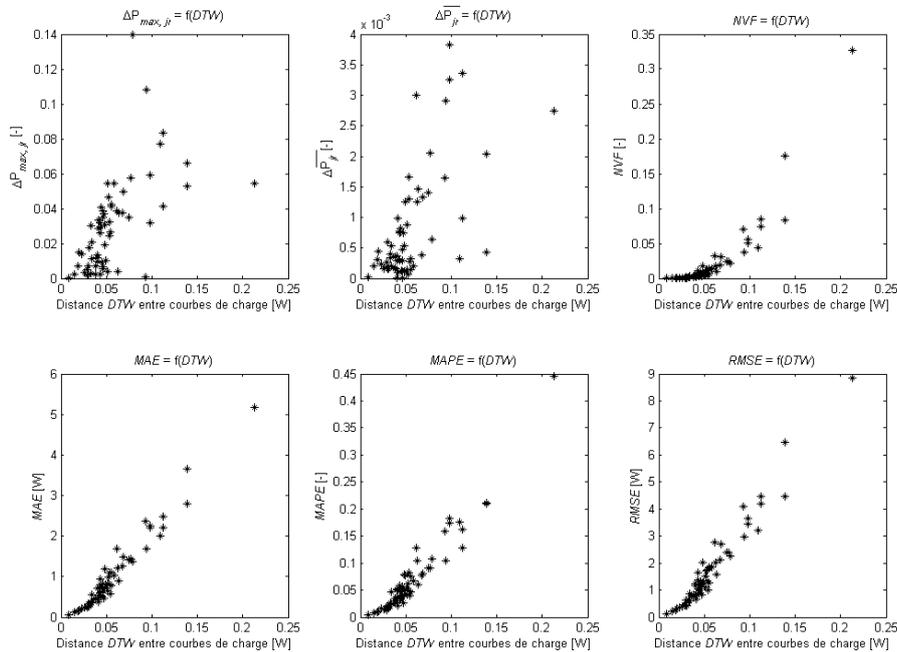


FIGURE D.1 – Relations entre la distance  $DTW$  et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de sèche-linge obtenues par simulation (6500 courbes de charge unitaires)

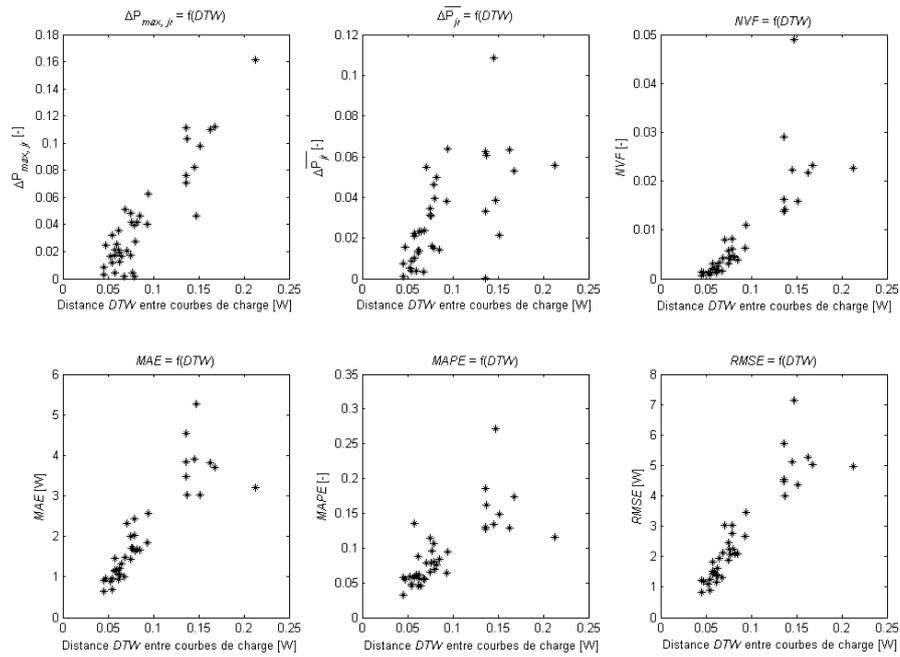


FIGURE D.2 – Relations entre la distance  $DTW$  et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de télévision issues de la campagne REMODECE (3949 courbes de charge unitaires)

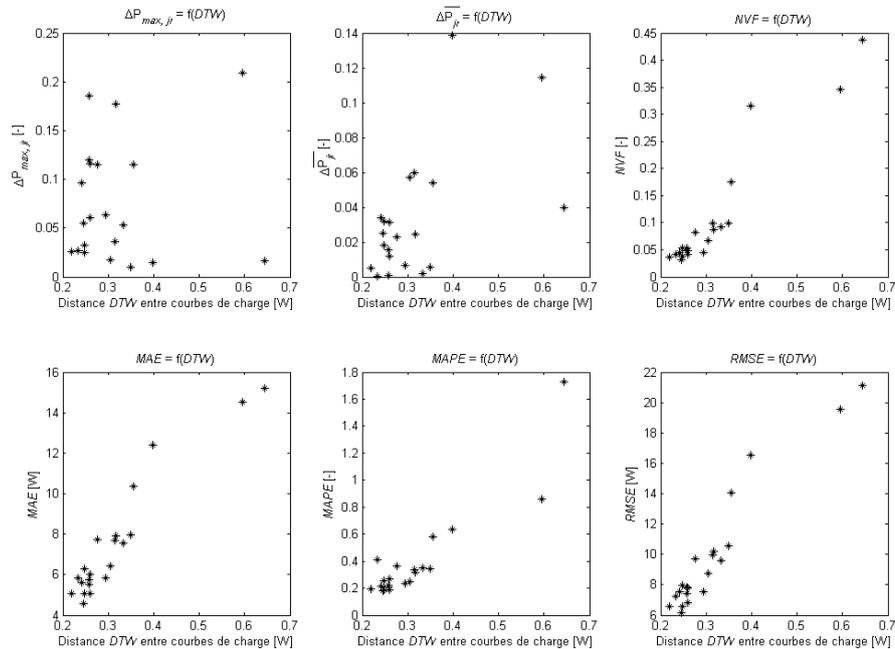


FIGURE D.3 – Relations entre la distance  $DTW$  et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de machine à laver issues de la campagne REMODECE (2189 courbes de charge unitaires)

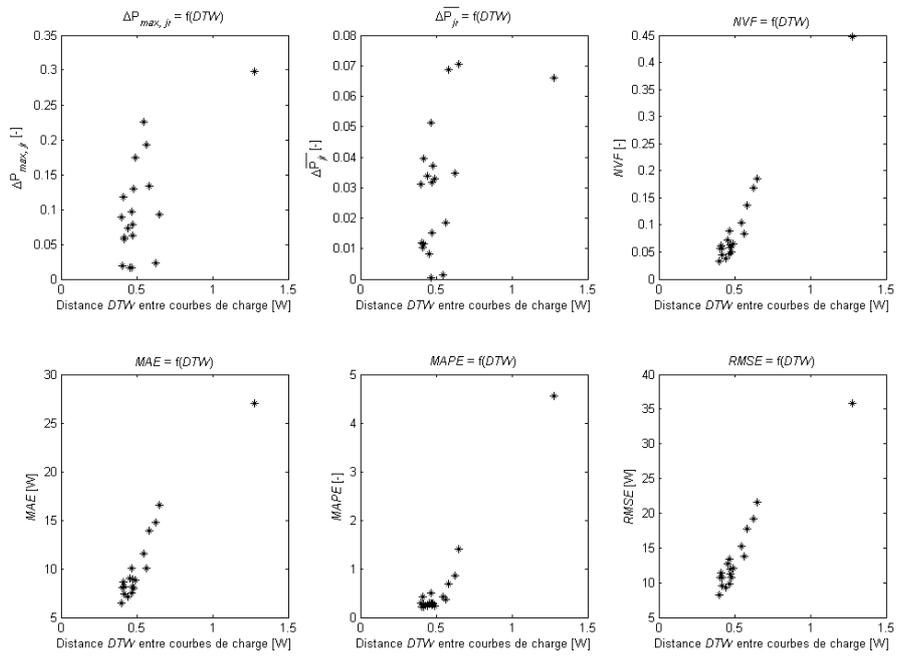


FIGURE D.4 – Relations entre la distance  $DTW$  et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de lave-vaisselle issues de la campagne REMODECE (1961 courbes de charge unitaires)



## Annexe E

# Étude des écarts d'origine temporelle et en puissance entre $CDCMF$

Les figures E.1 à E.4 illustrent la part des écarts relatifs au déphasage temporel et au décalage en puissance entre courbes de charge moyennes foisonnées construites à partir de données mesurées et simulées. Le calcul des  $CDCMF$  a fait appel aux 6 processus décrits au sein du paragraphe 4.3 du chapitre 4. Plus particulièrement, les graphiques ci-dessous représentent la décomposition du foisonnement pour le sèche-linge (65  $CDCMF$  construites à partir de simulations du scénario principal ASL – cf. figure E.1), la télévision (40  $CDCMF$  issues de mesures lors de la campagne REMODECE – cf. figure E.2), le lave-linge (22  $CDCMF$  construites à partir des mesures de la campagne REMODECE – cf. figure E.3) et enfin le lave-vaisselle (20  $CDCMF$  établies à partir des résultats de la campagne REMODECE – cf. figure E.4).

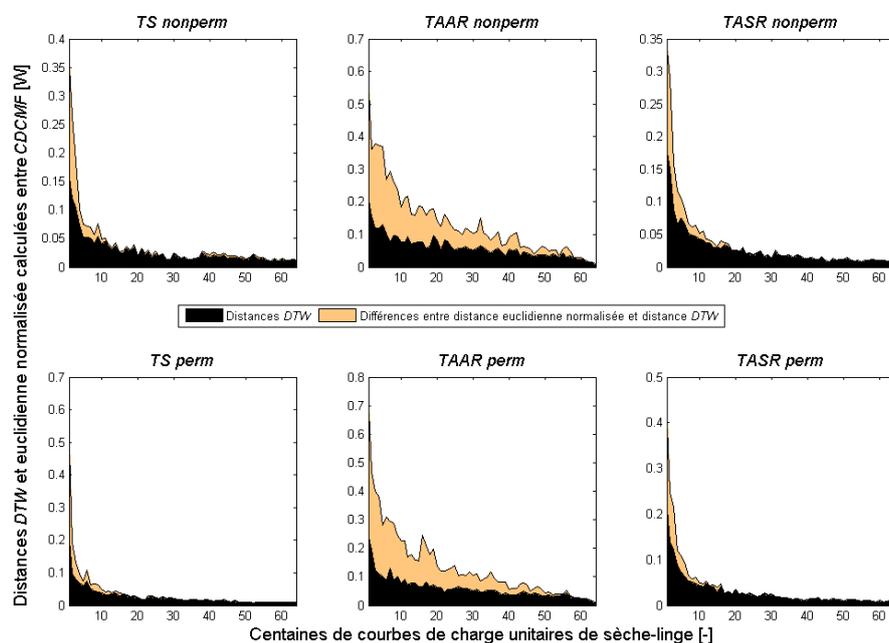


FIGURE E.1 – Visualisation de la part des écarts d'origine temporelle et en puissance entre  $CDCMF$  de sèche-linge, construites par différents processus, à partir de 6500  $CDCU$  simulées

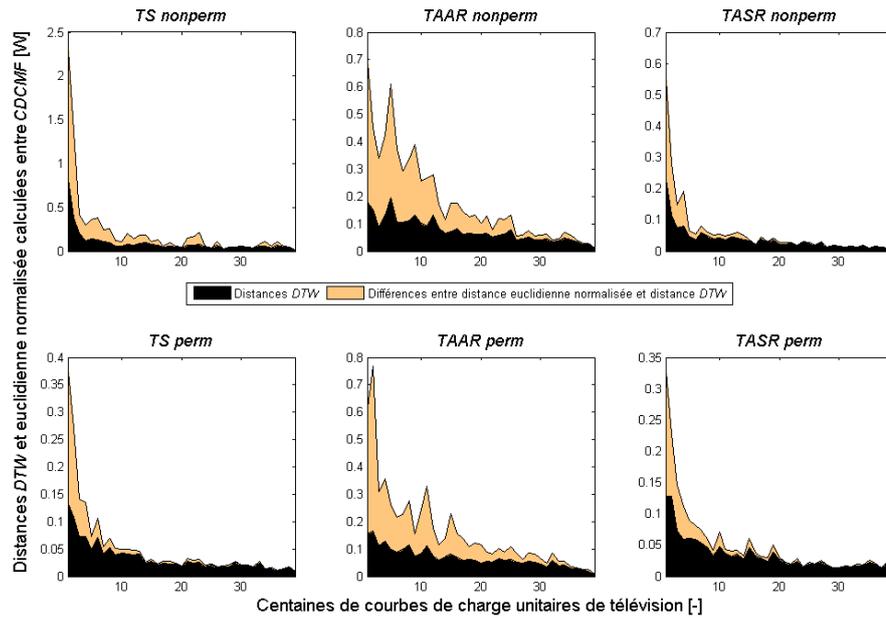


FIGURE E.2 – Visualisation de la part des écarts d'origine temporelle et en puissance entre *CDCMF* de télévision, construites par différents processus, à partir de 3949 *CDCU* mesurées.

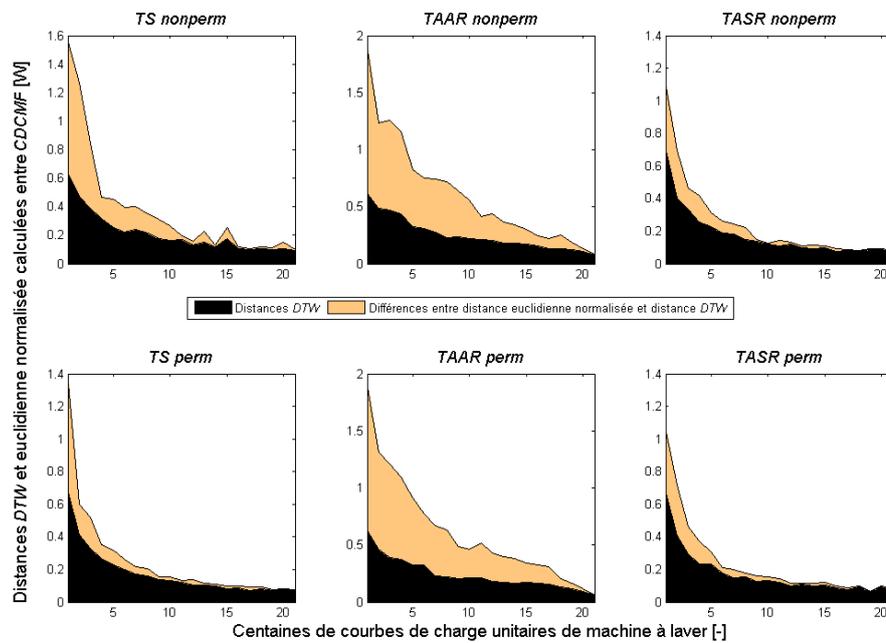


FIGURE E.3 – Visualisation de la part des écarts d'origine temporelle et en puissance entre *CDCMF* de machine à laver, construites par différents processus, à partir de 2189 *CDCU* mesurées.

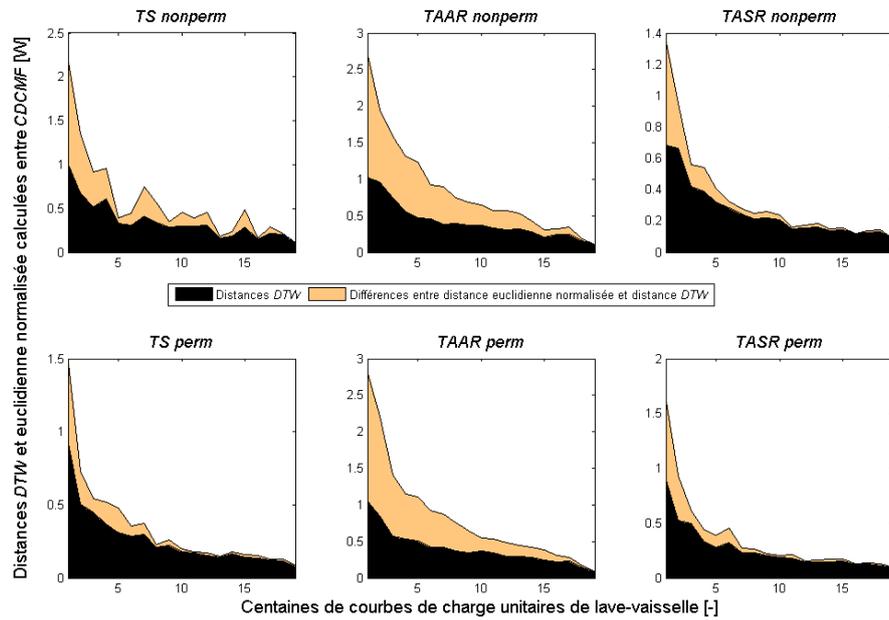


FIGURE E.4 – Visualisation de la part des écarts d’origine temporelle et en puissance entre  $CDCMF$  de lave-vaisselle, construites par différents processus, à partir de 1961  $CDCU$  mesurées.



# Annexe F

## Analyses complémentaires sur des courbes de charge mesurées de télévision issues de la base de données REMODECE

Dans cet annexe, nous détaillons les deux études que nous avons menées sur des courbes de charge normées de télévision (profils de visionnement) issues de la campagne de mesures REMODECE.

La partie [F.1](#) expose les résultats de l'ACP conduite sur les données normées, ACP suivie d'une CAH sur les coordonnées des observations dans la nouvelle base fournie par l'ACP. La section [F.2](#) quant à elle, présente les détails de la CAH opérée seule sur les profils de visionnement. Plus précisément, deux CAH ont été conduites, respectivement sans puis avec fixation en entrée d'algorithme du nombre de classes désirées.

### F.1 ACP sur courbes de charge mesurées normées puis CAH sur les résultats de l'ACP

Comme évoqué à l'issue du paragraphe [4.4.5](#) du chapitre [4](#), l'utilisation de courbes de charge normées (et donc adimensionnelles<sup>1</sup>) laisse supposer en théorie que l'analyse qui en est faite est indépendante de l'influence liée au type de télévision. Celui-ci conditionne largement le niveau des appels de puissance des appareils et plus indirectement de la présence d'une veille. De ce fait, il est attendu que cette nouvelle analyse différencie les courbes de charge uniquement en fonction de leur forme.

L'analyse en composantes principales que nous avons effectuée sur les profils de visionnement donne lieu à un ensemble de vecteurs propres, chacun d'eux étant associé à une valeur propre. Comme ce fut le cas dans la partie [4.4.4](#), nous ne nous intéresserons simplement qu'aux résultats présentés dans le tableau [F.1](#).

	F1	F2	F3	F4	F5
Valeur propre	36,148	16,776	11,549	9,305	6,844
Pourcentage d'inertie expliquée [%]	25,102	11,650	8,020	6,462	4,752
Pourcentage cumulé d'inertie expliquée [%]	25,102	36,752	44,773	51,234	55,987

TABLEAU F.1 – Résultats sur les cinq premiers axes de l'ACP menée sur les données normées

1. Rappelons au passage que la puissance réduite de la télévision, qui est utilisée dans cette étude, correspond au rapport exprimé en pourcentage entre l'appel de puissance de cet appareil à un instant donné et la puissance journalière maximale de ce même équipement

La figure F.1 indique l'ensemble des valeurs propres calculées par l'ACP ainsi que la répartition sur les 144 axes de la variabilité du nuage de points.

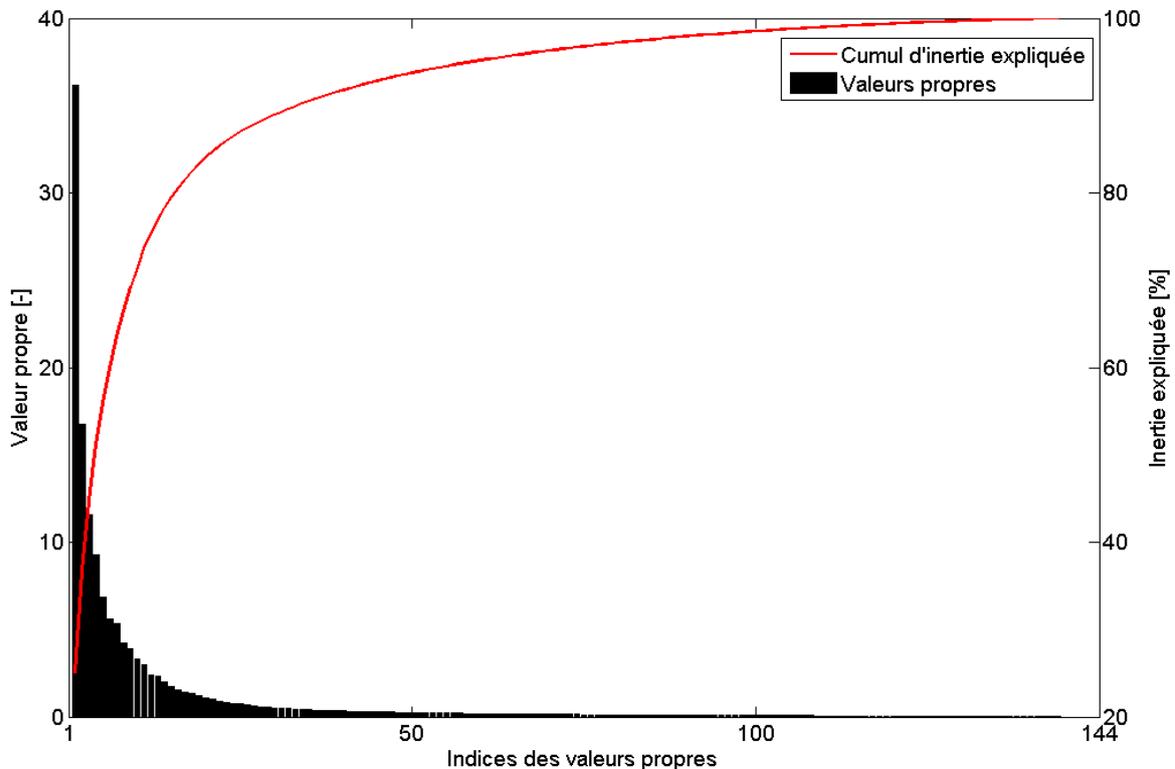


FIGURE F.1 – Valeurs propres issues de l'ACP sur données normées et pourcentage cumulé d'explicitation de la variabilité du nuage de points

Les figures<sup>2</sup> F.2 et F.3 représentent les projections sur les trois premiers axes de l'ACP des variables manipulées.

À l'image des résultats de l'ACP menée dans la partie 4.4.4 du chapitre 4, les figures F.2 et F.3 témoignent d'une corrélation apparente entre les 144 points 10 min qui constituent les profils journaliers de visionnement. Une fois encore, nous remarquons que ces variables se répartissent respectivement au sein de deux cadrans (projection sur F1 et F2) et de quatre cadrans (projection sur F2 et F3) du cercle de corrélation. Dans les deux cas, la projection des variables aboutit à une trajectoire orientée où la répartition des points projetés n'est pas homogène.

Les résultats en matière de variabilité cumulée sont très similaires à la première ACP (56 % d'inertie expliquée en 5 axes), aucun changement flagrant n'est observé. En revanche, l'étude des contributions et des cosinus carrés des variables montre que les axes ne sont pas créés par les mêmes variables :

- le premier axe est constitué par les variables 1 à 106 (instants compris entre 0 h 00 et 17 h 30) ;
- le deuxième axe est principalement expliqué par les variables 107 à 120 (moments de la journée répartis entre 17 h 40 et 19 h 50) ;
- enfin le troisième axe est le fruit des variables 121 à 141 (partie de la journée comprise entre 20 h 00 et 23 h 20).

2. Sur ces figures, les pourcentages désignent les valeurs d'inertie du nuage expliquée par les axes

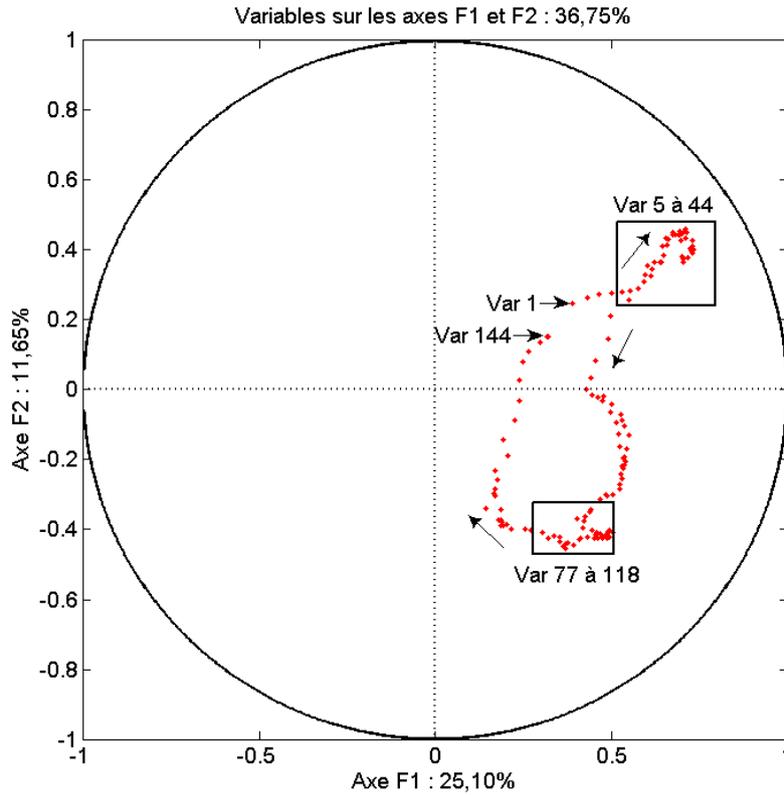


FIGURE F.2 – Cercle de corrélation entre les variables et les facteurs principaux 1 et 2 de l'ACP sur données normées

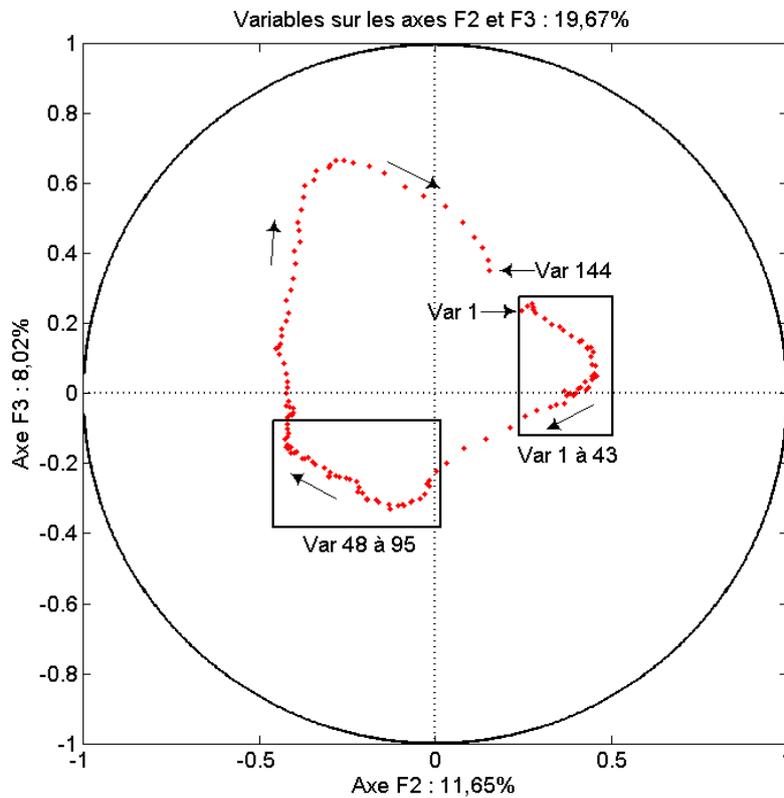


FIGURE F.3 – Cercle de corrélation entre les variables et les facteurs principaux 2 et 3 de l'ACP sur données normées

Avec les profils de visionnement, l'analyse en composantes principales n'est pas influencée par les niveaux de puissance des postes de télévision. En effet, elle n'est supposée être dépendante que de la forme de ces profils. Les trois premiers axes factoriels de cette ACP sont chacun construits par **une série continue**<sup>3</sup> de points 10 min, chacune d'elles correspond ainsi à une période infra-journalière parfaitement définie. Hasard de l'analyse ou non, les périodes identifiées s'enchaînent de manière chronologique conformément au rang des axes factoriels. Ces résultats semblent indiquer l'identification de probabilités différentes d'usage de la télévision au fil d'une journée.

Une fois l'ACP effectuée, nous avons sélectionné les coordonnées des observations dans la nouvelle base et avec ces valeurs nous opérons une CAH sans contraindre l'algorithme en matière de nombre de classes à obtenir. La classification donne alors lieu à quatre classes dont des éléments de caractérisation sont compilés dans le tableau F.2.

Classe	Objet central	Taille de la classe	Variance intra-classe	Distance au barycentre de la classe		
				minimale	moyenne	maximale
1	106_40_1_2797	2461	73,875	4,470	8,272	21,392
2	18_40_5_331	948	132,667	6,237	11,425	16,582
3	107_39_7_2827	336	163,303	6,891	12,514	20,864
4	100_40_2_2586	204	236,796	4,173	14,541	23,466

TABLEAU F.2 – Caractéristiques des classes de la CAH sur les résultats de l'ACP menée sur les données normées

La CAH donne lieu à quatre classes : la première est la plus conséquente du point de vue de l'effectif puisqu'elle rassemble 62 % de l'effectif total des profils de visionnement. De plus, comme ce fut le cas pour l'ACP conduite sur les données brutes (cf. partie 4.4.4 du chapitre 4), la taille des classes formées par l'algorithme diminue avec leur rang et la variance intra-classe augmente simultanément. La figure F.4 reproduit les profils de visionnement qui correspondent aux objets centraux des classes de cette CAH. Soulignons finalement que ces derniers ne témoignent que de deux types de télévision (les types 39 et 40, respectivement TV LCD  $\varnothing > 81$  cm et TV plasma  $\varnothing \leq 107$  cm). Sans doute, les profils de visionnement de ces deux types d'appareil sont les plus représentatifs de l'ensemble de l'échantillon considéré.

L'examen des objets centraux des classes de la CAH se révèle particulièrement intéressant. Tout d'abord, nous remarquons que les profils identifiés sont très différents. Les moins surprenants sont ceux relatifs aux deux premières classes : pour la classe 1, le profil de 106\_40\_1\_2797 est celui d'un « téléspectateur actif » pour un jour de semaine (le lundi). La télévision est utilisée le soir (de 20 h 20 à 22 h 20) et le reste du temps, elle demeure éteinte (la puissance réduite correspondante vaut 0).

Le profil de visionnement représentatif de la classe 2 correspond à un usage intensif du petit écran puisque l'appareil est allumé en continu de 9 h 40 à 22 h. Comme pour la classe 1, le profil de visionnement est celui d'un jour de semaine puisque 18\_40\_5\_331 est une mesure d'un vendredi.

Viennent ensuite deux profils assez particuliers : tout d'abord celui de la classe 3 (107\_39\_7\_2827), qui est le profil de visionnement d'un dimanche. Nous remarquons tout d'abord qu'une période de fonctionnement est visible depuis le début du jour jusqu'à 1 h 30 du matin. Durant celle-ci, les appels de puissance correspondent à 85 à 90 % de l'appel journalier maximal. Ensuite, une deuxième période plus longue débute et s'étend jusqu'à 4 h 20 du matin. Pendant celle-ci, la puissance réduite est de l'ordre de 12 à 15 %, autrement dit une puissance appelée aux alentours de 30 W pour cet appareil. Il semblerait que cette période ne soit pas de la veille seule puisqu'entre 4 h 30 et 7 h 40, la puissance réduite du téléviseur est plus faible encore que lors de la période précédente puisqu'elle fluctue aux alentours de 0,5 à 0,7 % (soit de 1,2 à 1,8 W pour la télévision en question). En faisant l'hypothèse que la troisième période corresponde à de la veille, la seconde période peut éventuellement témoigner d'un fonctionnement

3. Cette conclusion va à l'inverse des résultats de l'analyse factorielle précédente que nous avons détaillée au paragraphe 4.4.4 du chapitre 4

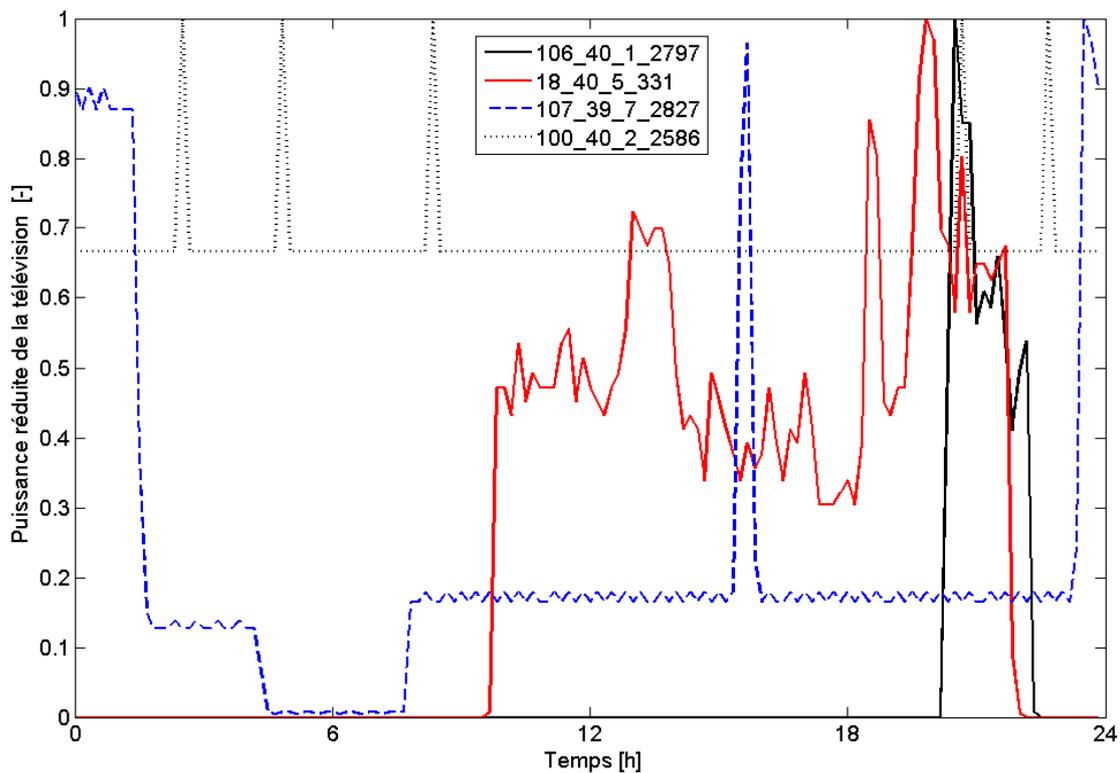


FIGURE F.4 – Profils de visionnement des objets centraux des classes de consommateurs issues de la CAH sur les résultats de l'ACP menée sur données normées

« dégradé » de la télévision (luminosité peu élevée et/ou niveau sonore faible voire mode muet). Cette hypothèse semble être cohérente avec le reste du profil journalier de visionnement : en effet de 7 h 50 à 23 h 10 (en dehors d'un pic de fonctionnement entre 15 h 30 et 15 h 50), la télévision semble être allumée dans un mode à nouveau dégradé puisque la puissance réduite avoisine les 16 à 18 %. Au delà de 23 h 10, un visionnement « classique » de la télévision a lieu puisque la puissance réduite durant cette période est supérieure à 90 %.

Enfin le profil de visionnement, objet central de la classe 4, est également de forme particulière. En effet, celui-ci correspond à une droite horizontale, à un niveau élevé de puissance réduite (autour de 66 %), droite qui est perturbée par cinq pics dans la journée répartis de façon non homogène. Dans ce cas, la seule étude du profil est insuffisante et il faut donc retourner au niveau de la courbe de charge pour se rendre compte qu'en réalité, 100\_40\_2\_2586 est le profil d'un téléviseur qui demeure en veille toute la journée ; les sauts d'appels de puissance résultant de l'échantillonnage trop faible des mesures.

Les objets centraux de cette CAH témoignent de quatre types de comportement :

- téléspectateur du soir – pas de veille ;
- consommateur « télévision bruit de fond » (continûment en journée et en soirée) – pas de veille ;
- consommateur « télévision compagnon » (appareil en marche pratiquement continue à toute heure du jour et de la nuit) – présence d'une veille ;
- pas de visionnement durant le jour observé – présence d'une veille.

La comparaison des figures 4.36 et F.4 montre que les objets centraux identifiés dans les deux études de cas sont assez proches :

- 73\_38\_6\_1774 se rapproche de 106\_40\_1\_2797 dans la manière d'utiliser la télévision en soirée

uniquement. La présence d'une veille et d'une courte période de visionnement très tôt le matin dans le premier cas étant les différences les plus notables entre ces deux objets. Soulignons également que la période d'utilisation de la télévision pour l'objet central 106\_40\_1\_2797 est incluse dans celle de l'objet central 73\_38\_6\_1774. Elle est en outre de durée inférieure ;

- 13\_40\_1\_276 et 18\_40\_5\_331 sont deux profils de téléspectateurs dont le visionnement peut être qualifié d'intensif au cours de la journée. Notons par ailleurs que comme précédemment la période de visionnement relative à 18\_40\_5\_331 est incluse dans celle correspondante à 13\_40\_1\_276 ;
- enfin dans une certaine mesure, 109\_38\_5\_2914 suggère une certaine proximité avec 100\_40\_2\_2586, une grande partie de chacun de ces objets centraux pouvant être approximés par une droite horizontale.

À l'image du travail réalisé dans la partie 4.4.4, nous allons opérer une étude approfondie de la constitution des classes. Tout d'abord, le tableau F.3 révèle les durées moyennes journalières de fonctionnement de la télévision dans chacun des modes marche/arrêt/veille<sup>4</sup>.

Classe	Durée quotidienne moyenne par mode de fonctionnement [h]		
	Mode veille	Mode arrêt	Mode marche
1	6,49	13,53	3,98
2	4,84	9,84	9,32
3	4,98	10,25	8,77
4	1,88	4,09	18,02
<b>Total</b>	<b>5,73</b>	<b>11,88</b>	<b>6,39</b>

TABLEAU F.3 – Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH menée sur les résultats de l'ACP conduite sur données normées

Le tableau F.3 nous indique que la première classe regroupe principalement des téléspectateurs qui regardent la télévision de manière « modérée » puisque la durée moyenne de fonctionnement dans cette classe est bien inférieure à la moyenne de l'échantillon (3,98 h contre 6,39 h). En revanche, les autres classes semblent rassembler les téléspectateurs assidus puisque la durée de fonctionnement moyenne de ces classes est supérieure ou égale à 8,77 h.

Comme nous l'avons suggéré en introduction de cette partie, l'analyse de courbes de charge normées devrait théoriquement rendre invisible l'influence du type de télévision sur la constitution des classes. Pour estimer la validité de cette hypothèse, nous étudions donc au travers des tableaux F.4 et F.5 respectivement la constitution des classes par type de télévision et la répartition de ces différentes gammes au sein des classes.

Classe	Composition des classes par type de télévision [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1	5,0	21,0	5,7	1,1	29,8	4,9	30,0	2,4
2	1,6	12,3	11,3	0,2	29,6	5,2	34,8	5,0
3	0,0	14,6	10,1	0,6	9,5	19,0	31,3	14,9
4	0,0	23,5	10,3	2,0	43,1	1,0	13,7	6,4
<b>Total</b>	<b>3,5</b>	<b>18,5</b>	<b>7,6</b>	<b>1,0</b>	<b>28,7</b>	<b>6,0</b>	<b>30,4</b>	<b>4,3</b>

TABLEAU F.4 – Composition des classes par type de télévision de la CAH menée sur les résultats de l'ACP sur données normées

L'analyse du tableau F.4 nous montre que la classe 1 n'attire pas spécifiquement un type de télévision de manière particulière : en effet sa composition par type d'appareil est semblable à celle de l'échantillon complet de télévisions, la différence étant de 2,5 % dans le cas le plus défavorable. En revanche, les classes 2 et 3 semblent rejeter les télévisions de petite taille au profit des grands modèles (types 36, 38, 40 et dans une moindre mesure 42 pour la classe 2 et les types 36, 39, 40 et 42 pour la classe 3). Enfin la

4. Les hypothèses de codage de chacun des modes de fonctionnement restent les mêmes que ceux évoqués page 208

quatrième classe semble privilégier les tailles intermédiaires d'écran (les types 35 et 38 constituant plus de 65 % de cette classe).

En comparant les tableaux 4.14 et F.4, il est possible de souligner que les classes formées par le même processus sur données brutes et normées aboutit à des classes semblables mais non identiques. Dans le premier cas en effet, les groupes formés sont plus tranchés en matière de composition car ils attirent à eux plus spécifiquement un ou plusieurs types d'appareils. Cette tendance est moindre dans l'étude sur données normées. Cependant, elle demeure toujours comme le suggèrent les valeurs du tableau F.5.

Classe <sup>5</sup>	Répartition des types de télévision au sein des classes [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1 (62,3)	89,1	70,8	46,4	77,1	64,7	51,3	61,4	35,3
2 (24,0)	10,9	16,0	35,4	5,7	24,8	20,8	27,5	27,6
3 (8,5)	0,0	6,7	11,3	5,7	2,8	27,1	8,7	29,4
4 (5,2)	0,0	6,6	7,0	11,4	7,8	0,8	2,3	7,6

TABLEAU F.5 – Répartition des types de télévision au sein des classes de la CAH menée sur les résultats de l'ACP sur données normées

Le tableau F.5 indique comment les télévisions se répartissent en pourcentage dans les différentes classes. La comparaison du poids de la classe (rapport de la taille de la classe sur celle de l'échantillon total) avec la répartition de chaque type de télévision au sein de chaque classe apporte des éléments supplémentaires pour la caractérisation de celles-ci. Pour la classe 1 notamment, il est plus facile d'identifier que celle-ci attire en priorité les petites et moyennes télévisions (types 34, 35, 37 et 39) au détriment des autres (types 36, 39, 40 et 42). Des constatations identiques à celles que nous avons formulées précédemment peuvent être faites au sujet des autres classes (2, 3 et 4) ; il nous semble inutile de les préciser à nouveau.

En définitive les tableaux F.4 et F.5 témoignent de l'influence toujours visible du type de télévision sur la classification de courbes de charge normées. Au delà des « simples » niveaux de puissance appelée par les différents appareils, il se peut qu'il y ait des différences en matière de forme des appels de puissance en fonction des types d'appareils (stabilité de la puissance appelée en fonctionnement, présence d'une veille, niveaux de puissance réduite par mode de fonctionnement...), formes qui seraient donc mis en lumière par la classification.

Dernière source de variabilité que nous puissions identifier sur les profils de visionnement, la composition des classes par type de jour est présentée au travers du tableau F.6.

Classe	Composition des classes par type de jour [%]						
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
1	14,6	15,4	13,9	14,8	14,8	13,9	12,6
2	14,3	12,4	15,7	13,1	12,8	13,7	17,9
3	14,3	11,3	12,2	10,1	13,4	16,7	22,0
4	12,3	12,7	12,3	14,2	14,2	20,1	14,2
<b>Total</b>	<b>14,4</b>	<b>14,2</b>	<b>14,1</b>	<b>13,9</b>	<b>14,2</b>	<b>14,4</b>	<b>14,8</b>

TABLEAU F.6 – Composition des classes par type de jour de la CAH menée sur les résultats de l'ACP sur données normées

Le tableau F.6 suggère une certaine tendance de la première classe à attirer les mardis, les jeudis et dans une moindre mesure les vendredis et en revanche à délaisser les samedis mais surtout les dimanches. À l'opposé, la deuxième classe privilégie les mercredis mais plus largement les dimanches au détriment principalement des mardis et vendredis et de manière secondaire des jeudis et samedis. La classe 3 quant à elle délaisse clairement les mardis, mercredis, jeudis et de manière moins marquée les vendredis au profit des samedis mais surtout des dimanches. Enfin la classe 4 est celle des samedis. A contrario, les

5. Les nombres entre parenthèses indiquent la proportion respective en pourcentage de chaque classe dans l'effectif total des profils de visionnement

trois premiers jours de la semaine sont abandonnés de manière relative par cette classe.

L'étude de la composition des classes suivant la durée journalière de marche de la télévision, les types d'appareil et les jours de la semaine aboutit aux constatations ci-dessous :

- la première classe rassemble des téléspectateurs assez modérés en matière de visionnement quotidien. Ils disposent d'une télévision petite à moyenne et la regardent préférentiellement lors de jours de semaine « similaires<sup>6</sup> » ;
- la seconde classe regroupe des téléspectateurs assidus, disposant d'un poste de grande taille et qui s'en servent principalement lors de journées potentielles de repos (mercredi et dimanche) ;
- la troisième classe semble être également représentative de téléspectateurs assidus, possédant une télévision de technologie récente et largement énergivore. Ceux-ci l'utilisent principalement le week-end ;
- la dernière classe témoigne d'un usage intensif du petit écran, pour des télévisions de technologie plutôt ancienne et dont l'utilisation se concentre en fin de semaine.

Avant de conclure sur cette première analyse menée sur des courbes de charge normées, précisons que nous ne faisons pas apparaître ici les barycentres des classes formées. En effet, ils ne représentent aucune réalité relative à l'usage de la télévision du fait notamment de la présence de valeurs négatives dans ces objets mathématiques.

Cette étude nous a permis de constater que malgré la normalisation des courbes de charge et donc l'utilisation de profils de visionnement, l'influence du type de télévision demeure perceptible lorsque la composition des classes est analysée. Ceci suggère que le type d'appareil a une influence sur la forme des appels de puissance. De plus, l'analyse de données normées ne semble pas être plus propice, au contraire, à l'identification de l'influence du type de jour sur la manière dont la télévision est regardée, que ne l'était le travail sur courbes de charge brutes.

La dernière étape de notre étude expérimentale a consisté en une classification directe (i. e. sans ACP préalable) des données normées. Nous présentons les résultats de ce travail dans le paragraphe F.2.

## F.2 CAH sur les courbes de charge mesurées normées

Comme nous l'avons annoncé ci-dessus, nous avons effectué une classification ascendante hiérarchique directement sur les données normées. Comme précédemment, nous utilisons la méthode d'agrégation de Ward et dans une première approche, nous ne fixons pas un nombre a priori de classes à obtenir. Par opposition à la CAH avec fixation préalable du nombre de classes que nous traiterons par la suite et que nous appellerons « CAH contrainte » nous nommons cette présente classification « CAH libre ».

### Analyse de la CAH libre

L'algorithme aboutit à trois classes dont les principales caractéristiques sont présentées dans le tableau F.7.

Classe	Objet central	Taille de la classe	Variance intra-classe	Distance au barycentre de la classe		
				minimale	moyenne	maximale
1	21_36_3_438	2861	12,158	1,862	3,374	7,672
2	18_40_5_331	929	16,762	2,397	4,057	5,984
3	100_40_2_2586	159	18,567	0,878	3,986	7,116

TABLEAU F.7 – Caractéristiques des classes de la CAH libre menée sur les données normées

6. cf. note 9 page 114

La première classe regroupe autour de 72,4 % des profils de visionnement, la seconde contient 23,6 % des éléments de l'échantillon de données alors que la dernière classe ne compte que pour 4 % de l'effectif total. De plus, il est utile de souligner que deux des trois objets centraux ont déjà été rencontrés dans les études précédentes. En effet, les objets centraux des classes 2 et 3 étaient respectivement ceux des classes 2 et 4 de la précédente CAH (classification sur les résultats de l'ACP menée sur les données normées – voir tableau F.2).

Les objets centraux des classes de cette CAH ont été représentés au sein de la figure F.5.

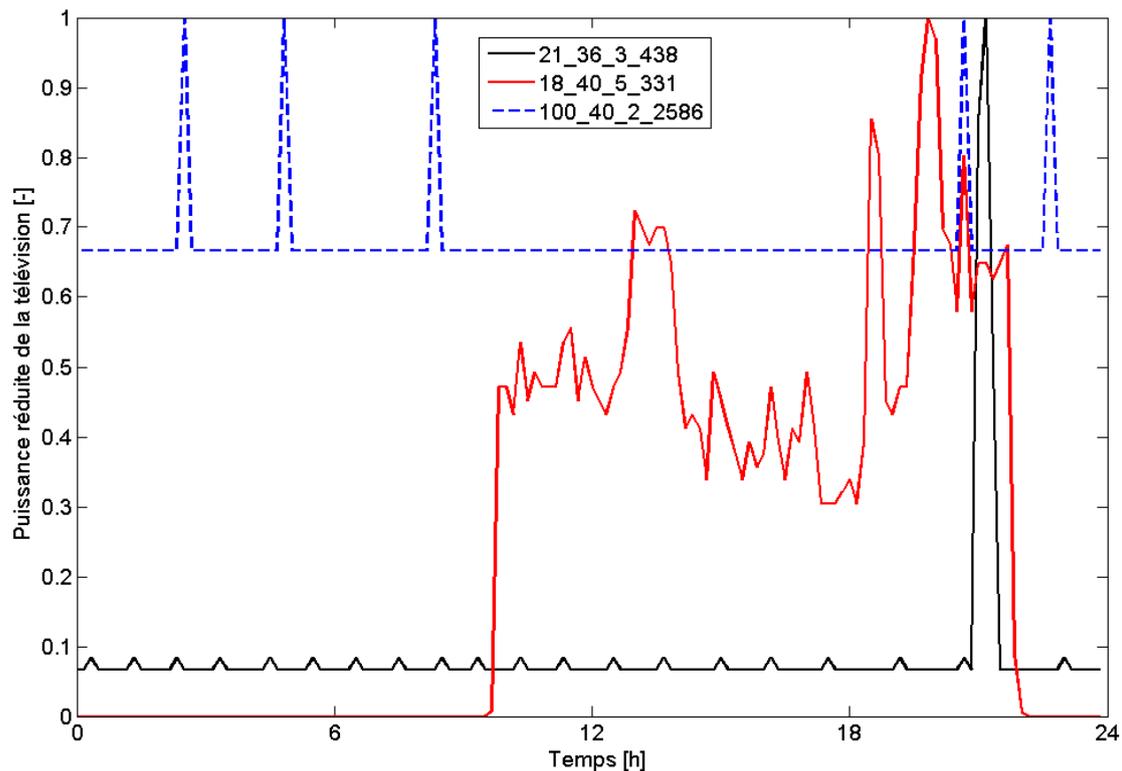


FIGURE F.5 – Profils de visionnement des objets centraux des classes de consommateurs issues de la CAH libre menée sur les données normées

La figure F.5 présente trois profils de visionnement bien distincts : nous ne rappellerons pas les remarques que nous avons faites concernant les profils 18\_40\_5\_331 et 100\_40\_2\_2586 (nous invitons le lecteur à consulter l'analyse de la figure F.4 page 252) et nous ne décrivons donc en détails que l'objet central de la classe 1. 21\_36\_3\_438 est le profil de visionnement d'un mercredi. Un bandeau de puissance réduite est remarquable sur ce profil. Celui-ci se situe autour de 6,5 % (7,2 W pour la télévision ici en jeu) mais il est légèrement bruité : quelques pics atteignant environ 8,5 % (9 W) sont détectables. Ceci semble être l'effet de l'échantillonnage choisi pour les mesures. De plus, le profil présente une courte période de visionnement de 21 h 00 à 21 h 20<sup>7</sup> durant laquelle la puissance réduite atteint son maximum c'est-à-dire 1 (ce qui correspond à une puissance de 108 W dans ce cas). Il s'agit ici du profil d'un téléspectateur actif dont la durée de visionnement est faible.

7. Sans doute le film programmé ce jour-là n'avait qu'un intérêt très limité...

Les objets centraux de la CAH libre sur données normées représentent trois types d'utilisation journalière de la télévision :

- téléspectateur du soir – présence d'une veille ;
- consommateur « télévision bruit de fond » (continûment en journée et en soirée) – pas de veille ;
- pas de visionnement durant le jour observé – présence d'une veille.

Étudions désormais la composition des classes : tout d'abord, nous les considérerons sous un aspect durée journalière de fonctionnement selon les différents modes. Cette analyse est retranscrite au travers du tableau F.8.

Classe	Durée quotidienne moyenne par mode de fonctionnement [h]		
	Mode veille	Mode arrêt	Mode marche
1	6,41	13,10	4,49
2	4,51	9,57	9,93
3	0,59	3,43	19,98
<b>Total</b>	<b>5,73</b>	<b>11,88</b>	<b>6,39</b>

TABLEAU F.8 – Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH libre menée sur les données normées

Le tableau F.8 nous montre clairement que les classes se différencient fortement en matière de durée de fonctionnement (mode marche). La première semble correspondre à nouveau à des téléspectateurs « raisonnables » en matière de durée journalière moyenne de visionnement (4,49 h dans cette classe contre 6,39 h en moyenne dans l'ensemble de l'échantillon). En revanche, il y a une gradation nette de la durée de fonctionnement (mode marche) avec le rang de la classe : la durée moyenne pour la classe 2 étant de 9,93 h et celle de la classe 3 de 19,98 h.

L'étude de la composition des classes par type de télévision et la répartition des appareils dans les classes mène à l'établissement des tableaux F.9 et F.10.

Classe	Composition des classes par type de télévision [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1	4,7	21,0	5,9	1,2	27,9	6,0	29,0	4,2
2	0,3	11,5	12,6	0,1	26,9	6,8	37,6	4,2
3	0,0	14,5	9,4	0,0	54,1	1,3	13,8	6,9
<b>Total</b>	<b>3,5</b>	<b>18,5</b>	<b>7,6</b>	<b>1,0</b>	<b>28,7</b>	<b>6,0</b>	<b>30,4</b>	<b>4,3</b>

TABLEAU F.9 – Composition des classes par type de télévision de la CAH libre menée sur les données normées

Le tableau F.9 nous indique tout d'abord que la constitution de la première classe s'éloigne assez peu de celle de l'ensemble de l'échantillon. Toutefois, les petites télévisions (types 34 et 35) semblent être privilégiées. En revanche les plus grands modèles (types 36 et 38) sont moins captés par cette classe.

Très clairement, la classe 2 est celle des téléviseurs de taille moyenne à grande en particulier les types 36 et 40, le type 39 n'étant que très légèrement sur-représenté dans cette classe par rapport à l'échantillon total. En revanche les types 34, 35, 37 et 38 sont délaissés par cette classe.

La troisième classe quant à elle exclut les petits modèles d'appareils (types 34 et 37). Ensuite elle attire modérément les types 36 et 42 mais elle plébiscite le type 38. En revanche, elle délaisse les types 35 et 39 mais plus largement le type 40. Cette classe privilégie notablement les téléviseurs de moyenne et de grande diagonale et de technologie récente mais cette préférence semble dépendre de la technologie (grande taille en CRT, taille intermédiaire en LCD et de nouveau grande taille en plasma).

8. Les nombres entre parenthèses indiquent la proportion respective en pourcentage de chaque classe dans l'effectif total des profils de visionnement

Classe <sup>8</sup>	Répartition des types de télévision au sein des classes [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1 (72,4)	97,8	82,2	56,3	97,1	70,4	72,5	69,1	70,6
2 (23,6)	2,2	14,6	38,7	2,9	22,0	26,7	29,1	22,9
3 (4,0)	0,0	3,1	5,0	0,0	7,6	0,8	1,8	6,5

TABLEAU F.10 – Répartition des types de télévision au sein des classes de la CAH libre menée sur les données normées

Du point de vue de la répartition des types de télévision sur les différentes classes, le tableau F.10 souligne l'attrait de la classe 1 pour les téléviseurs de petite à moyenne taille, les conclusions formulées au sujet du tableau F.9 se trouvent donc confirmées. De la même manière, les classes 2 et 3 captent les téléviseurs gourmands en électricité.

D'autre part, la constitution des classes suivant le type de jour peut être considérée au moyen du tableau F.11. Nous pouvons noter que la classe 1 est formée par les sept types de jour de manière relativement homogène, la plus grande différence n'étant que de 1,1 % pour le dimanche par rapport à la constitution de l'échantillon total. Cependant, nous remarquons une légère tendance à attirer les jours « scolaires » de semaine, i. e. lundi, mardi, jeudi et vendredi au détriment des autres.

À l'inverse, la constitution de la seconde classe révèle des préférences plus marquées envers les jours où les ménages ont potentiellement plus de temps libre. Dans un ordre croissant, les mercredis, les samedis et surtout les dimanches semblent être privilégiés. Pour finir, la troisième classe attire nettement les samedis (quatre points au dessus d'une distribution uniforme) et délaisse principalement les mardis et mercredis, les proportions des autres types de jour étant assez voisines de celles de la population globale de profils de visionnement.

Classe	Composition des classes par type de jour [%]						
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
1	14,8	14,9	13,9	14,3	14,6	13,7	13,7
2	13,0	12,3	14,7	12,9	12,8	15,9	18,3
3	14,5	13,2	12,6	13,2	13,8	18,2	14,5
<b>Total</b>	<b>14,4</b>	<b>14,2</b>	<b>14,1</b>	<b>13,9</b>	<b>14,2</b>	<b>14,4</b>	<b>14,8</b>

TABLEAU F.11 – Composition des classes par type de jour de la CAH libre menée sur les données normées

La CAH libre menée sur les données normées met en exergue trois classes de comportement d'utilisation de la télévision :

- la classe 1 regroupe des téléspectateurs sobres, possédant une télévision « moyenne » (en matière de taille de diagonale) et relativement indifférenciée (en type de technologie). Ils l'utilisent indistinctement quel que soit le jour de la semaine<sup>9</sup> ;
- la classe 2 rassemble des individus assidus face à leur poste de télévision. Celui-ci est de taille moyenne à grande et son usage se concentre sur des jours potentiels de repos (mercredi, samedi et dimanche) ;
- enfin la classe 3 semble représentative des comportements extrêmes<sup>10</sup> de consommateurs. Ces derniers disposent d'une télévision plutôt récente et de taille moyenne à grande et privilégient la fin de semaine pour en faire l'utilisation.

9. Les conclusions que nous pouvons faire au sujet de cette classe souffrent de l'effet de taille de celle-ci (la classe 1 regroupe près de 3 profils de visionnement sur 4 de l'échantillon – c'est la plus grande classe obtenue sur l'ensemble de nos analyses) d'où le caractère « moyen » de chacun des critères observés **sauf** en ce qui concerne la durée journalière de visionnement

10. Ce caractère est à tempérer du fait que la classe semble capter les journées au cours desquelles il n'y a pas de fonctionnement de la télévision *stricto sensu* en dehors du mode veille

Comme ce fut le cas dans la partie 4.4.5 du chapitre 4, nous représentons au travers de la figure F.6 les barycentres des classes de la CAH libre. Ces éléments sont une autre manière de résumer visuellement les classes formées par l'algorithme hiérarchique.

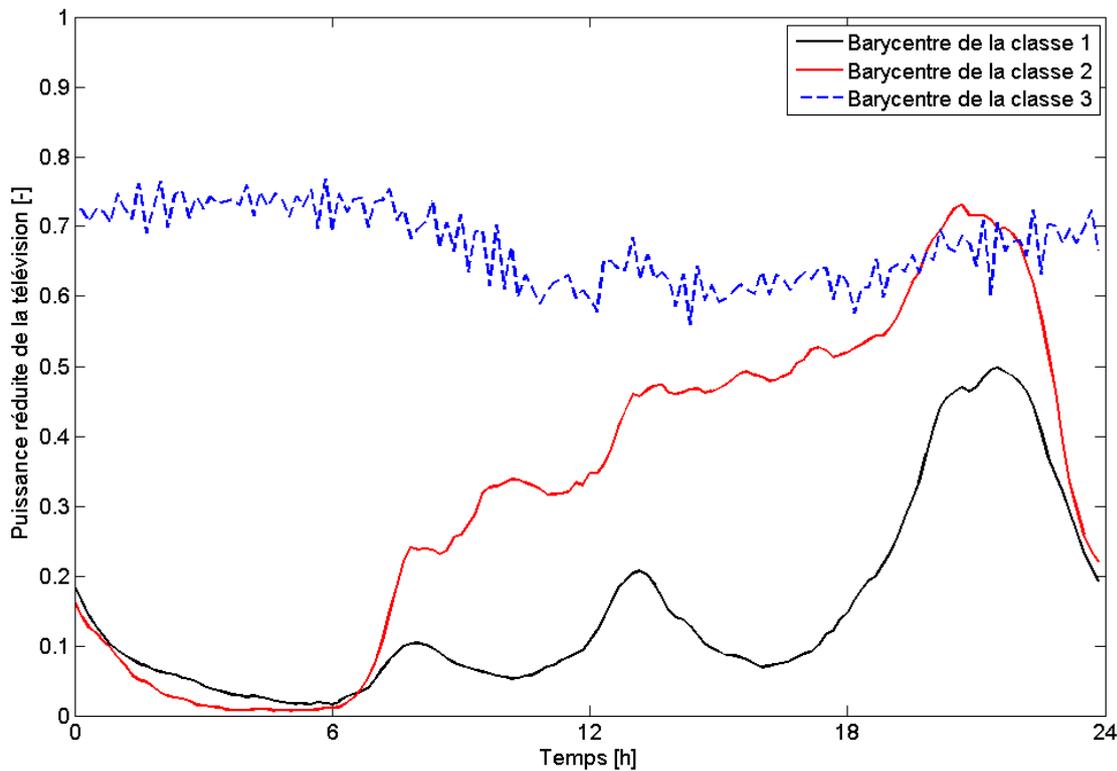


FIGURE F.6 – Barycentres des classes de la CAH libre menée sur les données normées

Le graphique F.6 laisse apparaître une ressemblance en forme mais non en niveau des barycentres des classes 1 et 2. Le profil de la première classe révèle trois périodes privilégiées de visionnement du petit écran (matin, midi et soir) avec des creux entre chacun d'eux. Des constatations similaires peuvent être formulées au sujet du deuxième profil à la différence près que les creux sont largement comblés. Dans les deux cas, le minimum du profil est atteint la nuit. Nous pouvons souligner un alignement notable de deux maximums locaux (autour de 7 h 50 et de 13 h 10) de chacun de ces profils. Les maximums globaux, atteints respectivement à 21 h 50 pour la classe 1 et à 20 h 40 pour la classe 2 sont par contre légèrement déphasés.

Ces résultats sont en cohérence avec les conclusions que nous avons formulées au sujet de la composition des classes par type de jour. En effet, la classe 1 est davantage représentative des jours scolaires de semaine alors que la seconde classe favorise les samedis, dimanches et mercredis.

Le profil de la classe 3, quant à lui, semble être constitué de deux parties : la première en forme d'arc de cercle correspond à la nuit (durant laquelle le maximum du profil est atteint) et au début de la journée (jusqu'à 10 h 50 environ). Durant cette partie, la tendance est plutôt baissière. La seconde partie du profil peut se résumer en un segment de droite de pente positive. Cet aspect complètement différent des profils des deux autres classes et les conclusions que nous avons tirées de la figure F.5 au sujet de l'objet central de la classe 3 nous poussent à croire que le barycentre est représentatif des veilles des téléviseurs.

### Analyse de la CAH contrainte

Après avoir conduit la CAH sur les données normées sans avoir imposé un nombre prédéterminé de classes, nous avons cherché à « forcer » l’algorithme hiérarchique à mettre en lumière l’influence du type de jour sur les profils de visionnement, l’idéal étant bien entendu d’avoir chaque classe représentative d’un jour-type différent. C’est donc pour cela que nous avons effectué une CAH devant aboutir à sept groupes de profils de visionnement de télévision. Une fois encore la méthode d’agrégation de Ward a été choisie.

Le tableau F.12 résume les caractéristiques générales des classes formées.

Classe	Objet central	Taille de la classe	Variance intra-classe	Distance au barycentre de la classe		
				minimale	moyenne	maximale
1	130_40_6_3798	929	7,110	1,492	2,588	5,508
2	113_40_3_3057	792	10,813	2,007	3,216	5,719
3	117_38_4_3219	899	9,958	1,038	2,990	6,654
4	85_40_6_2112	462	13,257	2,073	3,593	5,972
5	18_40_3_329	467	16,623	2,278	4,037	5,759
6	88_42_5_2233	241	16,703	2,371	3,994	6,311
7	100_40_2_2586	159	18,567	0,878	3,986	7,116

TABLEAU F.12 – Caractéristiques des classes de la CAH contrainte menée sur les données normées

Une première lecture rapide du tableau F.12 nous indique que les objets centraux des classes ne représentent pas l’ensemble des types de jour. En effet, le lundi et le dimanche n’y sont pas répertoriés. D’autre part et comme ce fut le cas pour les autres CAH, la classe 1 est la plus importante en matière d’effectif puisqu’elle représente 23,5 % de la population. La hiérarchie des classes par ordre décroissant d’importance se poursuit ensuite avec les classes 3 (22,8 %), 2 (20,1 %), 5 (11,8 %), 4 (11,7 %), 6 (6,1 %) et 7 (4 %). Soulignons que la plus petite classe a le même effectif que son équivalent de la CAH libre ce qui souligne une certaine robustesse de ce groupe de profils résiduels indépendamment du nombre de classes en sortie<sup>11</sup>. À noter également que l’évolution de la variance intra-classe (variance qui témoigne de l’homogénéité au sein du groupe d’éléments) ne suit pas exactement l’ordre établi par l’effectif des classes. Néanmoins, une tendance (attendue) à l’augmentation de cette variance avec le rang de la classe est constatée.

La figure F.7 représente les objets centraux des classes formées par la CAH contrainte. Parmi les sept profils de visionnement, celui représentatif de la dernière classe (100\_40\_2\_2586) est d’ores et déjà connu<sup>12</sup>. Celui-ci représente une consommation de veille.

D’autre part, le profil 18\_40\_3\_329 correspond à un mercredi du consommateur 18. Ce dernier était au centre de la classe 2 (18\_40\_5\_331) de la CAH libre. Ce profil très chahuté témoigne d’un visionnement intensif de la télévision puisque celle-ci n’est éteinte qu’entre 3 h 10 et 6 h 40 du matin. La puissance réduite lors du fonctionnement est assez variable entre 0,3 et 1 soulignant une fluctuation au cours de la journée des réglages de l’appareil (vraisemblablement le niveau sonore en fonction de l’attention portée à l’écran), les pointes de puissance réduite étant notables le matin, le midi et en soirée.

Le profil central de la classe 1 (130\_40\_6\_3798) est celui d’un téléspectateur actif qui n’est face à son poste qu’en soirée (de 19 h 40 jusqu’à la fin de la journée). En fonctionnement, la puissance réduite est fluctuante entre 0,4 et 1 alors qu’en dehors de ce mode, elle est identiquement nulle.

L’objet central de la classe 2 (113\_40\_3\_3057) est assez voisin du profil précédent. Il s’agit là encore d’un téléspectateur qui n’utilise son poste qu’en soirée (de 18 h 40 à 23 h 20), une très courte consommation en début de nuit (de 0 h 00 à 0 h 10) est également décelable (couchage tardif la veille). Entre ces deux périodes, l’appareil est en veille : ce mode n’appelle qu’une très faible puissance réduite de 0,002.

À l’image du profil représentatif de la classe 7, l’objet central de la classe 3 (117\_38\_4\_3219) se distingue des autres par un appel de puissance réduite relativement constant tout au long de la journée (cette puissance est légèrement bruitée et fluctue autour de 0,1) hormis lors d’une courte période de fonctionnement

11. Nous verrons par la suite qu’il ne s’agit pas de la seule caractéristique commune aux deux classes

12. Voir remarques associées à la figure F.5

comprise entre 17 h et 17 h 30 durant laquelle la puissance réduite atteint son maximum.

L'objet central de la classe 4 (85\_40\_6\_2112) constitue une sorte de profil intermédiaire entre les objets centraux des classes 1 et 5. Le consommateur 85 utilise abondamment sa télévision pendant la journée et la soirée (entre 12 h 50 et 22 h 50), la puissance réduite étant alors fluctuante entre 0,27 et 1. En dehors de cette période, celle-ci est identiquement nulle.

Enfin le profil représentatif de la classe 6 (88\_42\_5\_2233) peut être qualifié d'intermédiaire entre les objets centraux des classes 2 et 5. La première période de visionnement a lieu tôt dans la nuit (de minuit à 2 h 20) et la seconde débute en soirée (19 h) et se prolonge jusqu'au jour suivant. Lors du visionnement de l'appareil, celui-ci appelle une puissance réduite supérieure à 0,5. En dehors de ces deux périodes, la puissance réduite est identiquement nulle soulignant l'absence de veille.

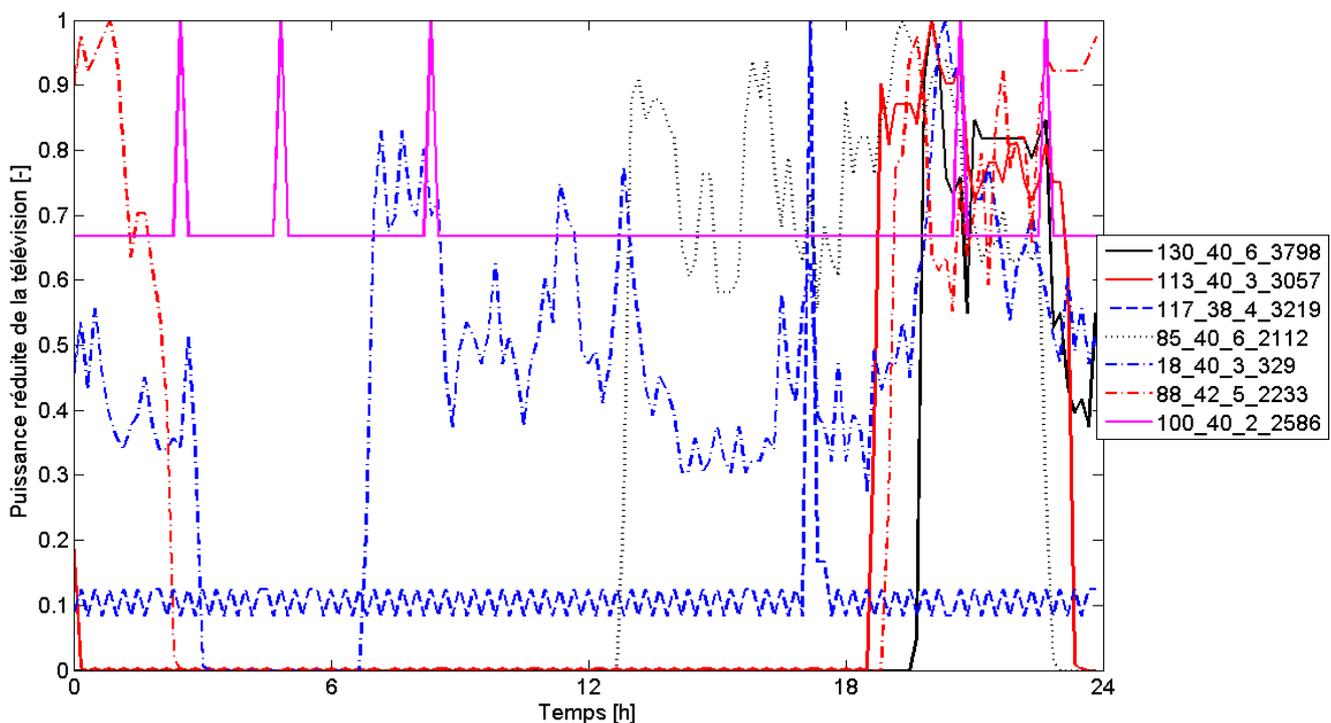


FIGURE F.7 – Profils de visionnement des objets centraux des classes de consommateurs issues de la CAH contrainte menée sur les données normées

En résumé, ces profils témoignent bien de différents comportements :

- téléspectateur du soir se prolongeant la nuit – pas de veille ;
- téléspectateur du soir sans prolongement durant la nuit – pas de veille ;
- téléspectateur du début de soirée – présence d'une veille ;
- téléspectateur de l'après-midi et de la soirée – pas de veille ;
- téléspectateur très assidu (du début du jour au milieu de la nuit puis du matin à la fin de la journée) – pas de veille ;
- téléspectateur intensif (du début du jour jusqu'au milieu de la nuit et du soir avec prolongement la nuit suivante) – pas de veille ;
- veille continue.

Le tableau F.13 renseigne les durées quotidiennes de fonctionnement par mode et pour chaque classe de

la CAH contrainte. Si comme précédemment, nous nous limitons à l'étude de la durée de fonctionnement en mode marche, nous pouvons remarquer que :

- les classes 1, 2 et 3 rassemblent des téléspectateurs dont la durée journalière moyenne d'usage de la télévision est inférieure à la durée moyenne de la population. Ceci corrobore les remarques que nous avons faites au sujet des objets centraux de ces classes (cf. figure F.7) ;
- les classes 4 et 6 regroupent des téléspectateurs qui regardent le petit écran un peu plus longtemps que la moyenne de l'échantillon ;
- enfin les classes 5 et 7 sont celles qui semblent attirer les téléspectateurs très fidèles à leur poste de télévision avec des durées journalières moyennes bien au dessus de celle de l'échantillon global. D'après ce que nous avons constaté à l'aide de la figure F.7, cette dernière remarque semble s'appliquer davantage à la classe 5 qu'à la classe 7 dont l'objet central est un profil de veille.

Classe	Durée quotidienne moyenne par mode de fonctionnement [h]		
	Mode veille	Mode arrêt	Mode marche
1	7,56	12,92	3,52
2	5,70	12,43	5,86
3	6,15	14,71	3,15
4	4,62	10,76	8,62
5	4,39	8,39	11,22
6	5,28	9,96	8,76
7	0,59	3,43	19,98
<b>Total</b>	<b>5,73</b>	<b>11,88</b>	<b>6,39</b>

TABLEAU F.13 – Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH contrainte menée sur les données normées

La composition des classes en fonction du type de télévision semble moins tranchée que pour la CAH libre (le nombre plus important de classes peut bien entendu avoir son influence). Cependant, nous pouvons remarquer au travers des tableaux F.14 et F.15 que la première classe demeure attrayante pour les télévisions de petite et de moyenne taille (types 34, 35, 40 mais surtout 37).

La seconde classe favorise les petits téléviseurs CRT (type 34) et les télévisions LCD de taille intermédiaire à grande (types 38 et 39).

La classe 3 se rapproche de la première classe car elle privilégie les petits CRT (type 34), les petits LCD (type 37) mais elle favorise avant tout les CRT de taille intermédiaire (type 35).

La classe 4 est clairement celle des grands téléviseurs CRT (type 36) et LCD (types 38 et 39).

Comme la précédente, la classe 5 favorise les grands modèles CRT (type 36) mais plus spécifiquement elle est la classe des téléviseurs plasma (types 40 et 42).

La classe 6 n'attire principalement que les téléviseurs les plus gourmands en électricité (types 39 et 42). Enfin la dernière classe attire les grands modèles de toutes les technologies (types 36, 38 et 42).

Pour résumer, comme à l'image des deux autres CAH menées sur des données normées, l'influence du type de télévision est perçue au travers des profils de visionnement. Cela nous laisse penser que la gamme d'une télévision conditionne bien évidemment le niveau de ses appels de puissance mais aussi la forme de la courbe de charge associée.

Finalement, l'analyse du tableau F.16 constitue le point central dans le cadre de cette CAH contrainte. En effet, cette dernière a été entreprise justement pour savoir si le processus de classification était capable d'identifier l'influence du type de jour sur les profils de visionnement. Nous proposons ici l'étude de cette hypothèse.

13. Les nombres entre parenthèses indiquent la proportion respective en pourcentage de chaque classe dans l'effectif total des profils de visionnement

Classe	Composition des classes par type de télévision [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1	4,6	19,4	5,4	1,6	28,2	3,1	34,0	3,7
2	6,3	12,0	5,9	0,6	36,0	8,5	27,7	3,0
3	4,7	33,1	5,0	1,4	22,9	3,1	26,4	3,3
4	0,0	12,6	12,6	0,0	38,7	9,1	25,1	1,9
5	0,6	10,5	12,6	0,2	15,2	4,5	49,9	6,4
6	0,0	12,0	11,6	0,4	19,1	19,5	24,1	13,3
7	0,0	14,5	9,4	0,0	54,1	1,3	13,8	6,9
<b>Total</b>	<b>3,5</b>	<b>18,5</b>	<b>7,6</b>	<b>1,0</b>	<b>28,7</b>	<b>6,0</b>	<b>30,4</b>	<b>4,3</b>

TABLEAU F.14 – Composition des classes par type de télévision de la CAH contrainte menée sur les données normées

Classe <sup>13</sup>	Répartition des types de télévision au sein des classes [%]							
	Type 34	Type 35	Type 36	Type 37	Type 38	Type 39	Type 40	Type 42
1 (23,5)	31,2	24,6	16,6	42,9	23,1	12,3	26,3	20,0
2 (20,1)	36,2	13,0	15,6	14,3	25,1	28,4	18,2	14,1
3 (22,8)	30,4	40,7	14,9	37,1	18,1	11,9	19,7	17,6
4 (11,7)	0,0	7,9	19,2	0,0	15,8	17,8	9,7	5,3
5 (11,8)	2,2	6,7	19,5	2,9	6,3	8,9	19,4	17,6
6 (6,1)	0,0	4,0	9,3	2,9	4,1	19,9	4,8	18,8
7 (4,0)	0,0	3,1	5,0	0,0	7,6	0,8	1,8	6,5

TABLEAU F.15 – Répartition des types de télévision au sein des classes de la CAH contrainte menée sur les données normées

Les deux premières classes sont nettement celles des jours de semaine (cette tendance est légèrement plus marquée pour la classe 1 que pour la classe 2). À l'inverse, les classes 3 à 6 favorisent largement les samedis et les dimanches. La dernière classe quant à elle est assez équilibrée en matière de type de jour mais elle attire assez spécifiquement les samedis.

Classe	Composition des classes par type de jour [%]						
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
1	15,7	16,0	15,0	16,8	14,6	10,4	11,4
2	15,8	17,2	14,9	13,5	15,2	11,9	11,6
3	12,5	12,7	12,0	13,5	15,1	17,8	16,5
4	10,8	13,4	13,9	13,9	13,4	15,6	19,0
5	15,2	11,1	15,6	12,0	12,2	16,3	17,6
6	17,0	11,6	14,1	10,4	11,2	17,0	18,7
7	14,5	13,2	12,6	13,2	13,8	18,2	14,5
<b>Total</b>	<b>14,4</b>	<b>14,2</b>	<b>14,1</b>	<b>13,9</b>	<b>14,2</b>	<b>14,4</b>	<b>14,8</b>

TABLEAU F.16 – Composition des classes par type de jour de la CAH contrainte menée sur les données normées

La CAH contrainte opérée sur les données normées aboutit à sept classes d'utilisation du petit écran :

- la première classe regroupe des téléspectateurs très modérés, qui disposent d'une télévision petite à moyenne utilisée préférentiellement lors des quatre premiers jours de la semaine ;
- la seconde classe rassemble des utilisateurs raisonnables du petit écran qui possèdent soit une petite télévision de technologie ancienne ou bien une télévision récente et de taille moyenne à grande. Celle-ci est regardée plutôt en début de semaine ;
- la classe 3 semble représentative de téléspectateurs très modérés. Ils sont équipés d'une télévision petite à moyenne et l'utilisent davantage au cours de la seconde moitié de la semaine ;
- la quatrième classe réunit des téléspectateurs assidus, ayant une télévision plutôt récente et de taille moyenne à grande. Ils la regardent principalement le week-end ;
- la classe 5 regroupe des téléspectateurs faisant un usage intensif de la télévision. Celle-ci est plutôt énergivore (technologie récente et de grande taille) et elle est utilisée prioritairement lors de journées potentielles de repos (mercredi, samedi et dimanche) ;
- la sixième classe rassemble des utilisateurs assidus du petit écran. Dans cette catégorie, il s'agit principalement de télévisions de grande taille et gourmandes en énergie qui sont regardées de manière prépondérante en tout début et en fin de semaine (lundi, samedi et dimanche) ;
- enfin il n'est pas nécessaire de décrire la dernière classe puisqu'elle correspond **exactement** <sup>14</sup> à la classe 3 de la CAH libre.

Les remarques que nous avons faites au sujet du tableau F.16 sont confirmées par l'étude de la figure F.8 qui représente les barycentres des classes de la CAH contrainte. Il est notamment remarquable de souligner l'allure semblable des profils des deux premières classes qui témoignent d'un visionnement modéré lors de la journée (hormis le pic méridien de la classe 2) et une prédominance de l'usage de la télévision en soirée. Les profils barycentriques des classes 3 à 6 sont très différents les uns des autres (utilisation plus ou moins importante pendant la nuit, heures de début du visionnement diurne non concomitantes...) soulignant une plus grande hétérogénéité des jours de fin de semaine. Enfin le barycentre de la classe 7 est identique à celui de la plus petite classe de la CAH libre. Ce profil représente les veilles.

Cette dernière CAH tente d'identifier les différents types de jour. Du fait de la structure de l'échantillon et de la présence de nombreuses personnes retraitées, ceci entache les possibilités de conclure. En effet, les personnes âgées sont enclines à regarder leur téléviseur indépendamment du jour de la semaine.

14. L'identité de ces deux classes concerne l'ensemble des paramètres descriptifs, l'objet central et le barycentre

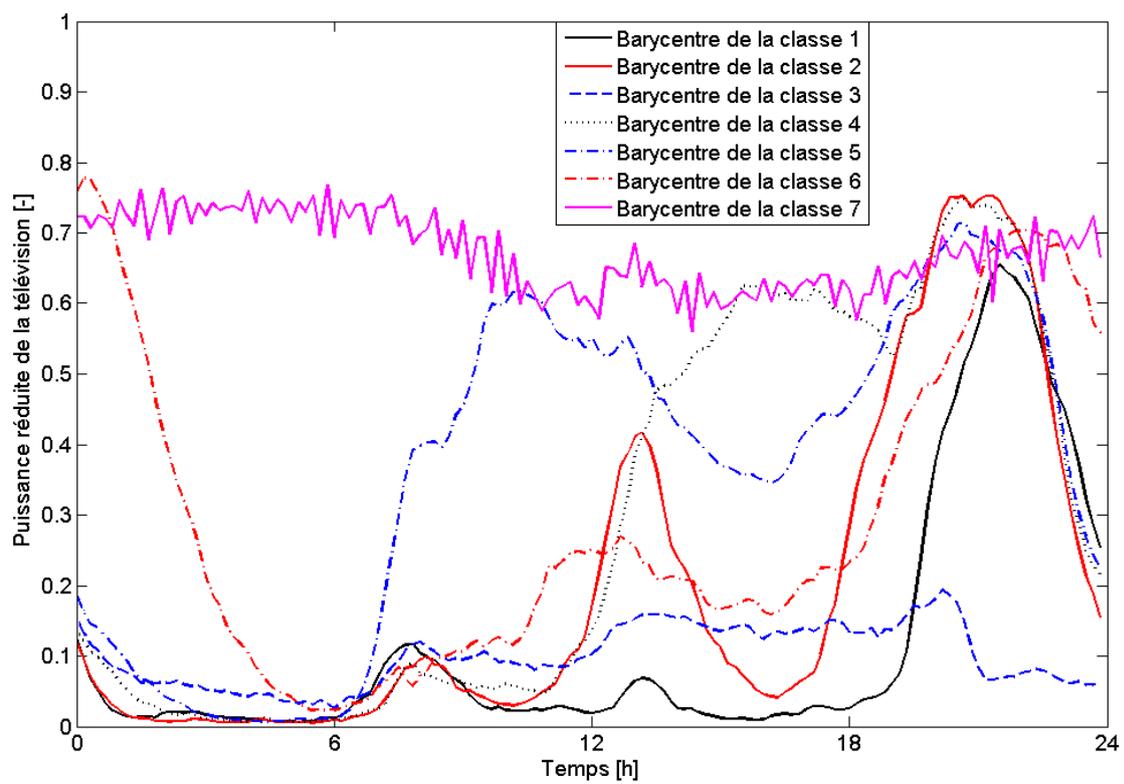


FIGURE F.8 – Barycentres des classes de la CAH contrainte menée sur les données normées

# Bibliographie

- [ACPF12] J.M. ABREU, F. CÂMARA PEREIRA et P. FERRÃO : Using pattern recognition to identify habitual behavior in residential electricity consumption. *Energy and buildings*, 49:479–487, 2012.
- [AD03] R. ANGIOLETTI et H. DESPRETZ : Maîtrise de l'énergie dans les bâtiments. Définitions, usages, consommation. Rapport technique BE 9 020, Techniques de l'ingénieur, 2003.
- [AD04] R. ANGIOLETTI et H. DESPRETZ : Maîtrise de l'énergie dans les bâtiments. Techniques. Rapport technique BE 9 021, Techniques de l'ingénieur, 2004.
- [AECL90] E. AUDUREAU, P. ENJALBERT et F. CERRO-LUIS : *Logique temporelle*. Masson, 1990.
- [AEE08] ADEME, EDF et ENERTECH : Campagne de mesures des appareils de production de froid et des appareils de lavage dans 100 logements. Projet AEE 2008. Campagne de mesures, Avril 2008.
- [AEI51] AEIC : *Manuals of procedures for load survey*. Association of Edison Illuminating Companies, 1<sup>ère</sup> édition, Février 1951.
- [AEI61] AEIC : *Manuals of procedures for load survey*. Association of Edison Illuminating Companies, 2<sup>nde</sup> édition, Mai 1961.
- [AEI01] AEIC : *Manuel. Études de profils de consommation*. Association of Edison Illuminating Companies. Load Research Committee, 2<sup>nde</sup> édition, 2001.
- [AND02] P. ANDRITSOS : Data clustering techniques. Rapport technique, University of Toronto. Department of Computer Science, 2002.
- [ASK84] J. AIGNER, D., C. SOROOSHIAN et P. KERWIN : Conditional demand analysis for estimating residential end-use load profiles. *The Energy Journal*, 5(3): 81–97, 1984.
- [ASR<sup>+</sup>09] M. M. ARMSTRONG, M. C. SWINTON, H. RIBBERINK, I. BEAUSOLEIL-MORRISON et J. MILLETTE : Synthetically derived profiles for representing occupant-driven electric loads in Canadian housing. *Journal of Building Performance Simulation*, 2(1):15–30, 2009.
- [BAR45] C. W. BARY : Coincidence factor relationships of electric service load characteristics. Rapport technique Trans. 64, Agence Internationale de l'Énergie, 1945.
- [BC94] D.J. BERNDT et J. CLIFFORD : Using dynamic time warping to find patterns in time series. In *Workshop on knowledge discovery in databases. AAAI-94*, 1994.
- [BEN82] J-P. BENZÉCRI : *Histoire et préhistoire de l'analyse des données*. Dunod, 1982.
- [BFGL92] R. BARTELS, G. FIEBIG, D., M. GARBEN et R. LUMSDAINE : An end-use electricity load simulation model : Delmod. *Utilities Policy*, 2(1): 71–82, 1992.
- [BFL<sup>+</sup>99] P. BASTARD, D. FARGUE, P. LAURIER, B. MATHIEU, M. NICOLAS et P. ROOS : *Électricité. Voyage au coeur du système*. Eyrolles, 1<sup>ère</sup> édition, Décembre 1999.

- [BFMH<sup>+</sup>94] B. BOARDMAN, D. FAVIS-MORTLOCK, M. HINELLS, K. LANE, K. MILNE, E. SMALL, V. STRANG et J. WADE : DECADE. First year report. Energy and Environment Programme. Rapport technique, Environment Change Unit. University of Oxford, 1994.
- [BRA96] P. BRAUD : Eau chaude sanitaire. Rapport technique B 9 190, Techniques de l'ingénieur, 1996.
- [BRO06] J-S. BROC : *L'évaluation ex-post des opérations locales de maîtrise de la demande en énergie. État de l'art, méthodes bottom-up, exemples appliqués et approches du développement d'une culture pratique de l'évaluation*. Thèse de doctorat, École des Mines de Paris, 2006.
- [BY02] N. BAKER et R. YAO : An integrated energy design tool. In *Proceedings of the Design with the Environment*, pages 669–673, Toulouse, France, 2002. LT Europe, Conférence PLEA.
- [CAY11] J-M. CAYLA : *Les ménages sous la contrainte carbone. Exercice de modélisation prospective des secteurs résidentiels et transports avec TIMES*. Thèse de doctorat, École Nationale Supérieure des Mines de Paris, 2011.
- [CER03] CERTU : Mémento technique du bâtiment pour le chargé d'opération de constructions publiques. Rapport technique, Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques, 2003.
- [CER10] CEREN : Secteur résidentiel. Suivi du parc et des consommations d'énergie. Rapport technique, CEREN, 2010.
- [CGLP94] A. CAPASSO, W. GRATTIERI, R. LAMEDICA et A. PRUDENZI : A bottom-up approach to residential load modeling. *IEEE Transaction on Power Systems*, 9(2):957–964, Mai 1994.
- [CGP99] E. CLARKE, O. GRUMBERG et D.A. PELED : *Model checking*. The MIT Press, 1999.
- [CRE11] CRE : Dossier d'évaluation de l'expérimentation Linky, Juin 2011.
- [CRE12] CRE : Consultation publique de la Commission de Régulation de l'Énergie du 6 mars 2012 sur la structure des tarifs d'utilisation des réseaux publics de transport et de distribution d'électricité, 2012.
- [CRR10] F. CALABRESE, J. READES et C. RATTI : Eigenplaces : segmenting space through digital signatures. *IEEE Pervasive Computing*, 9.1:78–84, 2010.
- [DEF01] DEFRA : The market transformation programme. Rapport technique, Department of the Environment, Food and Rural Affairs, 2001.
- [DUM79] J. DUMONTET : *Étude de la preuve des programmes numériques : contrôle, prévision et minimisation des erreurs de calcul*. Université Paris 6, 1979.
- [ECE08] ECEEE : Preparatory Studies for Eco-design Requirements of EuPs. LOT 13 : Domestic Refrigerators & Freezers. Final Report. Tasks 6-7. Rapport technique Tender TREN/D1/40-2005, ECEEE, Mars 2008.
- [ECE10] ECEEE : Is efficient sufficient? The case for shifting our emphasis in energy specifications to progressive efficiency and sufficiency, 2010. Disponible à l'adresse [http://www.ecee.org/sufficiency/ecee\\_Progressive\\_Efficiency.pdf](http://www.ecee.org/sufficiency/ecee_Progressive_Efficiency.pdf). Consultée le 9 août 2012.
- [ECO00] ECONOMIC AND SOCIAL DATA SERVICE : United Kingdom time of use survey, 2000. Consultée le 28 février 2012.
- [EG91] EDF-GDF : Prévention du risque électrique. Carnet de prescriptions au personnel, Janvier 1991.
- [EL 09] M. EL GUEDRI : *Caractérisation aveugle de la courbe de charge électrique. Définition, classification et estimation des usages dans les secteurs résidentiel et tertiaire*. Thèse de doctorat, Université Paris-Sud 11, 2009.

- [ENE08] ENERTECH : Mesures de la consommation des usages domestiques de l'audiovisuel et de l'informatique. Rapport technique, ADEME - EDF - Union Européenne, 2008. Accessible à l'adresse [http://www.energiepositive.info/data/document/remodece\\_rapport\\_final.pdf](http://www.energiepositive.info/data/document/remodece_rapport_final.pdf). Consultée le 10 mai 2012.
- [EPR91] EPRI : End-use load shape estimation : methods and validation. Rapport technique CU-7127, Electric Power Research Institute, 1991.
- [ErD08] ERDF : Documentation technique de référence comptage. Rapport technique ERDF-NOI-CPT\_01E, Électricité Réseau Distribution de France, Avril 2008.
- [ErD10a] ERDF DISTRIBUTION : Glossaire du site internet d'ErDF, Avril 2010. <http://www.erdfdistribution.fr/Synchrone>. Consultée le 2 avril 2010.
- [ErD10b] ERDF DISTRIBUTION : Glossaire du site internet d'ErDF, Février 2010. [http://www.erdfdistribution.fr/electricite-reseau-distribution-france/glossaire-130119.html&wrd\\_s\\_firstletter=P](http://www.erdfdistribution.fr/electricite-reseau-distribution-france/glossaire-130119.html&wrd_s_firstletter=P). Consultée le 1<sup>er</sup> février 2010.
- [GAB12] A. GRANDJEAN, J. ADNOT et G. BINET : A review and an analysis of the residential electric load curve models. *Renewable and Sustainable Energy Reviews*, 16(9):6539–6565, 2012.
- [GE04] GRD-EDF : Règles de calcul des énergies estimées. Rapport technique, Gestionnaire du Réseau de Distribution d'Électricité De France, Jan. 2004.
- [GEL96] C.W. GELLINGS : Then and now. The perspective of the man who coined the term « DSM »' *Energy Policy*, 24(4): 285–288, 1996.
- [GPR74] J. GIRAUD, P. PAMART et J. RIVERAIN : *Les nouveaux mots dans le vent*. Larousse, 1974.
- [Gra12] GRAND DICTIONNAIRE TERMINOLOGIQUE : Article « Gestion de la charge », 2012. [http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=8395422](http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8395422). Consultée le 3 juillet 2012.
- [GV11] A. GRANDJEAN et C. VUILLECARD : How does behavior impact the load curve . . . and the other way around? The TV case. *In Proceedings of the 5th BECC conference*, 2011.
- [HAM42] R. F. HAMILTON : Synthetic or equivalent load curves. *AIEE Journal*, 61(6):369–381, Juin 1942.
- [HAM44] R. F. HAMILTON : The summation of load curves. *Transactions of the American Institute of the Electrical Engineers*, 63(10):729–735, 1944.
- [HON10] W-C. HONG : Application of chaotic ant swarm optimization in electric load forecasting. *Energy Policy*, 38(10):5830–5839, 2010.
- [HUG83] T.P. HUGHES : L'électrification de l'Amérique. *Culture technique*, 10:21–41, 1983.
- [IEA11] IEA : CO<sub>2</sub> emissions from fuel combustion, 2011.
- [IEA12] IEA : Key world energy statistics, 2012.
- [INS03] INSEE : Nomenclature des professions et catégories socioprofessionnelles, 2003.
- [INS06] INSEE : Tableaux de l'économie française, 2005-2006.
- [INS07] INSEE : Insee première n°1153. Enquêtes annuelles de recensement de 2004 à 2006, Juillet 2007.
- [INS10a] INSEE : Enquête emploi du temps, 2010. Consultée le 28 février 2012.
- [INS10b] INSEE : Insee première n°1320. Projections de population à l'horizon 2060, Octobre 2010.
- [INS10c] INSEE : Tableaux de l'économie française, 2010.
- [INS10d] INSEE : Taux d'équipement des ménages pour les principaux équipements d'électricité spécifique, 2010. Données INSEE retravaillées par EDF R&D.
- [INS11a] INSEE : Insee première n°1340. Deux ménages sur trois disposent d'internet chez eux, Mars 2011.

- [INS11b] INSEE : Tableaux de l'économie française, 2011.
- [INS12a] INSEE : Définitions et méthodes. Article « solde migratoire », 2012. [http ://www.insee.fr/fr/methodes/default.asp?page=definitions/solde-migratoire.htm](http://www.insee.fr/fr/methodes/default.asp?page=definitions/solde-migratoire.htm). Consultée le 6 juillet 2012.
- [INS12b] INSEE : Définitions et méthodes. Article « solde naturel », 2012. [http ://www.insee.fr/fr/methodes/default.asp?page=definitions/liste-definitions.htm](http://www.insee.fr/fr/methodes/default.asp?page=definitions/liste-definitions.htm). Consultée le 6 juillet 2012.
- [INS12c] INSEE : Évolution de la population jusqu'en 2012, 2012. [http ://www.insee.fr/fr/themes/tableau.asp?reg\\_id=0&ref\\_id=NATnon02145](http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=NATnon02145). Consultée le 29 octobre 2012.
- [INS12d] INSEE : Évolution de la taille des ménages, 2012. [http ://www.insee.fr/fr/themes/tableau.asp?reg\\_id=0&ref\\_id=AMFd1](http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=AMFd1). Consultée le 29 octobre 2012.
- [INS12e] INSEE : Résultats du recensement de la population - 2009. Ménages par taille du ménage, sexe et âge de la personne de référence - Niveau agrégé, Août 2012. Consultée le 30 août 2012.
- [INS12f] INSEE : Tableaux de l'économie française, 2012.
- [JD06] M. JACCARD et M. DENNIS : Estimating home energy decision parameters for a hybrid energy-economy policy model. *Environmental Modeling and Assessment*, 11:91–100, 2006.
- [KP98] E. J. KEOGH et M. J. PAZZANI : An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. *In Fourth International Conference on Knowledge Discovery and Data Mining*, 1998.
- [KP99] E. J. KEOGH et M. J. PAZZANI : Scaling up dynamic time warping to massive datasets. *In Proceedings of the 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases*, 1999.
- [KP01] E. J. KEOGH et M. J. PAZZANI : Derivative dynamic time warping. *In Proceedings of the SIAM International Conference on Data Mining*, 2001.
- [LAN85] M. LANDRY : Le foisonnement en basse tension domestique. Premiers essais de quantification directe. Note interne HR 24-1653, EDF R&D, Août 1985.
- [LAN87] M. LANDRY : Le foisonnement des courbes de charge de la clientèle domestique sans chauffage principal électrique. Note interne HR/24 1995, EDF R&D, Décembre 1987.
- [LB09] M-H. LAURENT et G. BINET : Courbe de charge et effacement. Note interne, EDF R&D Enerbat, Décembre 2009.
- [LEF07] V. LEFIEUX : *Modèles semi-paramétriques appliqués à la prévision des séries temporelles. Cas de la consommation d'électricité*. Thèse de doctorat, Université Rennes 2, 2007. [http ://tel.archives-ouvertes.fr/tel-00179866](http://tel.archives-ouvertes.fr/tel-00179866).
- [LM94] J-L. LE MOIGNE : *La théorie du système général. Théorie de la modélisation*. Presses Universitaires de France, 1994.
- [MED12] MEDIAMETRIE : Médiamat mensuel - avril 2012, 2012.
- [MIN10] MINISTÈRE DE L'ÉCOLOGIE, DE L'ÉNERGIE, DU DÉVELOPPEMENT DURABLE ET DE LA MER, EN CHARGE DES TECHNOLOGIES VERTES ET DES NÉGOCIATIONS SUR LE CLIMAT : Arrêté du 26 octobre 2010 relatif aux caractéristiques thermiques et aux exigences de performance énergétique des bâtiments nouveaux et des parties nouvelles de bâtiments. Journal Officiel de la République Française, Octobre 2010.
- [MÜL07] M. MÜLLER : *Information retrieval for music and motion*, chapitre 4, pages 69–84. 2007.

- [MNP96] I. MANSOURI, M. NEWBOROUGH et D. PROBERT : Energy consumption in UK households : impact of domestic electrical appliances. *Applied Energy*, 54(3):211–285, 1996.
- [MOR04] E. MORIN : *La méthode*. Seuil, 1977-2004. 6 Tomes.
- [NC04] J-P. NAKACHE et J. CONFAIS : *Approche pragmatique de la classification : Arbres hiérarchiques, Partitionnements*. Technip, 2004.
- [ORP99] M. ORPHELIN : *Méthodes pour la reconstitution de courbes de charges agrégées des usages thermiques de l'électricité*. Thèse de doctorat, École Nationale Supérieure des Mines de Paris, Mai 1999.
- [PAT89] M. PATIN : Le foisonnement des courbes de charge en BT domestique : influence de la température. Note interne HR-24/2315, EDF R&D, Octobre 1989.
- [PCR<sup>+</sup>89] R. G. PRATT, C. C. CONNER, E. E. RICHMAN, K. G. RITLAND, W. F. SANDUSKY et M. E. TAYLOR : Description of electric energy use in single-family residences in the pacific northwest, End-use Load and Consumer Assessment Program (ELCAP). Rapport technique DOE/BP-13795-21, Bonneville Power Administration, Juillet 1989.
- [PKG11] F. PETITJEAN, A. KETTERLIN et P. GANCARSKI : A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, 44(3):678 – 693, 2011.
- [PL06] J.V. PAATERO et P.D. LUND : A model for generating household electricity load profiles. *International Journal of Energy Research*, 30(5):273–290, 2006.
- [PP80] M. PARTI et C. PARTI : The total and appliance-specific conditional demand for electricity in the household sector. *Bell Journal of Economics*, 11(1):309–321, 1980.
- [PSL<sup>+</sup>09] A. PRUDENZI, A. SILVESTRI, R. LAMEDICA, M. C. FALVO et M. REGOLI : A domestic electric load simulator including psychological aspects of demand. In *ECEEE*, Novembre 2009.
- [PSR09] A. PRUDENZI, A. SILVESTRI et M. REGOLI : A residential electric load simulator to support demand management strategies in competitive electricity markets. In *Energy Efficiency in Domestic Appliances and Lighting*, volume 1, pages 36–45, Berlin, Juin 2009.
- [REH46] A. F. REHBERG : Residential Load Study. Rapport technique 14, Edison Electric Institute, 1946.
- [REP00] REPUBLIQUE FRANCAISE : Loi n° 2000-108 du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité, 2000.
- [RT10] I. RICHARDSON et M. THOMSON : Domestic electricity demand model - simulation example. <http://hdl.handle.net/2134/5786>, Janvier 2010. <http://hdl.handle.net/2134/5786>. Consultée le 31 août 2010.
- [RTE06] RTE : Cahier des charges des publications relatives à la mesure des productions d'électricité en France, Octobre 2006. [http://clients.rte-france.com/htm/fr/vie/telecharge/cahier\\_des\\_charges\\_transparence-v\\_annexe\\_CP.pdf](http://clients.rte-france.com/htm/fr/vie/telecharge/cahier_des_charges_transparence-v_annexe_CP.pdf). Consultée le 6 janvier 2010.
- [RTE12a] RTE : Aperçu mensuel sur l'énergie électrique, Septembre 2012.
- [RTE12b] RTE : Bilan prévisionnel de l'équilibre offre-demande de l'électricité en France, 2012.
- [RTE12c] RTE : Historique des consommations journalières en puissance, Octobre 2012. [http://clients.rte-france.com/lang/fr/visiteurs/vie/vie\\_stats\\_conso\\_inst.jsp](http://clients.rte-france.com/lang/fr/visiteurs/vie/vie_stats_conso_inst.jsp). Consultée le 30 octobre 2012.
- [RTE12d] RTE : Production électrique en France, 2012. <http://clients.rte-france.com/lang/fr/visiteurs/vie/prod/accueil.jsp>. Consultée le 14 octobre 2012.
- [RTE12e] RTE : Site internet du Réseau de Transport d'Électricité, 2012. <http://www.rte-france.com/fr/>. Consultée le 14 octobre 2012.

- [RTIC10] I. RICHARDSON, M. THOMSON, D. INFELD et C. CLIFFORD : Domestic electricity use : a high-resolution energy demand model. *Energy and Buildings*, 42(10):1878–1887, 2010.
- [SAP06] G. SAPORTA : *Probabilités, analyse des données et statistique*. Technip, 2<sup>de</sup> édition, 2006.
- [SC71] H. SAKOE et S. CHIBA : A dynamic programming approach to continuous speech recognition. In *Proceedings of the Seventh International Congress on Acoustics, Budapest*, volume 3, pages 65–69, Budapest, 1971. Akadémiai Kiadó.
- [SC78] H. SAKOE et S. CHIBA : Dynamic programming algorithm optimization for spoken word recognition. *Transactions on ASSP*, 26:43–49, 1978.
- [SN06] G. SÉGUIER et F. NOTELET : *Électrotechnique industrielle*. Lavoisier, 2006.
- [STA09] R. STAMMINGER : *Synergy potential of smart domestic appliances in renewable energy systems*. Shaker, 2009.
- [STA10] STATISTICS CANADA : General Society Survey – Time Use, 2010. Consultée le 22 août 2012.
- [STO05] M. STOKES : *Removing barriers to embedded generation : a fine-grained load model to support low voltage network performance analysis*. Thèse de doctorat, De Montfort University, Janvier 2005.
- [SU09] L. G. SWAN et V. I. UGURSAL : Modeling of end-use energy consumption in the residential sector : a review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8):1819–1835, 2009.
- [THW85] K. TRAIN, J. HERRIGES et R. WINDLE : Statistically adjusted engineering models of end use load curves. *Energy*, 10(10): 1103–1111, 1985.
- [TRA92] E. TRAIN, K. : An assessment of the accuracy of statistically adjusted engineering (SAE) models of end-use load curves. *Energy*, 17(7):713–723, 1992.
- [TUF10] S. TUFFERY : *Data mining et statistique décisionnelle. L’intelligence des données*. Technip, 3<sup>ème</sup> édition, 2010.
- [UFE09a] UFE : Caractérisation du parc de production français à la pointe. Une approche européenne, Novembre 2009.
- [UFE09b] UFE : Défis climatiques. Nouveaux enjeux électriques. Synthèse des propositions de l’UFE pour mettre l’électricité au service de la croissance verte et de la réduction des émissions de CO<sub>2</sub>, Février 2009.
- [UNI73] UNIPEDE : *Étude des courbes de charge dans l’économie électrique. Manuel de théorie et de méthodologie pratique*. Comité d’études du développement des applications de l’énergie électrique, 1973.
- [UNI04] UNITED NATIONS. ECONOMIC AND SOCIAL AFFAIRS : World population to 2300, 2004.
- [UNI12] UNITED STATES CENSUS BUREAU : American Time Use Survey, 2012. Consultée le 28 février 2012.
- [WAD<sup>+</sup>85] C.J. WILLMOTT, S.G. ACKLESON, R.E. DAVIS, J.J. FEDDEMA, K.M. KLINK, D.R. LEGATES, J. O’DONNELL et C.M. ROWE : Statistics for the evaluation and comparison of models. *Journal of geophysical research*, 90(C5):8995–9005, Septembre 1985.
- [WAL82] C. F. WALKER : *A residential electrical load model*. Thèse de doctorat, University of New Hampshire, 1982.
- [WIL81] C.J. WILLMOTT : On the validation of models. *Physical Geography*, 2:184–194, 1981.
- [WIL82] C.J. WILLMOTT : Some comments on the evaluation of model performance. *Bulletin American Meteorological Society*, 63:1309–1313, 1982.

- [WLV<sup>+</sup>09] J. WIDEN, M. LUNDH, I. VASSILEVA, E. DAHIQUIST, K. ELLEGARD et E. WACKELGARD : Constructing load profiles for household electricity and hot water from time-use data. Modelling approach and validation. *Energy and Buildings*, 41(7):753–768, Février 2009.
- [WNW09] J. WIDEN, M. NILSSON, A. et E. WACKELGARD : A combined Markov-chain and bottom-up approach to modelling of domestic lighting demand. *Energy and Buildings*, 41(10):1001–1012, 2009.
- [WP85] C.F. WALKER et J.L. POKOSKI : Residential load shape modeling based on customer behavior. *IEEE Transaction on Power Apparatus and Systems*, 104(7):1703–1711, Juillet 1985.
- [WRI84] R.L. WRIGHT : Model Based Statistical Sampling for Electric Utility Load Research. Rapport technique EA 3389, RP 2279-3, EPRI, Palo Alto, 1984.
- [WW09] J. WIDEN et E. WACKELGARD : A high-resolution stochastic model of domestic activity patterns and electricity demand. *Applied Energy*, 87(6):1880–1892, Juin 2009.
- [YBM02] R. YAO, N. BAKER et M. McEVOY : A simplified thermal resistance network model for building thermal simulation : Sim2002. In *Proceedings of the eSim*, pages 135–140, Montréal, Canada, Septembre 2002.
- [YS05] R. YAO et K. STEEMERS : A method of formulating energy load profile for domestic buildings in the UK. *Energy and Buildings*, 37(6):663–671, 2005.
- [ZOU03] A. ZOUGHAIB : *High efficiency design for domestic refrigerators*. Thèse de doctorat, École des Mines de Paris, 2003.



# Liste des Figures

<b>Introduction générale</b>	<b>1</b>
1 Représentation graphique d'une courbe de charge et de l'énergie correspondante . . . . .	2
2 Courbe de charge nationale pour 4 types de jour ouvré/non ouvré et hiver/été [RTE12c] .	3
3 Évolution de la consommation d'électricité spécifique des résidences principales françaises entre 1999 et 2010 [CER10] . . . . .	4
4 Capacité installée en MW et par filière du parc français de production d'électricité en septembre 2012 [RTE12a] . . . . .	5
5 Production française d'électricité par filière en septembre 2012 [RTE12a] . . . . .	5
6 Représentation graphique d'une monotone de puissance et principaux moyens de production associés par période [UFE09b] . . . . .	5
7 Prévisions de RTE et réalisation d'une courbe de charge journalière nationale . . . . .	7
8 Finalités des consommations d'électricité du secteur résidentiel français (153,1 TWh) en 2010 à climat normal [CER10] . . . . .	11
<b>Chapitre 1</b>	<b>13</b>
1.1 Top-down vs bottom-up models . . . . .	16
1.2 Proposed classification of the load curve models . . . . .	17
1.3 Bloc diagram of Aigner et al.'s model [ASK84] . . . . .	19
1.4 Bloc diagram of Bartels et al.'s model [BFGL92] . . . . .	20
1.5 Bloc diagram of Yao and Steemers's model [YS05] . . . . .	23
1.6 Bloc diagram of Stokes's model [STO05] . . . . .	24
1.7 Bloc diagram of Paatero and Lund's model [PL06] . . . . .	28
1.8 Bloc diagram of Walker and Pokoski's model [WP85] . . . . .	32
1.9 Bloc diagram of Capasso et al.'s model [CGLP94] . . . . .	34
1.10 Bloc diagram of Armstrong et al.'s model [ASR <sup>+</sup> 09] . . . . .	36
1.11 Bloc diagram of Widén et al.'s model [WLV <sup>+</sup> 09] . . . . .	38
1.12 Bloc diagram of Widén and Wäckelgård's model [WW09] . . . . .	40
1.13 Bloc diagram of Richardson et al.'s model [RTIC10] . . . . .	43
1.14 Bloc diagram of Train et al.'s model [THW85] . . . . .	45
1.15 Cross analysis diagram of the load curve models . . . . .	52
<b>Chapitre 2</b>	<b>55</b>
2.1 Fractions par usage de la consommation énergétique finale (toutes énergies) du secteur résidentiel français (411,4 TWh) en 2010 à climat normal [CER10] . . . . .	58
2.2 Agencement des fonctions F 1 à F 4 de notre méthode de reconstitution de courbes de charge	59
2.3 Illustration des événements-clés et des périodes de simulation . . . . .	60
2.4 Schéma d'entrées/sorties de la fonction F 2 . . . . .	61
2.5 Mécanisme d'obtention des n-uplets du parc habité . . . . .	61
2.6 Représentation du découpage d'un territoire en zones géographiques . . . . .	62

2.7	Liens entre les concepts « usage », « équipement » et « équipement-type » dans le cas général (a) et sur un exemple (b) . . . . .	65
2.8	Branche chauffage de l'arbre d'analyse fonctionnelle . . . . .	67
2.9	Évolution des taux d'équipement des ménages français en réfrigérateur et congélateur entre 1954 et 2009 [INS10d] . . . . .	69
2.10	Évolution des taux d'équipement des ménages français en usages blancs (lave-linge, lave-vaisselle et sèche-linge) entre 1954 et 2009 [INS10d] . . . . .	70
2.11	Évolution du taux d'équipement des ménages français en micro-ordinateur entre 1990 et 2010 [INS10d] . . . . .	72
2.12	Évolution du taux d'équipement des ménages français en télévision entre 1954 et 2010 [INS10d] . . . . .	72
2.13	Constitution de générations d'équipements-types année par année . . . . .	83
2.14	Schéma d'entrées/sorties de la fonction F 1 . . . . .	86
2.15	Évolution du taux de pénétration de trois équipements-types dans le parc habité . . . . .	87
2.16	Schéma d'entrées/sorties de la fonction F 3 . . . . .	89
2.17	Apparence mathématique (partielle) d'un jour-type . . . . .	90
2.18	Processus d'obtention d'un scénario Time Of Use pour un n-uplet donné . . . . .	93
2.19	Méthode d'obtention du profil de charge d'un équipement-type . . . . .	97
2.20	Mise à l'échelle de la courbe de charge journalière de la machine à laver pour un n-uplet spécifié d'un parc habité . . . . .	98
2.21	Sommation des courbes de charge relatives à l'électricité spécifique pour un n-uplet donné . . . . .	99
2.22	Sommation des courbes de charge d'électricité spécifique tous n-uplets confondus . . . . .	100
2.23	Concaténation sur $n_{jr}$ jours de la courbe de charge d'électricité spécifique tous n-uplets confondus . . . . .	100
2.24	Schéma d'entrées/sorties de la fonction F 4 . . . . .	105
<b>Chapitre 3</b>		<b>109</b>
3.1	Schéma-blocs du modèle de courbe de charge développé . . . . .	111
3.2	Segmentation d'un jour-type en intervalles infra-journaliers . . . . .	113
3.3	Puissance électrique moyenne au pas 1 min d'un cycle à 30 °C et 1400 tours/min d'une machine à laver . . . . .	122
3.4	Courbe de charge journalière d'une télévision mesurée au pas 10 min issue de REMODECE . . . . .	123
3.5	Appels de puissance de plaques à induction mesurés au pas 10 s pour différentes intensités d'utilisation (identifiées de 1 à 12) . . . . .	124
3.6	Courbe de charge journalière d'un réfrigérateur mesurée au pas 10 min issue de REMODECE . . . . .	125
3.7	Courbes de charge moyennes foisonnées de réfrigérateurs, avec et sans décalage temporel à $t = 0$ s, obtenues à partir de 100 simulations . . . . .	127
3.8	Profil de disponibilité et résumé d'un jour-type <b>Semaine</b> pour un ménage-type <b>Actif</b> . . . . .	131
3.9	Exemples de profils de disponibilité pour un ménage-type <b>Actif</b> composé d'une seule personne lors d'un jour-type <b>Semaine</b> . . . . .	132
3.10	Processus de sélection de différents jeux d'instantants caractéristiques pour un ménage constitué de plus d'une personne à responsabilité . . . . .	134
3.11	Exemples de profils agrégés de disponibilité d'un ménage composé de quatre personnes à responsabilité (deux personnes de type <b>Actif</b> , un individu de type <b>ActifDomicile</b> et un membre de type <b>Inactif</b> ) au cours d'un jour-type <b>Semaine</b> . . . . .	135
3.12	Illustration des tests de compatibilité en équipements et en temps . . . . .	140
3.13	Étapes 1 à 4 du processus de recensement des instants éligibles au sein d'un intervalle pour le déclenchement de l'usage d'un équipement . . . . .	147
3.14	Étapes 1 à 3 de la mise à jour de l'intervalle développé suite au placement de l'usage d'un équipement . . . . .	148

3.15	Étapes 1 à 4 de la mise à jour d'un résumé d'intervalle suite au placement de l'usage d'un équipement . . . . .	149
3.16	Synthèse de l'algorithme de placement des équipements domestiques et mise en perspective dans son contexte . . . . .	151
<b>Chapitre 4</b>		<b>153</b>
4.1	Illustration sur une courbe de charge journalière de quelques indicateurs de statistique descriptive . . . . .	155
4.2	Représentation de plusieurs courbes de charge journalières sous forme de boîtes à moustaches	156
4.3	Illustration de l'utilisation du <i>MAE</i> et du <i>MAPE</i> pour la comparaison de courbes de charge journalières . . . . .	157
4.4	Illustration du facteur de charge calculé sur une période $\Delta t = 24$ h dans le cas d'une courbe de charge « plate » (courbe de charge 1, $F_{P_{max}} = 0,945$ ) et d'une courbe de charge « chahutée » (courbe de charge 2, $F_{P_{max}} = 0,498$ ) . . . . .	159
4.5	Illustration du coefficient de foisonnement d'une courbe de charge agrégée résultante de la somme de trois courbes de charge individuelles . . . . .	160
4.6	Exemple d'alignement optimal entre deux séries temporelles $Q$ et $V$ grâce à l'algorithme <i>DDTW</i> . . . . .	163
4.7	Exemple de chemin de déformation entre deux séquences dont l'alignement est étudié [KP99]	164
4.8	Représentation d'une séquence $\mathbf{S}$ à partir de la série temporelle $S$ . . . . .	166
4.9	Modélisation du cycle unitaire de puissance de la machine à laver pour le scénario principal	169
4.10	Modélisation du cycle unitaire de puissance de sèche-linge pour le scénario principal . . .	169
4.11	<i>CUP</i> de machine à laver et de sèche-linge issus des simulations du scénario principal . . .	170
4.12	Modification du cycle unitaire de puissance de machine à laver à iso-consommation : assimilation du $CUP_1$ au $^{inst}CUP_1$ et passage du $^{inst}CUP_1$ au $^{inst}CUP_2$ . . . . .	171
4.13	Modélisation du cycle unitaire de puissance de lave-vaisselle pour le scénario étoffé . . . .	172
4.14	Modélisation du cycle unitaire de puissance de télévision pour le scénario étoffé . . . . .	173
4.15	Exemples de quelques courbes de charge unitaires de machine à laver issues des simulations du scénario principal ASL . . . . .	174
4.16	Exemples de quelques courbes de charge unitaires de sèche-linge issues des simulations du scénario principal ASL . . . . .	174
4.17	Illustration de quelques courbes de charge moyennes foisonnées de machine à laver calculées à partir de courbes de charge unitaires simulées pour le scénario principal ASL . . . . .	176
4.18	Illustration de quelques courbes de charge moyennes foisonnées de sèche-linge calculées à partir de courbes de charge unitaires simulées pour le scénario principal ASL . . . . .	177
4.19	Illustration de courbes de charge moyennes foisonnées de machine à laver issues du $^{inst}$ scénario principal ASL* et de sa variante à iso-consommation pour quatre niveaux d'agrégation . . .	180
4.20	Matrice $M$ et chemin de déformation $W$ pour les cas <i>CDCMF</i> 100-500 (a), <i>CDCMF</i> 500-1000 (b), <i>CDCMF</i> 1000-2000 (c) et <i>CDCMF</i> 2000-6500 (d) . . . . .	182
4.21	Matrice de distances <i>DTW</i> entre courbes de charge moyennes foisonnées de machine à laver construites par sélection séquentielle des courbes de charge unitaires issues du scénario principal ASL . . . . .	184
4.22	Matrice de distances <i>DTW</i> entre courbes de charge moyennes foisonnées de sèche-linge construites par sélection séquentielle des courbes de charge unitaires issues du scénario principal ASL . . . . .	185
4.23	Matrice de distances <i>DTW</i> entre courbes de charge moyennes foisonnées de machine à laver construites par sélection aléatoire de courbes de charge unitaires issues du scénario principal ASL . . . . .	186
4.24	Matrice de distances <i>DTW</i> entre courbes de charge moyennes foisonnées de sèche-linge construites par sélection aléatoire de courbes de charge unitaires issues du scénario principal ASL . . . . .	187

4.25	Distribution des distances <i>DTW</i> entre courbes de charge moyennes foisonnées de machine à laver construites à partir de plusieurs effectifs de courbes de charge unitaires . . . . .	188
4.26	Distribution des distances <i>DTW</i> entre courbes de charge moyennes foisonnées de sèche-linge construites à partir de plusieurs effectifs de courbes de charge unitaires . . . . .	188
4.27	Relations entre la distance <i>DTW</i> et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de machine à laver obtenues par simulation (6500 <i>CDCU</i> ) . . . . .	190
4.28	Relations entre la distance <i>DTW</i> et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de machine à laver obtenues par simulation et par des mesures (respectivement 6500 et 2189 courbes de charge unitaires) . . . . .	191
4.29	Visualisation de la part des écarts d'origine temporelle et en puissance entre <i>CDCMF</i> de machine à laver issues de simulation (a) et de mesures (b) . . . . .	192
4.30	Visualisation de la part des écarts d'origine temporelle et en puissance entre <i>CDCMF</i> de machine à laver construites par différents processus à partir de <i>CDCU</i> simulées. . . . .	194
4.31	Exemple de dendrogramme . . . . .	202
4.32	Valeurs propres issues de l'ACP menée sur les données brutes et pourcentage cumulé d'explicitation de la variabilité du nuage de points . . . . .	203
4.33	Cercle de corrélation entre les variables et les facteurs principaux 1 et 2 de l'ACP sur données brutes . . . . .	205
4.34	Cercle de corrélation entre les variables et les facteurs principaux 2 et 3 de l'ACP sur données brutes . . . . .	205
4.35	Dendrogramme de la CAH conduite sur les résultats de l'ACP menée sur données brutes . . . . .	206
4.36	Courbes de charge des objets centraux des classes de consommateurs issues de la CAH sur les résultats de l'ACP menée sur données brutes . . . . .	207
4.37	Courbes de charge des objets centraux des classes de consommateurs issues de la CAH menée sur données brutes . . . . .	212
4.38	Barycentres des classes de la CAH menée sur les données brutes . . . . .	214
<b>Annexe A</b>		<b>225</b>
A.1	Décomposition d'une puissance apparente en puissance active et réactive . . . . .	227
A.2	Schéma électrique de deux charges ayant une puissance active égale mais des puissances réactives distinctes . . . . .	228
A.3	Diagramme de Fresnel associé au circuit électrique de la figure A.2 . . . . .	228
<b>Annexe C</b>		<b>235</b>
C.1	Représentation schématique d'un assemblage dans Dymola et des relations entre modèles et fichiers de données . . . . .	236
<b>Annexe D</b>		<b>239</b>
D.1	Relations entre la distance <i>DTW</i> et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de sèche-linge obtenues par simulation (6500 courbes de charge unitaires) . . . . .	239
D.2	Relations entre la distance <i>DTW</i> et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de télévision issues de la campagne REMODECE (3949 courbes de charge unitaires) . . . . .	240
D.3	Relations entre la distance <i>DTW</i> et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de machine à laver issues de la campagne REMODECE (2189 courbes de charge unitaires) . . . . .	240
D.4	Relations entre la distance <i>DTW</i> et 6 indicateurs d'intérêt sur des courbes de charge moyennes foisonnées de lave-vaisselle issues de la campagne REMODECE (1961 courbes de charge unitaires) . . . . .	241

<b>Annexe E</b>	<b>243</b>
E.1 Visualisation de la part des écarts d'origine temporelle et en puissance entre <i>CDCMF</i> de sèche-linge, construites par différents processus, à partir de 6500 <i>CDCU</i> simulées . . . . .	243
E.2 Visualisation de la part des écarts d'origine temporelle et en puissance entre <i>CDCMF</i> de télévision, construites par différents processus, à partir de 3949 <i>CDCU</i> mesurées. . . . .	244
E.3 Visualisation de la part des écarts d'origine temporelle et en puissance entre <i>CDCMF</i> de machine à laver, construites par différents processus, à partir de 2189 <i>CDCU</i> mesurées. . . . .	244
E.4 Visualisation de la part des écarts d'origine temporelle et en puissance entre <i>CDCMF</i> de lave-vaisselle, construites par différents processus, à partir de 1961 <i>CDCU</i> mesurées. . . . .	245
<b>Annexe F</b>	<b>247</b>
F.1 Valeurs propres issues de l'ACP sur données normées et pourcentage cumulé d'explicitation de la variabilité du nuage de points . . . . .	248
F.2 Cercle de corrélation entre les variables et les facteurs principaux 1 et 2 de l'ACP sur données normées . . . . .	249
F.3 Cercle de corrélation entre les variables et les facteurs principaux 2 et 3 de l'ACP sur données normées . . . . .	249
F.4 Profils de visionnement des objets centraux des classes de consommateurs issues de la CAH sur les résultats de l'ACP menée sur données normées . . . . .	251
F.5 Profils de visionnement des objets centraux des classes de consommateurs issues de la CAH libre menée sur les données normées . . . . .	255
F.6 Barycentres des classes de la CAH libre menée sur les données normées . . . . .	258
F.7 Profils de visionnement des objets centraux des classes de consommateurs issues de la CAH contrainte menée sur les données normées . . . . .	260
F.8 Barycentres des classes de la CAH contrainte menée sur les données normées . . . . .	264



# Liste des Tableaux

<b>Chapitre 1</b>	<b>13</b>
1.1 End-use families and appliances considered in Stokes's model [STO05]	26
1.2 Reading grid (first part) of the studied load curve reconstitution models	48
1.3 Reading grid (second part) of the studied load curve reconstitution models	49
<b>Chapitre 2</b>	<b>55</b>
2.1 Principe de visualisation du parc habité sous forme d'une matrice n-uplets $\times$ éléments-types	79
<b>Chapitre 3</b>	<b>109</b>
3.1 Exemple d'intervalle développé	115
3.2 Exemple de résumé d'intervalle	115
3.3 Exemple de forme condensée de l'usage d'un équipement	119
3.4 Exemple de forme développée de l'usage d'un équipement	119
3.5 Impact du décalage temporel sur la valeur des indicateurs usuels décrivant les courbes de charge moyennes foisonnées des réfrigérateurs issues de simulations	126
3.6 Liste des équipements à placer avec leurs paramètres fonctionnels pour un jour-type donné	129
3.7 Exemple de matrice de macro-compatibilité (jour-type <b>Semaine</b> )	136
3.8 Exemple d'usage d'un équipement avec des $d_{PNDU}$ et $d_{Tampou$ non nulles (8 et 32 min respectivement)	142
3.9 Configurations possibles pour le calcul du nombre de cas favorables dans un intervalle infra-journalier pour le placement d'un équipement	144
3.10 Évaluation des composantes du vecteur <b>ReportNecessaire</b> selon les configurations possibles de recherche des cas favorables dans les intervalles infra-journaliers	145
<b>Chapitre 4</b>	<b>153</b>
4.1 Valeurs pour le scénario principal des paramètres à renseigner au niveau des modèles d'équipements domestiques à cycle de fonctionnement	169
4.2 Éléments de modélisation des cycles unitaires de puissance utilisés dans le scénario principal	169
4.3 Valeurs pour le scénario étoffé des paramètres à renseigner au niveau des modèles d'équipements domestiques à cycle de fonctionnement et à états	172
4.4 Éléments de modélisation des cycles unitaires de puissance du lave-vaisselle et de la télévision utilisés dans le scénario étoffé	173
4.5 Valeurs de quelques indicateurs unaires pour des <i>CDCMF</i> simulées de lave-linge et de sèche-linge	178
4.6 Valeurs de quelques indicateurs binaires entre <i>CDCMF</i> simulées de lave-linge et de sèche-linge	178
4.7 Explicitation des calculs réalisés pour la recherche des relations entre la mesure <i>DTW</i> et les indicateurs d'intérêt entre courbes de charge moyennes foisonnées	189

4.8	Types de télévision principale dans notre analyse de courbes de charge . . . . .	198
4.9	Structure des données de courbes de charge analysées . . . . .	198
4.10	Tableau individus $\times$ variables . . . . .	200
4.11	Résultats sur les cinq premiers axes de l'ACP conduite sur les données brutes . . . . .	203
4.12	Caractéristiques des classes de la CAH sur les résultats de l'ACP menée sur les données brutes . . . . .	206
4.13	Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH menée sur les résultats de l'ACP conduite sur données brutes . . . . .	209
4.14	Composition des classes par type de télévision de la CAH menée sur les résultats de l'ACP sur données brutes . . . . .	209
4.15	Répartition des types de télévision au sein des classes de la CAH menée sur les résultats de l'ACP sur données brutes . . . . .	209
4.16	Composition des classes par type de jour de la CAH menée sur les résultats de l'ACP sur données brutes . . . . .	210
4.17	Caractéristiques des classes de la CAH sur données brutes . . . . .	211
4.18	Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH menée sur les données brutes . . . . .	213
4.19	Composition des classes par type de télévision de la CAH menée sur les données brutes . . . . .	213
4.20	Répartition des types de télévision au sein des classes de la CAH menée sur les données brutes . . . . .	214
4.21	Composition des classes par type de jour de la CAH menée sur les données brutes . . . . .	214
<b>Annexe A</b>		<b>225</b>
A.1	Paliers de tension sur le réseau de transport national . . . . .	229
<b>Annexe B</b>		<b>231</b>
B.1	Explicitations des calculs du nombre de cas favorables dans les intervalles infra-journaliers en fonction des configurations de placement des équipements . . . . .	232
B.2	Configurations possibles pour le balayage des instants éligibles dans les intervalles infra-journaliers . . . . .	233
<b>Annexe C</b>		<b>235</b>
C.1	Valeurs des paramètres à renseigner au niveau de l'assemblage ConstructionLongueur-Intervalles pour le scénario principal ASL . . . . .	236
<b>Annexe F</b>		<b>247</b>
F.1	Résultats sur les cinq premiers axes de l'ACP menée sur les données normées . . . . .	247
F.2	Caractéristiques des classes de la CAH sur les résultats de l'ACP menée sur les données normées . . . . .	250
F.3	Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH menée sur les résultats de l'ACP conduite sur données normées . . . . .	252
F.4	Composition des classes par type de télévision de la CAH menée sur les résultats de l'ACP sur données normées . . . . .	252
F.5	Répartition des types de télévision au sein des classes de la CAH menée sur les résultats de l'ACP sur données normées . . . . .	253
F.6	Composition des classes par type de jour de la CAH menée sur les résultats de l'ACP sur données normées . . . . .	253
F.7	Caractéristiques des classes de la CAH libre menée sur les données normées . . . . .	254

F.8 Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH libre menée sur les données normées . . . . .	256
F.9 Composition des classes par type de télévision de la CAH libre menée sur les données normées . . . . .	256
F.10 Répartition des types de télévision au sein des classes de la CAH libre menée sur les données normées . . . . .	257
F.11 Composition des classes par type de jour de la CAH libre menée sur les données normées	257
F.12 Caractéristiques des classes de la CAH contrainte menée sur les données normées . . . . .	259
F.13 Durées quotidiennes moyennes par mode de fonctionnement de la télévision et pour chaque classe de la CAH contrainte menée sur les données normées . . . . .	261
F.14 Composition des classes par type de télévision de la CAH contrainte menée sur les données normées . . . . .	262
F.15 Répartition des types de télévision au sein des classes de la CAH contrainte menée sur les données normées . . . . .	262
F.16 Composition des classes par type de jour de la CAH contrainte menée sur les données normées . . . . .	262



# Glossaire

## **acceptabilité humaine du pilotage**

Renseigne le niveau d'acceptation d'un individu vis-à-vis du pilotage d'un de ses équipements au sein de son logement.

## **ACP**

Analyse en Composantes Principales.

## **ADEME**

Agence De l'Environnement et de la Maîtrise de l'Énergie.

## **agrégation**

Dans le cadre d'un modèle de reconstitution de la courbe de charge électrique, l'agrégation désigne l'opération de sommer différents résultats unitaires restitués par le modèle pour une certaine période définie dans le scénario modélisé (voir également [niveau d'agrégation](#)).

## **année de calage**

Pour la modélisation de la [courbe de charge](#), l'année de calage correspond à la dernière année pour laquelle des données réalisées sont disponibles aussi bien en énergie qu'en puissance.

## **apport interne**

Il s'agit de dégagements de chaleur au sein d'un logement dont la source peut être d'origine humaine (métabolisme des occupants), technique (dissipation thermique des [équipements-types](#), pertes de production d'ECS) ou bien liés à une activité domestique (préparation culinaire par exemple).

## **Après-midi**

[Intervalle infra-journalier](#) délimité par les [instants caractéristiques](#)  $t_4$  et  $t_2$ .

## **ASL**

Au Sens Large.

## **assemblage**

Dans le logiciel dans lequel notre méthode de reconstitution de courbe de charge a été implémentée en partie, un assemblage désigne un regroupement de modèles élémentaires connectés entre eux. Un assemblage forme tout ou partie (une étape) de la simulation d'un scénario modélisé.

## **autorisation de dépassement**

Pour un équipement donné vis-à-vis d'un [intervalle infra-journalier](#) spécifié, l'autorisation de dépassement renseigne la permission pour cet appareil d'être déclenché dans cet intervalle et de voir son usage être conduit à la fois dans l'intervalle accueillant son instant de déclenchement mais aussi et en partie durant l'[intervalle infra-journalier](#) qui lui est postérieur.

## **base unique**

La base unique désigne le seul dispositif de chauffage présent dans une habitation lorsque celui-ci assure à lui-seul l'intégralité des besoins de chaleur du logement, et ce, quelles que soient les spécificités géométriques et morphologiques de l'habitation d'une part et les conditions climatiques extérieures d'autre part. Aucun autre système de production de chaleur n'a d'utilité dans ce cas.

## **BBC**

Bâtiment Basse Consommation.

**bottom-up**

Se dit d'une modélisation qui part d'une unité de base non agrégée. Dans notre cas, celle-ci correspond à l'équipement électrique domestique individuel au sein d'un ménage.

**BT**

Basse Tension.

**bâtiment-type**

Dans le cadre de la modélisation de la [courbe de charge](#), un bâtiment-type est une habitation dont les caractéristiques typologiques, dimensionnelles, thermiques et/ou techniques sont représentatives d'un segment significatif du parc de logements considéré.

**CAH**

Classification Ascendante Hiérarchique.

**capacité de l'équipement**

Valeur caractéristique d'un équipement, le plus souvent exprimée en masse ou en volume, qui renseigne la quantité maximale de service pouvant être fournie lors de son fonctionnement normal (par exemple : la quantité de linge pouvant être lavé par un cycle de machine à laver ou bien le volume utile d'un réfrigérateur).

**cas favorable**

Un cas favorable est un élément d'une [forme développée d'intervalle](#) infra-journalier, non précisément connu dans le temps, qui peut accueillir le déclenchement de l'usage d'un équipement. Un cas favorable englobe, par nature, les caractéristiques du concept de [cas possible](#).

**cas possible**

Dans le cadre de l'algorithme stochastique de placement des équipements domestiques, un cas possible désigne un élément temporel – une case – de la [forme développée d'un intervalle](#) infra-journalier.

**CDCMF**

Courbe De Charge Moyenne Foisonnée.

**CDCU**

Courbe De Charge Unitaire.

**CEREN**

Centre d'Études et de Recherches Économiques sur l'Énergie.

**CESI**

Chauffe-Eau Solaire Individuel.

**chauffage d'appoint**

Désigne l'appareil de production de chaleur supplémentaire au système de chauffage principal d'une habitation qui devient nécessaire lorsque ce dernier n'est plus capable temporairement (par exemple lors de conditions climatiques exceptionnelles) d'assurer les besoins de chaleur d'une habitation. Un chauffage d'appoint peut également s'avérer indispensable si la distribution de la chaleur au sein d'un logement ne satisfait pas, en tout endroit de l'habitation, le confort de ses occupants.

**chemin de déformation**

Dans l'algorithme de calcul de la distance *DTW* entre deux séries temporelles, le chemin de déformation correspond à une série contigüe des éléments de la matrice de distances *M*. Cette série d'éléments respecte un certain nombre de contraintes et qui forme un alignement entre les deux séries considérées.

**chronique de puisage**

Synonyme de [profil de puisage](#).

**chronique météorologique**

Données climatiques pour une ville ou une [zone géographique](#) donnée généralement disponibles à

l'échelle annuelle et au pas horaire ou tri-horaire (standard Météo-France) comprenant a minima des informations sur la température extérieure, la nébulosité et le rayonnement solaire direct et diffus.

### chronogramme probabilisé d'usage

Périodes infra-journalières au cours desquelles est autorisée, **n-uplet** par **n-uplet**, une activité domestique faisant appel à l'usage d'un ou plusieurs équipement(s) électrique(s). Ces périodes sont issues du **jour-type** simulé et sont probabilisées conformément aux caractéristiques du **n-uplet** considéré. L'agrégation des chronogrammes probabilisés d'usage, pondérés par les **poids dans l'effectif** des **n-uplets**, coïncide avec les probabilités d'activités domestiques définies au niveau du **jour-type**.

### clé de répartition

Une clé de répartition correspond au **taux d'équipement** d'une population pour un certain **équipement-type**.

### coefficient de foisonnement

Le coefficient de foisonnement est le rapport entre la puissance maximale concomitante de  $n$  consommateurs par la somme des puissances maximales individuelles (non concomitantes) de chacun des  $n$  consommateurs.

### compatibilité de l'équipement

Vis-à-vis d'un **intervalle infra-journalier**, la compatibilité de l'équipement désigne pour lui l'opportunité d'être déclenché dans cet intervalle.

### confort thermique

Le confort thermique est une notion physiologique faisant référence aux perceptions sensorielles des occupants d'un logement en matière d'humidité de l'air (hygrométrie), des **effets de paroi** et du **gradient thermique en température**. Dans les climats tempérés, le confort thermique implique le chauffage des locaux durant l'hiver et en mi-saison ainsi que leur ventilation pour assurer un taux de renouvellement d'air suffisant. En outre le rafraîchissement en période de forte chaleur et le conditionnement d'air sont deux opérations qui peuvent s'y ajouter selon la fonction des locaux considérés.

### consigne spatio-temporelle

Il s'agit d'un paramétrage des températures de consigne qui se rapportent aux modes de gestion des usages chauffage et climatisation et qui traduit le comportement du **ménage-type** considéré.

### COP

COefficient de Performance.

### couple structurant

Dans le cadre de la modélisation de la **courbe de charge**, un couple structurant désigne l'association entre un **bâtiment-type** et un **ménage-type**.

### courbe de charge

Représentation de la puissance électrique  $P$  [W] en fonction du temps  $t$  [s] (les multiples ou sous-multiples de ces deux grandeurs – watt et seconde – peuvent être utilisés). Elle traduit l'évolution au cours du temps de la consommation d'électricité.

### Courbe De Charge Moyenne Foisonnée

La Courbe De Charge Moyenne Foisonnée (**CDCMF**) de  $n$  individus est construite en sommant les  $n$  courbes de charge individuelles de chacun d'eux et en divisant le profil de charge obtenu par  $n$ . La courbe de charge moyenne foisonnée peut être considérée comme étant équivalente à la courbe de charge de l'individu marginal.

### courbe de charge normée

Courbe de charge qui engendre une consommation unitaire d'énergie. À partir d'une courbe de charge donnée, la courbe de charge normée correspondante est obtenue en divisant chaque point de la courbe de charge considérée par la valeur de la consommation énergétique engendrée par cette courbe de charge.

**Courbe De Charge Unitaire**

Courbe de charge journalière d'un [équipement-type](#) pour un individu d'un [n-uplet](#).

**courbe de production**

Traduit l'évolution dans le temps de la production d'électricité sur un réseau.

**CRT**

Cathode Ray Tube.

**CSP**

Catégorie SocioProfessionnelle.

**CUP**

[Cycle Unitaire de Puissance](#).

**Cycle Unitaire de Puissance**

Motif singulier de [courbe de charge](#) d'un [équipement-type](#) pour un [mode de fonctionnement](#) donné.

**DDTW**

Derivative Dynamic Time Warping.

**degré de simultanéité**

Le degré de simultanéité des appels de puissance électrique pour un certain appareil est équivalent à la probabilité de déclenchement de cet équipement.

**DEL**

Diode Électro-Luminescente.

**disponibilité requise**

Pour un équipement, la disponibilité requise désigne le besoin exprimé en [niveau de disponibilité](#) inhérent à l'utilisation de cet appareil.

**double base**

Lorsqu'un logement présente des spécificités (du fait par exemple de sa géométrie ou de son agencement) telles qu'un unique dispositif de chauffage ne peut assurer seul les besoins de chaleur de l'habitation, le système énergétique requis dans ce cas s'appelle une double base.

**DTW**

Dynamic Time Warping.

**durée d'usage**

Pour un équipement donné, la durée d'usage correspond à la somme des périodes de besoin de disponibilité en dehors du fonctionnement de l'équipement, des éventuelles phases d'attente avant et/ou après la marche de l'appareil et du temps de fonctionnement de l'équipement.

**durée tampon de l'usage d'un équipement**

La durée tampon de l'usage d'un équipement correspond à la première séquence continue, si elle existe, d'un besoin nul de disponibilité qui suit la partie non déplaçable de l'usage de cet équipement.

**durée à considérer**

Pour un [équipement de tête](#), il s'agit de la durée à considérer correspondante à l'usage d'un possible équipement séquentiellement lié après l'[équipement de tête](#), fonctionnant durant le même jour que ce dernier.

**ECS**

Eau Chaude Sanitaire.

**ED**

Entretien Domestique.

**EDF**

Électricité De France.

**effacement**

Politique de **MDE** qui consiste à adapter temporairement le fonctionnement d'un certain type d'appareils, d'un usage – généralement le chauffage – ou d'une catégorie de consommateurs au cours de périodes infra-journalières.

**effet boomerang**

Suite à une mesure de **MDE** et plus précisément de management de la charge électrique, l'effet boomerang désigne la possible compensation en puissance qui peut survenir à l'issue de la période de contrôle de la charge électrique (reprise de la consommation). Classiquement, vouloir écrêter une pointe de consommation électrique ne fait que la déplacer dans le temps si cela ne l'amplifie pas de surcroît.

**effet de paroi**

Malaise ressenti par l'occupant d'un logement à cause du rayonnement thermique d'une paroi de son habitation. Généralement, il s'agit d'une paroi extérieure et froide (type baie vitrée).

**effet rebond**

Suite à une mesure de **MDE**, l'effet rebond désigne un effet compensateur qui peut réduire les gains de consommation théoriques attendus par cette mesure. Il se matérialise par exemple par une augmentation de la capacité des systèmes ou par la hausse du niveau de confort souhaité.

**efficacité énergétique**

L'efficacité énergétique d'un système est le rapport entre la quantité de service rendu par son utilisation et la consommation énergétique engendrée par ce système lors de son fonctionnement.

**élément-type**

Un élément-type désigne indistinctement un **bâtiment-type**, un **ménage-type** ou encore un **équipement-type**.

**éligibilité d'un couple structurant**

Propension d'un couple **bâtiment-type/ménage-type** à être doté d'un équipement donné. Cette propension est évaluée à partir de règles logiques.

**éligibilité d'un intervalle infra-journalier**

L'éligibilité d'un intervalle infra-journalier vis-à-vis d'un équipement donné équivaut à la possibilité pour lui d'accueillir son instant de déclenchement.

**ENR**

Énergies Renouvelables.

**équipement auto-lié**

Au sein d'un ménage donné, un équipement est dit auto-lié si d'une part un seul appareil de ce type équipe le foyer considéré et si d'autre part plusieurs épisodes de fonctionnement de cet équipement ne peuvent se faire que séquentiellement.

**équipement de tête**

Lorsqu'un équipement est lié avec d'autres, celui-ci est dit de tête s'il conditionne le fonctionnement de ces autres appareils.

**équipement lié**

Des équipements sont dits liés s'il existe une dépendance ou a minima une relation de temporalité dans le cadre de l'usage que l'on fait de chacun d'eux.

**équipement-type**

Appareil domestique dont les caractéristiques techniques et/ou de fonctionnement sont représentatives d'une famille d'équipements du marché à une période donnée.

**équipement-type fixe**

**Équipement-type** « fixé » au bâtiment qui assure un des usages suivants : chauffage, climatisation ou **ECS**. Il est donc intimement lié au bâtiment dans lequel il est installé.

**équipement-type nomade**

Équipement-type indépendant du logement dans lequel il est installé. Il peut par exemple accompagner les occupants d'un foyer lors d'un déménagement. La possession d'un tel équipement est davantage fonction des occupants d'un ménage que ne l'est celle d'un [équipement-type fixe](#).

**ErDF**

Électricité Réseau Distribution de France.

**état initial**

Le résumé d'un [intervalle infra-journalier](#) est dit dans son état initial lorsqu'aucun [événement](#) n'est placé dans l'[intervalle infra-journalier](#) correspondant.

**événement**

Dans le cadre de la modélisation des concepts de la méthode de reconstitution de la courbe de charge, un événement correspond au couple d'informations [instant de déclenchement, [signature d'équipement](#)] relatif à l'utilisation d'un appareil électrique domestique.

**événement-clé**

Il s'agit d'une rupture dans le [parc habité](#) modélisé. Il peut affecter n'importe lequel (lesquels) de ses constituants ([éléments-types](#)) que ce soit directement ou non.

**facteur de charge**

Le facteur de charge (de la puissance maximale) est le rapport entre la puissance moyenne et la puissance maximale d'une courbe de charge.

**facteur de puissance**

Le facteur de puissance est égal au rapport entre la puissance active et la puissance apparente. Il est un indicateur de la qualité du courant circulant dans un réseau électrique.

**filtre**

Dans le cadre de la modélisation des [équipements-types](#), un filtre représente un critère ou une caractéristique au moyen duquel des distinctions et des rapprochements entre équipements peuvent être opérés de manière à aboutir à une [modélisation générique](#) par famille d'[équipements-types](#).

**flux net**

Étant donné un périmètre d'étude et un ensemble d'objets considéré, le flux net de ces objets entre deux instants distincts  $t$  et  $t + \Delta t$  par rapport au périmètre d'étude correspond à la somme algébrique des objets entrant dans le périmètre (comptés positivement) et des objets sortant de ce périmètre (comptés négativement).

**foisonnement**

Il s'agit d'un phénomène naturel se produisant au sein d'un réseau électrique et qui traduit le degré de concomitance (et de non-concomitance) de la consommation d'électricité entre différents usagers.

**forme condensée de l'usage d'un équipement**

Dans le cadre de l'algorithme de placement des équipements, la forme condensée de l'usage d'un équipement désigne un tableau contenant les informations relatives aux périodes  $d_\beta$ , c'est-à-dire leur durée respective et le niveau de responsabilité correspondant dans chaque période.

**forme développée d'intervalle**

La forme développée d'un [intervalle infra-journalier](#) est un tableau comprenant « [longueur d'intervalle](#) » éléments qui renseignent le [niveau de disponibilité](#) au sein de cet [intervalle infra-journalier](#).

**forme développée de l'usage d'un équipement**

Dans le cadre de l'algorithme de placement des équipements, la forme développée de l'usage d'un équipement désigne le tableau construit en dilatant les informations de la [forme condensée de l'usage d'un équipement](#).

**gestion de la charge**

Ensemble des actions prises par un distributeur d'énergie pour modifier la demande en fonction de contraintes financières et techniques liées à la production et à la distribution de l'énergie.

**gradient thermique de consommation énergétique**

Variation de la consommation électrique d'un parc suite à l'évolution de la température extérieure. Ce gradient s'exprime en W/ °C.

**gradient thermique en température**

Dans une zone habitée, il s'agit de la différence de température existante entre le plafond et le sol par manque de brassage d'air.

**HC**

Hygiène Corporelle.

**HT**

Haute Tension.

*IA*

Index of Agreement.

*IAC*

Intervalle À Considérer.

*IC*

Intervalle Choisi.

*IE*

Intervalle Éligible.

*IL*

Intervalle Limite.

**indicateur binaire**

Un indicateur binaire désigne une grandeur statistique qui caractérise des courbes de charge prises deux à deux.

**indicateur unaire**

Un indicateur unaire désigne une grandeur statistique qui caractérise des courbes de charge prises individuellement.

**instant caractéristique**

Instant de la journée qui délimite deux [intervalles infra-journaliers](#) consécutifs.

**instant éligible de déclenchement**

Dans le cadre de l'algorithme stochastique de placement des équipements domestiques, un instant éligible de déclenchement pour un équipement donné désigne un élément de la [forme développée d'intervalle](#) infra-journalier précisément connu dans le temps et qui peut accueillir le déclenchement d'un appareil ménager. Un instant éligible de déclenchement englobe les caractéristiques du concept de [cas favorable](#) et le dépasse.

**intervalle infra-journalier**

Période au sein d'un [jour-type](#) délimitée en durée par deux instants caractéristiques et comportant le [niveau de disponibilité](#) des occupants d'un [ménage-type](#).

**intervalle infra-journalier éligible**

Dans le cadre de l'algorithme stochastique de placement des équipements, un intervalle infra-journalier éligible pour un équipement donné désigne un intervalle pouvant accueillir l'instant de déclenchement de l'usage de cet équipement.

**intervalle limite**

Dernier [intervalle infra-journalier](#) d'un [jour-type](#) dans lequel un équipement peut être utilisé.

**intervalle à considérer**

Dans le cadre de la modélisation d'un équipement, l'intervalle à considérer témoigne des contraintes éventuellement existantes pour le déclenchement d'un appareil lors d'une journée considérée.

**IT**

Intervalle Traité (considéré).

**jeu initial d'instants caractéristiques**

Dans la modélisation des [profils de disponibilité](#), un jeu initial d'instants caractéristiques correspond à la première série d'instants  $t_\alpha$  sélectionnés (plusieurs séries devant être choisies) entre les bornes paramétrées au sein du modèle du [jour-type](#).

**jeu supplémentaire d'instants caractéristiques**

Dans le cadre de la génération de profils élaborés de disponibilité, les jeux supplémentaires d'instants caractéristiques désignent les séries d'instants notés  $t_{\alpha+}$  qui sont choisis dans des zones bornées en partie par les instants caractéristiques du [jeu initial](#) et par les limites paramétrées au niveau du modèle du [jour-type](#).

**jour-type**

Un jour-type est caractérisé par des périodes infra-journalières où sont renseignées les probabilités d'activités domestiques ou d'usages, ceci de manière transverse à l'ensemble des ménages de la population considérée.

**LA**

L'Avage.

**LBC**

Lampe Basse Consommation.

**LC**

Logement Collectif.

**LCD**

Liquid Cristal Display.

**LED**

Light-Emitting Diode.

**liaison**

Une liaison entre deux équipements représente un possible lien logique ou une dépendance qui conditionne la temporalité du fonctionnement de l'un de ces appareils par rapport à l'autre.

**lien hypertexte**

Ce qui vous permet d'arriver jusqu'ici.

**lien simultané**

Pour deux équipements liés, cette modalité de [liaison](#) implique que le fonctionnement de ces appareils est généralement concomitant.

**lien séquentiel**

Pour deux équipements liés, cette modalité de [liaison](#) implique que le fonctionnement d'un de ces appareils précède la marche de l'autre équipement.

**longueur d'intervalle**

La longueur d'un [intervalle infra-journalier](#) est égale à la division entière entre d'une part la différence entre les bornes de l'intervalle ( $t_{Final} - t_{Initial}$ ) et d'autre part la résolution temporelle de la modélisation/simulation.

**LRCM**

Loisirs, Repos, Communications et Mobilité.

**macro-compatibilité d'une tâche domestique**

La macro-compatibilité d'une tâche domestique pour un [intervalle infra-journalier](#) donné correspond au statut d'autorisation/d'interdiction de la [macro-tâche](#) face à ce même intervalle.

**macro-tâche**

Une macro-tâche désigne une activité domestique considérée à l'échelle macroscopique c'est-à-dire en considérant tous les ménages du territoire étudié.

**MAE**

Mean Absolute Error.

**MAPE**

Mean Absolute Percentage Error.

**Matin**

**Intervalle infra-journalier** délimité par les **instants caractéristiques**  $t_1$  et  $t_3$ .

**Matrice de MAcro-Compatibilité**

Une Matrice de MAcro-Compatibilité (*MMAC*) contient les permissions (interdictions et autorisations) relatives à la conduite des **macro-tâches** domestiques qui sont définies a priori et à l'échelle agrégée.

**Matrice de MIcro-Compatibilité**

Une Matrice de MIcro-Compatibilité (*MMIC*) contient les permissions (interdictions et autorisations) relatives à la conduite des **micro-tâches** domestiques c'est-à-dire à l'échelle du ménage individuel.

**Maîtrise de la Demande en Énergie**

« La Maîtrise de la Demande en Énergie (**MDE**) désigne les actions conduites par les pouvoirs publics et par les producteurs et/ou distributeurs d'énergie, destinées à inciter et parfois à encadrer la manière avec laquelle les usagers d'un secteur d'activités utilisent et/ou consomment l'énergie. Dans ce cas on va notamment chercher à modifier les appels de puissance des énergies non stockables et qui doivent être distribuées par réseau directement aux points de consommation » [BRO06].

**MDE**

Maîtrise De la demande en Énergie.

**MI**

Maison Individuelle.

**micro-compatibilité d'une tâche domestique**

La micro-compatibilité d'une tâche domestique pour un **intervalle infra-journalier** donné correspond au statut d'autorisation/d'interdiction de la **micro-tâche** face à ce même intervalle.

**micro-tâche**

Une micro-tâche désigne une activité domestique considérée à l'échelle individuelle c'est-à-dire en ne considérant qu'un seul ménage.

**Midi**

**Intervalle infra-journalier** délimité par les **instants caractéristiques**  $t_3$  et  $t_4$ .

**MMAC**

Matrice de MAcro-Compatibilité.

**MMIC**

Matrice de MIcro-Compatibilité.

**mode de fonctionnement**

Paramétrage d'un **équipement-type** programmé par l'utilisateur suivant lequel ce dernier fonctionne de manière discrète (choix du type de cycle d'une machine à laver : 30 °C, 60 °C, 90 °C...) ou continue (réglage de la température de consigne d'un réfrigérateur, de la luminosité et du niveau sonore d'une télévision...).

**mode de gestion**

Relativement aux appareils de chauffage/climatisation, un mode de gestion correspond au niveau d'intensité programmé par l'utilisateur pour satisfaire le niveau de confort souhaité. Pour le chauffage par exemple, les modes de gestion classiques sont « confort », « réduit » et « hors-gel ».

**modélisation générique**

À l'inverse de la modélisation exhaustive qui cherche à représenter chaque élément de manière spécifique en prenant en compte l'ensemble de ses particularités, la modélisation générique a pour objectif de réduire le nombre de modèles nécessaires à la représentation d'un système. Pour cela des similarités entre les entités à modéliser sont recherchées. Des [filtres](#) sont alors à choisir pour définir quels sont les points d'intérêt principaux des objets à modéliser.

**monotone de puissance**

Représentation ordinale et non chronologique des puissances électriques appelées par un parc au cours d'une période donnée.

*MTP*

Micro-Tâche Parente.

**multi-équipement**

Possession de deux appareils ou plus de même nature (soit dans la résidence principale, soit dans la résidence secondaire) et assurant le même usage.

**ménage agrégé**

Un ménage agrégé désigne un foyer considéré de manière indivisible c'est-à-dire que les membres qui le constituent ne sont pas pris en compte de manière individuelle.

**ménage-type**

Foyer aux caractéristiques démographiques, économiques, sociales et comportementales représentatif d'un ensemble significatif de ménages de la population étudiée.

**n-uplet**

Un n-uplet est un [ménage-type](#) habitant dans un [bâtiment-type](#) et disposant d'un ensemble d'[équipements-types](#). Un n-uplet correspond également à un [triplet](#) doté d'[équipements-types fixes](#) et d'[équipements-types nomades](#).

**niveau d'agrégation**

Le niveau d'agrégation d'une [Courbe De Charge Moyenne Foisonnée](#) équivaut à l'effectif de [courbes de charge unitaires](#) qui a été utilisé pour construire cette dernière.

**niveau de charge**

Taux d'utilisation en pourcentage de la capacité ou de la puissance nominale d'un équipement électrique. Un niveau de charge égal à 1 indique que l'appareil est utilisé au maximum de sa capacité, respectivement à sa puissance nominale.

**niveau de disponibilité**

Indicateur de mobilisation d'un ménage considéré de manière agrégée ou d'une personne individuelle vis-à-vis d'une activité domestique ou d'un usage donné. Il varie continûment entre 0 (le ménage considéré de manière agrégée ou la personne est occupé(e) par une mono-tâche) et 1 (le ménage considéré de manière agrégée/l'occupant est alors parfaitement disponible pour tout usage). De plus, la valeur -1 représente l'état de sommeil. Lorsque l'on considère chaque membre d'un ménage de manière individuelle, le niveau de disponibilité est défini pour chaque membre d'un [ménage-type](#) reconnu responsable, c'est-à-dire pouvant agir sur les équipements domestiques électriques.

**niveau de responsabilité**

Dans le cadre de la modélisation du [ménage-type](#), le niveau de responsabilité d'un membre du ménage détermine la possibilité ou non pour cet individu de déclencher des appareils domestiques.

**nomenclature**

Ensemble des éléments décrivant un [équipement-type](#) de façon à l'identifier de manière univoque.

**NOSE**

Nord-Ouest – Sud-Est.

**Nuit**

[Intervalle infra-journalier](#) délimité par les [instants caractéristiques](#)  $t_{End}$  et  $t_{StartBis}$ .

**NVF**

Normalized Variation Factor.

**opportunité technique de pilotage**

L'opportunité technique de pilotage d'un [équipement-type](#) est la possibilité matérielle pour lui d'une part d'être commandé autrement que par l'intervention humaine et d'autre part le cas échéant, elle indique les différentes potentialités dynamiques de commande offertes par cet équipement.

**PAC**

Pompe À Chaleur.

**paramètre de gestion**

Caractéristique intrinsèque et extrinsèque d'un [équipement-type](#) qui influence sensiblement la manière avec laquelle cet appareil peut être utilisé.

**paramètre fonctionnel**

Caractéristique intrinsèque d'un [équipement-type](#) qui influence de manière prépondérante ses appels de puissance lors de son fonctionnement.

**parc de référence**

Considérons une année  $n$  dans le cadre de la méthode de reconstitution des appels de puissance du secteur résidentiel. Le parc de référence pour cette année  $n$  correspond au dernier [parc habité](#) précédemment construit c'est-à-dire le [parc habité](#) établi à l'année  $n - 1$  puisque ce dernier est revu chaque année.

**parc habité**

Dans le cadre de la modélisation de la [courbe de charge](#), le parc habité correspond à un ensemble de [n-uplets](#) affectés de leur poids dans l'effectif.

**partie non déplaçable de l'usage d'un équipement**

La partie non déplaçable de l'usage d'un équipement correspond à la première séquence continue, si elle existe, d'un besoin non nul de disponibilité issue de la forme développée de l'usage d'un équipement.

**PCR**

Préparation Culinaire et Repas.

**PERCHE**

Pompe à chaleur En Relève de CHaudière Existante.

**performance thermique d'enveloppe**

La performance thermique d'enveloppe d'un bâtiment renseigne de la qualité de son isolation thermique avec l'environnement extérieur. Étant donné que le chauffage est l'usage domestique le plus consommateur d'énergie à l'échelle annuelle dans les zones climatiques tempérées, la performance thermique de l'enveloppe d'un logement conditionne grandement la [performance énergétique du bâtiment](#).

**performance énergétique du bâtiment**

La performance énergétique du bâtiment correspond à la consommation d'énergie d'une construction résidentielle pour une sélection d'usages et exprimée en kWh/m<sup>2</sup>.an.

**perturbation**

Au cours du fonctionnement d'un équipement domestique, une perturbation désigne la modification d'origine humaine ou environnementale des conditions d'utilisation de cet appareil. Comme son nom le suggère, la perturbation a une influence significative sur les appels de puissance de l'équipement qui la subit.

**pilotage**

Action de commander à distance l'instant de déclenchement d'un équipement.

**placement**

Le placement d'un **équipement-type** au sein d'un **jour-type** consiste à lui chercher un instant de déclenchement pour son usage. L'utilisation de cet équipement doit se faire de manière réaliste et compatible avec les contraintes du scénario modélisé (durée d'usage, équipements liés...).

**PNDU**

Partie Non Déplaçable de l'Usage (d'un équipement).

**poids dans l'effectif**

Proportion numérique d'un **n-uplet** en matière d'individus par rapport à l'ensemble de la population étudiée.

**poids de représentativité**

Proportion en matière d'effectif d'un **couple structurant** du **parc habité** par rapport à l'ensemble des **couples structurants**.

**pointe concomitante**

La pointe concomitante sur un réseau électrique correspond à la puissance maximale (maximum de consommation), pour une période de temps donnée, d'un ensemble d'utilisateurs raccordés à ce réseau. La pointe de puissance peut être relative à un certain équipement, à un usage donné ou bien encore à l'ensemble des besoins en électricité des utilisateurs (i. e. tous types d'usages confondus).

**pointe de puissance**

Sur un réseau électrique, la pointe de puissance désigne la consommation maximale d'électricité pour une période donnée (jour, mois, année...).

**poste de consommation**

Désigne un usage ou un regroupement d'usages domestiques.

**profil agrégé de disponibilité**

Le profil agrégé de disponibilité correspond à la connaissance à l'échelle du ménage de la disponibilité cumulée de chacun de ses membres. Il est calculé conformément aux **règles de cumul (ou d'additivité)**.

**profil d'occupation**

Évolution au cours du temps du nombre d'occupants d'un ménage présents à leur domicile. Généralement la seule présence au domicile n'est pas considérée telle quelle : on lui préfère par exemple la présence des occupants à leur domicile dans un état « éveillé » c'est-à-dire un état compatible avec l'utilisation d'appareils domestiques.

**profil de charge**

Pour un **équipement-type** donné, il s'agit de la courbe de charge moyenne foisonnée de cet appareil, tous modes de fonctionnement confondus, pour un individu équipé d'un n-uplet lors d'un jour simulé donné.

**profil de consommation**

Il s'agit de la courbe de charge représentative d'un type de clients équipés de compteur à index qui sont raccordés au réseau de distribution d'électricité. Cette courbe de charge est construite à partir de coefficients hebdomadaires, journaliers et demi-horaires.

**profil de disponibilité**

Évolution sur une période donnée (typiquement au cours d'un **jour-type**) du **niveau de disponibilité** d'un membre d'un **ménage-type**.

**profil de puisage**

Besoin en ECS exprimé en volume par unité de temps (le plus souvent en L/h) et délivré à une température de référence appelée **température de puisage** (par exemple 60 °C).

**profil de visionnement**

Pour une télévision, le profil de visionnement correspond à l'évolution journalière de la puissance

réduite de cet appareil (la puissance de référence dans notre cas est la puissance maximale journalière). Cette construction permet ainsi de comparer des comportements de téléspectateurs au travers de leurs courbes de charge, quel que soit le type de télévision qu'ils possèdent.

**profil exogène**

Un profil exogène désigne la connaissance de l'évolution au cours du temps et sur une durée donnée (par exemple et respectivement à l'échelle de l'heure et pendant une journée) de la valeur d'une variable d'intérêt. Dans le contexte de la modélisation de la [courbe de charge](#) électrique, des profils exogènes de disponibilité, de pilotage ou encore de tarif peuvent présenter un intérêt notable.

**profilage**

Méthode employée par le gestionnaire du réseau de distribution d'électricité dont le but est de calculer les consommations et/ou les productions des clients non sujets au comptage à courbe de charge.

**puissance active**

La puissance active est égale à la valeur moyenne de la [puissance instantanée](#) d'un courant électrique sinusoïdal.

**puissance apparente**

La puissance apparente d'un courant électrique est égale au produit de la tension efficace et du courant efficace.

**puissance de veille**

La puissance de veille d'un équipement est la puissance appelée par cet appareil lorsqu'il est en mode de fonctionnement veille.

**puissance instantanée**

La puissance instantanée d'un courant électrique correspond au produit des grandeurs instantanées tension et courant.

**puissance nominale**

Pour un équipement domestique, il s'agit de la puissance (thermique, mécanique ou électrique) définie et garantie par le constructeur comme pouvant être fournie (ou consommée) en marche continue tout en respectant les rendements utiles associés.

**puissance réactive**

La puissance réactive traduit les échanges d'énergie à valeur moyenne nulle entre une source et une inductance ou une capacité [SN06].

**puissance réduite**

La puissance réduite reçue ou fournie par un système à l'instant  $t$ ,  $P_r(t)$ , est égale au quotient de la puissance instantanée à l'instant  $t$ ,  $P(t)$ , divisée par la valeur de la puissance choisie comme référence (par exemple puissance moyenne ou maximale sur une période donnée).

**période**

Au sein d'une simulation de plusieurs années, une période est constituée de toutes les années comprises entre deux [événements-clés](#) consécutifs.

**REMODECE**

REsidential MOnitoring to Decrease Energy use and Carbon emissions in Europe.

**report nécessaire**

Désigne l'état d'obligation d'un [équipement-type](#) par rapport à un [intervalle infra-journalier](#) si son fonctionnement ne peut avoir lieu dans cet intervalle et l'oblige donc à être reporté ultérieurement.

**repère de l'équipement à considérer**

Désigne un instant dans une journée à partir duquel un équipement donné peut être déclenché. Cet instant tient compte des appareils qui lui sont éventuellement liés.

**RJT**

Résumé du Jour-Type.

**RMSE**

Root Mean Square Error.

**RT**

Règlementation Thermique.

**RTE**

Réseau de Transport d'Électricité.

**règle de cumul (ou d'additivité)**

Les règles de cumul (ou d'additivité) désignent les hypothèses permettant, à partir de profils individuels de disponibilité des membres d'un ménage, d'aboutir au [profil agrégé de disponibilité](#) de celui-ci.

**réaffectation**

La réaffectation d'un bâtiment correspond à un changement de sa valorisation (utilisation). Par exemple, des locaux industriels ou tertiaires peuvent être réaffectés en logements.

**régime établi**

Qualifie la situation stable d'un système ou d'un processus sur une période de temps donnée par opposition au régime transitoire.

**réglementation thermique**

Il s'agit d'un arrêté, paru au journal officiel de la République et mis à jour périodiquement, qui est relatif aux caractéristiques thermiques et aux exigences de performance énergétique des bâtiments nouveaux et des parties nouvelles de bâtiments.

**rénovation**

Processus d'amélioration d'une habitation ou d'éléments le constituant (systèmes énergétiques notamment) suite à la réalisation de travaux.

**résumé d'intervalle**

Le résumé d'intervalle comprend les instants de déclenchement et les [signatures d'équipements](#) des appareils qui fonctionnent au sein de cet intervalle. Le résumé d'intervalle est ordonné de manière chronologique.

**résumé du jour-type**

Tableau de six éléments comprenant le niveau de disponibilité caractéristique de chaque [intervalle infra-journalier](#).

**Réveil**

[Intervalle infra-journalier](#) délimité par les [instants caractéristiques](#)  $t_{start}$  et  $t_1$ .

**saisonnalité**

Il s'agit de l'influence que peuvent avoir les saisons vis-à-vis d'un phénomène, d'un processus ou d'une variable. Un exemple évident est la saisonnalité des besoins de chauffage.

**saturation**

Pour un équipement auto-lié, la saturation désigne le nombre maximum d'utilisations de cet équipement au sein d'un [intervalle infra-journalier](#) spécifié. Cette limite est fixée par le modélisateur.

**saut de classe**

En ce qui concerne la [performance thermique d'enveloppe](#) d'une habitation, un saut de classe désigne le passage d'une classe de performance initiale à une classe de performance meilleure.

**scénario principal (ou scénario principal *stricto sensu*)**

Le scénario principal désigne le cas de référence que nous avons établi pour la simulation de notre modèle de reconstitution de courbe de charge.

**scénario principal assimilé**

Le scénario principal assimilé est un scénario qui ne se différencie du [scénario principal \*stricto sensu\*](#) que par la valeur de la probabilité de déclenchement du sèche-linge. Dans ce scénario, elle ne vaut pas 1.

**scénario principal au sens large**

Désigne le regroupement du [scénario principal \*stricto sensu\*](#) et du [scénario principal assimilé](#). Il est noté scénario principal [ASL](#).

**scénario Time Of Use**

Instants de déclenchement – mise en marche – des équipements-types d’un n-uplet pour une journée donnée.

***SDTW***

Segmented Dynamic Time Warping.

**segment**

Un segment du parc résidentiel désigne une fraction de celui-ci c’est-à-dire un échantillon de ménages, de bâtiments ou, dans le cas de notre modélisation de n-uplets, représentatif ou non de l’ensemble du parc considéré.

**seuil d’absence**

Dans le cadre de la modélisation du [ménage-type](#), le seuil d’absence est la probabilité qu’un membre d’un [ménage-type](#) éligible aux absences, quitte son logement durant une certaine durée lors d’un jour donné.

**signature d’équipement**

La signature d’équipement est un élément numérique permettant de reconnaître l’appareil utilisé et la durée d’usage correspondante lors de son utilisation.

**SMA**

Système Multi-Agents.

**SO**

SOMmeil.

**Soir**

[Intervalle infra-journalier](#) délimité par les [instants caractéristiques](#)  $t_2$  et  $t_{End}$ .

**solde fatal en énergie**

Il s’agit de la quantité d’énergie occasionnée par l’utilisation des équipements non explicités dans la modélisation du [parc habité](#).

**solde migratoire**

« Le solde migratoire est la différence entre le nombre de personnes qui sont entrées sur le territoire et le nombre de personnes qui en sont sorties au cours d’une année donnée (ce concept est indépendant de la nationalité) » [\[INS12a\]](#).

**solde naturel**

« Le solde naturel est la différence entre le nombre de naissances et le nombre de décès enregistrés au cours d’une période. Ce solde peut aussi bien être positif (hausse de l’effectif de population) que négatif (réduction de la population dans son effectif) » [\[INS12b\]](#).

**sous-équipement**

Le sous-équipement correspond à la diminution du multi-équipement, autrement dit à celle du [taux d’équipement](#) des ménages équipés.

**statut d’activité**

Le statut d’activité d’un membre d’un foyer (ou du ménage considéré de manière agrégée) est une variable intermédiaire de l’activité principale de cette personne et de son emploi du temps quotidien (en particulier celui des jours ouvrés). Dans le cadre de la modélisation du [ménage-type](#), cette variable présente trois modalités **Actif**, **ActifDomicile** et **Inactif**.

**STEP**

Station de Transfert d’Énergie par Pompage.

**suréquipement**

Le suréquipement correspond à l’augmentation du multi-équipement, autrement dit à celle du [taux d’équipement](#) des ménages équipés.

**synchrone**

Somme de courbes de charge unitaires.

**TAAR**

Tirage Aléatoire Avec Remise.

**tableau de compatibilité équipement/intervalles**

Un tableau de compatibilité équipement/intervalles renferme la compatibilité d'un équipement – autrement dit l'opportunité d'utiliser cet appareil – vis-à-vis des différents intervalles infra-journaliers.

**TAD**

Travail À Domicile.

**TASR**

Tirage Aléatoire Sans Remise.

**taux d'équipement**

Pourcentage de la population totale équipée d'un bien de consommation donné.

**taux de pénétration**

Synonyme de [taux d'équipement](#).

**taux de sondage**

Le taux de sondage d'une population (ou d'une strate de population) est égale à la fraction à sonder de cette population (respectivement de cette strate de population) afin d'obtenir une estimation satisfaisante, c'est-à-dire au sein d'un intervalle de confiance fixé au préalable, de la grandeur d'intérêt pour l'ensemble de la population (de la strate de population).

**techno-explicite**

Se dit d'une modélisation qui cherche à prendre en considération les différences qu'il existe en matière technique entre les objets étudiés.

**temps caractéristique**

Ordre de grandeur de l'échelle temporelle à laquelle fonctionne un équipement domestique.

**température d'eau froide**

Température à laquelle l'eau froide est prélevée en entrée d'un système de production d'ECS.

**température de puisage**

Température de référence de l'eau chaude sanitaire produite par un équipement thermique (chaudière, ballon à accumulation...) et mesurée en sortie de cet appareil.

**tendance d'évolution**

Il s'agit de la manière avec laquelle sont modifiés qualitativement et/ou quantitativement les objets ([éléments-types](#) notamment) et les paramètres intervenant dans la modélisation de la [courbe de charge](#) au cours de la durée totale de la simulation.

**territoire**

Dans le modèle de reconstitution de [courbe de charge](#), un territoire désigne la maille géographique de la modélisation. Il peut soit s'agir d'une seule [zone géographique](#), d'un regroupement de [zones géographiques](#) ou bien encore du territoire national complet.

**THT**

Très Haute Tension.

**triplet**

Un triplet est un [couple structurant](#) doté d'un jeu d'[équipements-types fixes](#).

**TS**

Tirage Séquentiel.

**type de fonctionnement**

Pour un équipement domestique, le type de fonctionnement est une caractéristique immuable de l'appareil – à ne pas confondre avec la notion de [mode de fonctionnement](#) – qui définit son fonctionnement intrinsèque.

**UFE**

Union Française de l'Électricité.

**valeur efficace**

La valeur efficace d'une sinusoïde est obtenue en calculant le rapport entre la valeur maximale atteinte par la courbe sur une période (ou valeur « crête ») et  $\sqrt{2}$ .

**VE**

Véhicule Électrique.

**VHR**

Véhicule Hybride Rechargeable.

**VMC**

Ventilation Mécanique Contrôlée.

**zone du bâtiment**

Un [bâtiment-type](#) est constitué d'un ensemble de pièces rassemblées en plusieurs zones, a minima deux (bâtiment bi-zone). Celles-ci autorisent un réglage différencié des appareils de chauffage/climatisation entre une zone de pièces à vivre et une zone destinée au couchage par exemple.

**zone géographique**

Ensemble contigu de départements qui présentent des similitudes en matière de [chroniques météorologiques](#).

# Index

- A**  
absence,  
    seuil d' ..... 116  
ACP ..... 197  
activité,  
    statut d' ..... 64, 115  
agrégation ..... 6  
agrégation,  
    niveau d' ..... 13  
année de calage ..... 75  
apports internes ..... 95  
assemblage ..... 235
- B**  
base,  
    double ..... 67  
    unique ..... 66  
bâtiment,  
    performance énergétique d'un ..... 63  
    performance thermique d'enveloppe d'un .. 62  
    zone du ..... 95  
bâtiment-type ..... 60  
bibliothèque ..... 60  
bottom-up ..... 9, 13
- C**  
CAH ..... 198  
cas favorable ..... 141  
charge,  
    gestion de la ..... 9  
    profil de ..... 97  
chauffage d'appoint ..... 67  
chemin de déformation ..... 162  
chronogramme probabilisé d'usage ..... 64  
 $comp_t$  ..... 233  
 $comp_{Usage, i}$  ..... 233  
consigne spatio-temporelle ..... 95  
couple structurant ..... 60  
couple structurant,  
    éligibilité d'un ..... 78  
    poids de représentativité ..... 60  
courbe de charge ..... 1  
courbe de charge,  
    moyenne foisonnée ..... 96, 170, 175  
    normée ..... 103  
    unitaire ..... 74, 175  
courbe de production ..... 1
- D**  
 $DDTW$  ..... 162  
 $\Delta t_{Abs. Jour.}$  ..... 112  
 $\Delta t_{Départ}$  ..... 112  
 $\Delta t_{Prés. MÉR.}$  ..... 112  
diagonale NOSE ..... 181  
disponibilité,  
    niveau de ..... 93, 130  
    profil agrégé de ..... 133  
    requis ..... 117  
 $DTW$  ..... 162  
durée à considérer ..... 120, 138  
 $DureeAConsiderer$  ..... 120
- E**  
EDF ..... 6  
effacement ..... 8, 9, 75  
effet,  
    boomerang ..... 8  
    rebond ..... 9  
efficacité énergétique ..... 6  
élément-type ..... 60  
énergie,  
    maîtrise de la demande en ..... 8  
ENR ..... 4  
équipement,  
    à considérer ..... 120, 138  
    auto-lié ..... 118  
    autorisation de dépassement de l' ..... 118, 139  
    compatibilité de l' ..... 117  
    durée d'usage d'un ..... 117  
    forme condensée ..... 119  
    forme développée ..... 119  
    lié ..... 117  
    mode de fonctionnement d'un ..... 74  
    partie non déplaçable de l'usage d'un ..... 139  
    pointe concomitante journalière d'un ..... 175  
    report nécessaire ..... 117  
    saturation d'un ..... 139  
    signature de l' ..... 114, 117

- taux d' ..... 3
- équipement-type ..... 60
- équipement-type,
  - clé de répartition ..... 76
  - fixe ..... 65
  - nomade ..... 65
  - nomenclature ..... 74
  - paramètres de gestion ..... 74
  - paramètres fonctionnels ..... 74
  - vie d'un ..... 83
- ErDF ..... 6
- événement ..... 114
- événement-clé ..... 60
- F**
- filtre ..... 116
- flux net ..... 80
- foisonnement ..... 8
- foisonnement,
  - coefficient de ..... 102
- fonctionnement,
  - mode de ..... 74
  - type de ..... 116
- G**
- génération ..... 80
- gestion,
  - mode de ..... 90
- gradient,
  - de consommation énergétique ..... 106
- I**
- IA* ..... 158
- indicateur,
  - binaire ..... 155
  - unaire ..... 154
- IndiceIntervalleAConsiderer ..... 120
- instant caractéristique ..... 112
- instant éligible ..... 146
- intervalle à considérer ..... 120, 138
- intervalle,
  - forme développée d' ..... 114
  - résumé d' ..... 114
- IntervalleLimite ..... 120
- J**
- jour-type ..... 89
- jour-type,
  - résumé du ..... 131
- L**
- LBC ..... 6
- $l_{Intervalle}$  ..... 114
- M**
- macro-tâches ..... 128
- MAPE* ..... 102
- matrice,
  - de macro-compatibilité ..... 136
  - de micro-compatibilité ..... 137
- ménage,
  - agrégé ..... 110
- ménage-type ..... 60
- micro-tâches ..... 137
- modélisation générique ..... 117
- N**
- n-uplet ..... 61
- n-uplet,
  - pois dans l'effectif ..... 61
- nébulosité ..... 91
- niveau de responsabilité ..... 115
- NVF* ..... 32, 102
- P**
- parc de référence ..... 80
- parc habité ..... 60, 61
- parc résidentiel,
  - segment du ..... 167
- performance énergétique,
  - saut de classe ..... 81
- période ..... 60
- perturbation ..... 123
- pilotage ..... 75
- PNDU ..... 139
- pointe concomitante ..... 159
- poste de consommation ..... 101
- profil,
  - d'occupation ..... 130
  - de consommation ..... 10
  - de disponibilité ..... 114
  - de visionnement ..... 215
  - exogène ..... 120
- profilage ..... 10
- puisage ..... 68
- puisage,
  - chronique de ..... 90
  - profil de ..... 96
- puissance ..... 226
- puissance,
  - active ..... 2
  - apparente ..... 226
  - cycle unitaire de ..... 74
  - facteur de ..... 229
  - instantanée ..... 226
  - monotone de ..... 5

- pointe de ..... 4  
 réactive ..... 227  
 réduite ..... 2
- R**
- rayonnement solaire ..... 91  
 réaffectation ..... 80  
 règle,  
   de cumul ..... 133  
 réglementation thermique ..... 56  
 rénovation ..... 57  
 RepereEquipementAConsiderer ..... 120  
 report nécessaire ..... 144  
 résumé d'intervalle,  
   état initial d'un ..... 139  
*RMSE* ..... 157  
 RTE ..... 3
- S**
- saisonnalité ..... 3  
 scénario,  
   principal ..... 167  
   principal *stricto sensu* ..... 173  
   principal assimilé ..... 173  
   principal au sens large ..... 173  
   Time Of Use ..... 64, 92  
*SDTW* ..... 165  
 simultanément,  
   degré de ..... 177  
 SMA ..... 221  
 solde fatal ..... 66, 73  
 synchrone ..... 6, 97
- T**
- $t_1$  ..... 112  
 $t_2$  ..... 112  
 $t_3$  ..... 112  
 $t_4$  ..... 112  
 tableau de compatibilité équipement/intervalles 120  
 tâche domestique,  
   macro-compatibilité d'une ..... 136  
   micro-compatibilité d'une ..... 137  
 taux de pénétration ..... 65  
 taux de sondage ..... 195  
 température extérieure ..... 91  
 temps caractéristique ..... 114  
 $t_{End}$  ..... 112  
 tendance d'évolution ..... 87  
 territoire ..... 61  
 $t_{Final}$  ..... 114  
 $t_{Initial}$  ..... 114  
 top-down ..... 13
- triplet ..... 61  
 $t_{Start}$  ..... 112  
 $t_{StartBis}$  ..... 112
- U**
- UFE ..... 7  
 usage,  
   chronogramme probabilisé d' ..... 92  
   cuisson ..... 3  
   d'électricité spécifique ..... 3  
   thermique ..... 3
- V**
- valeur efficace ..... 226
- Z**
- zone géographique ..... 56, 62

## Introduction de non linéarités et non stationnarités dans les modèles de représentation de la demande électrique résidentielle

**Résumé :** La problématique développée dans la thèse est d'estimer, dans une démarche prospective et dans un but d'anticipation, les impacts en puissance induits par les ruptures technologiques et comportementales qui ne font pas aujourd'hui l'objet de mesures dans les panels. Pour évaluer les modifications sur les appels de puissance du parc résidentiel engendrées par ces profondes transformations, un modèle paramétrique, bottom-up, techno-explicite et agrégatif est donc nécessaire. Celui-ci serait donc destiné à la reconstitution, de manière non tendancielle, de la courbe de charge électrique résidentielle. Il permettrait ainsi de conduire la simulation de différents scénarios d'évolution contrastés. L'élaboration d'un tel modèle constitue le sujet de ce doctorat.

Pour répondre à cette problématique, nous proposons une méthode conceptuelle originale de reconstitution de courbe de charge. Sa mise en application centrée sur la génération de foisonnement d'origine comportementale a conduit à la mise en œuvre d'un certain nombre de concepts. Ce travail a abouti à l'élaboration d'un algorithme stochastique destiné à représenter le déclenchement réaliste des appareils domestiques. Différents cas d'application ont pu être testés et les résultats en puissance ont été étudiés. Plus particulièrement pour analyser le foisonnement visible à un niveau agrégé, nous avons mis en place une méthodologie nouvelle basée sur une distance adaptée aux courbes de charge. Finalement, nous avons cherché à identifier des comportements réels d'usage des appareils. Pour cela, nous avons conduit différents travaux de classification de courbes de charge.

**Mots clés :** modélisation bottom-up, courbe de charge électrique, foisonnement, équipements domestiques, algorithme stochastique

## Introducing non stationarities and non linearities in the residential load curve reconstitution models

**Abstract:** In this dissertation, we focus on the estimation of the impacts in terms of power demand caused by the technological and behavioural breaks that will affect the domestic sector in the future. These deep changes are not measured in the existing panels and the estimation is required for prospective (long-term) studies. To evaluate the very likely modifications of the domestic power demand that will follow previous influences, a bottom-up, technically-explicit and aggregative model is needed. This one aims at reconstituting the electric residential load curve according to a non-trending manner. Thanks to it, various evolution scenarios can be simulated. The purpose of this PhD is the elaboration of such a model.

A functional analysis was carried out to build up a new method to reconstitute the domestic electric load curve. Since the clarifying of the diversity represents one of the key points of our research, we decided to begin the modelling task with focus on it. More precisely, we elaborated a stochastic algorithm whose purpose is the realistic starting of domestic appliances. Some application cases have been tested. We studied the diversity affecting aggregated power demand and we propose a new methodology able to visualise and to analyse it. This method is based on a distance adapted to the load curve. Finally we tried to identify human behaviour concerning the use of appliances thanks to load curve classifications.

**Keywords:** bottom-up modelling, load curve, diversity, domestic appliances, stochastic algorithm