

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université 20 Août 1955 - Skikda



D012117004D

Faculté de Technologie
Département de Génie Electrique
Laboratoire d'Electrotechnique de Skikda

THÈSE
EN VUE DE L'OBTENTION DU DIPLOME DE
DOCTORAT

Domaine : Science et Technologie Filière : Génie Electrique
Spécialité : Electrotechnique Appliquée

Présentée par

LAOUAFI Abderrezak

Thème

**Contribution à la modélisation de la courbe de charge
électrique par des techniques intelligentes**

Soutenue publiquement le : 03/05/2017

Devant le Jury composé de :

M. BOUKADOUM Ahcene	Professeur	Université de Skikda	Président
M. MORDJAoui Mourad	MCA	Université de Skikda	Rapporteur
M. LEBAROUD Abdesselam	Professeur	ENP Constantine 3	Examineur
M. MEHASNI Rabia	Professeur	Université Constantine 1	Examineur
M. MEDOUED Ammar	MCA	Université de Skikda	Examineur
M. BOUDEBBOUZ Omar	MCA	Université de Skikda	Examineur

Année Universitaire 2016-2017

Remerciements

Ce travail a été effectué au laboratoire d'électrotechnique de la Faculté de technologie-Université de Skikda sous la direction de monsieur Mourad MORDJAOUI, Maître de Conférences au département de génie électrique de l'université de Skikda.

Je remercie tout d'abord Dieu tout puissant de m'avoir guidé durant toutes mes années d'études et qui m'a aidé à la réalisation de ce travail, en me donnant la force, la patience et la volonté.

Ensuite, je tiens à exprimer ma sincère gratitude et ma profonde reconnaissance à Monsieur Mourad MORDJAOUI, mon directeur de thèse pour m'avoir encadré, soutenu et m'orienté dans la préparation de ce travail. Sa disponibilité et son engagement scientifique m'ont grandement aidé tout au long ces années.

Je remercie également Monsieur Ahcene BOUKADOUM, Professeur à l'université de Skikda et directeur de laboratoire d'électrotechnique de la Faculté de technologie- Université de Skikda, qui m'a fait l'honneur de s'intéresser à ce travail proposé, et de présider le jury.

Je voudrais également adresser mes remerciements les plus sincères à monsieur Abdesselam LEBAROUD, Professeur à l'école nationale polytechnique de Constantine ; et à monsieur Rabia MEHASNI, Professeur à l'université de Constantine 1 qui ont accepté d'évaluer mon travail et participer au jury de thèse comme examinateurs.

Mes remerciements vont également à monsieur Ammar MEDOUED et monsieur Omar BOUDEBBOUZ, Maîtres de Conférences au département de génie électrique de l'université de Skikda, pour l'honneur qu'ils m'ont fait en acceptant d'examiner mon travail et pour les compétences et les qualités humaines qui m'ont été d'un soutien inestimable au cours de ces années.

Je tiens également à remercier Mme Farida LAOUAFI, M. Sahim KISSAR, M. Mohamed LASHAB, M. Bouzid BOUDJEMA et tous ceux qui m'ont aidé à établir cette thèse.

Dédicaces

Je dédie ce modeste effort, à mes très chers parents que Dieu les garde.
À mes sœurs, mon frère, ma nièce Kaoutar et à tous les membres de ma famille.
À mes collègues de l'Université de Skikda.
À tous mes amis.

ملخص:

يرتبط هذا العمل بمجال التنبؤ بالطاقة التي يطلبها المستهلكون في الشبكات الكهربائية. برغم العديد من الدراسات التي نشرت منذ سنوات الستينيات، لا يزال هذا المجال يعتبر كإشكاليه صعبة الحل، وهذا بسبب عدة اعتبارات وارتباطات معقدة. وبالتالي، فإن التوقعات الدقيقة للحمولة الكهربائية تمثل تحديا حقيقيا للباحثين، حيث يحاول كل باحث تحسين دقة التوقعات ويقارن أداء نموذج الخاص مع أداء نماذج أخرى سابقة. المساهمة الأساسية لهذا البحث تندرج إذا في هذا السياق من أجل تحسين أداء التنبؤ. في الحقيقة يركز عملنا أساسا على تطوير أساليب قادرة على ضمان التنبؤات المرضية على المديين القصير و القصير جدا. على كل حال، لقد أنجزنا دراسة أولية في ما يخص مديي التنبؤ المتوسط و الطويل، حيث كان الهدف من وراء هذه الدراسة هو تقييم مجموعة من النماذج الأساسية المستعملة بكثرة في مجال بحثنا. النتائج المستخلصة من هذه الدراسة أظهرت أنه من المستحسن استخدام أكثر من نموذج من أجل تحسين جودة التنبؤ. هذه الفكرة كانت عنصرا أساسيا في دراسة ثانية مرتبطة بالتوقع على المدى القصير لاستهلاك الكهرباء في الجزائر.

يرتكز النموذج المقدم من أجل التنبؤ على المدى القصير على استخدام منهجية هجينة توافمية ذات مرحلتين، حيث تتمحور الفكرة الرئيسية على توفير توقع أولي ومن ثم محاولة تحسينه من خلال نظام توقعي ثانوي. وأظهرت نتائج هذه الدراسة أداء أفضل مقارنة مع طريقة تنبؤ مرجعية، معروفة حاليا تحت مسمى طريقة Holt-Winters-Taylor .

النهج المتبع في التنبؤ على المدى القصير جدا، يركز بدوره، على فكرة استخدام العديد من النماذج التنبؤية. على كل حال، من أجل توقع الحملولة الكهربائية لبضع دقائق قادمة، لن يكون هناك وقت لمحاولة تحسين دقة التنبؤ. من الضروري إذا البحث عن أفضل أداء من البداية. لذلك، الفكرة الأساسية في هذه الحالة هي استخدام مزيج من التوقعات والبحث عن أفضل أداء انطلاقا من مجموعة من الأساليب الرئيسية. طبقنا هذه المنهجية على بيانات الحملولة في الوقت الحقيقي الخاصة بنظام الطاقة الفرنسي، وكذلك على البيانات الخاصة بولاية نيو ساوث ويلز بأستراليا. وقد أظهر هذا النظام نتائج أفضل بالمقارنة مع خوارزمية مرجعية، اسمها AWNN ، وطورت في عام ٢٠١٦ من قبل M. Rana .

الكلمات المفتاحية : حمولة كهربائية، تنبؤ، توازن عرض - طلب، مزج التوقعات، تقنيات توافمية

Résumé : Le présent travail s'inscrit dans la thématique de la prévision de la puissance appelée dans les réseaux électriques. Cette thématique a fait l'objet de recherches pendant plus de cinq décennies. Cependant, elle demeure une problématique assez difficile à cause du comportement non linéaire de la charge électrique et sa relation complexe avec plusieurs facteurs. Les prévisions satisfaisantes de charge représentent donc un défi important pour les chercheurs, où chaque auteur essaye d'améliorer la précision de prévision et compare la performance obtenue par son propre modèle avec les performances d'autres modèles issus de la littérature. Notre contribution s'inscrit donc dans ce cadre pour l'amélioration des performances de prévision. On s'intéresse principalement au développement des méthodes assurant les prévisions satisfaisantes pour les deux horizons de prévision à très court-terme et court-terme de la charge électrique.

Néanmoins, nous avons effectués une première étude pour les deux horizons temporels de prévision à moyen et à long-terme, dont le but est d'évaluer un ensemble de méthodes existantes. Les résultats de cette étude ont montré qu'il est définitivement conseillé d'utiliser à la fois plusieurs méthodes afin d'améliorer les performances. Cette idée été l'élément clé dans une deuxième étude consacrée à la prévision à court-terme de la consommation d'électricité en Algérie.

L'approche pour la prévision à court-terme consiste à utiliser une méthodologie hybride adaptative à deux étapes, dont l'idée clé est d'assurer une première prévision et d'essayer ensuite de l'améliorer par le biais d'un système de prévision secondaire. Les résultats de cette étude ont montré de meilleure performance par rapport à une méthode de prévision issue de la littérature, et connue actuellement sous le nom de la méthode de Holt-Winters-Taylor.

L'approche pour la prévision à très court-terme, a sans tour, consiste à exploiter cette idée d'utiliser plusieurs méthodes. Cependant, pour de tel cas de prévision pour quelques minutes à l'avance il n'y aura pas du temps pour tenter une amélioration de prévision. Il fallait donc chercher la meilleure performance dès un premier temps. Par conséquent, notre idée pour de tel cas est d'utiliser une combinaison de prévisions et de tirer la meilleure performance à partir d'un ensemble de méthodes primaires. Nous avons appliqués cette méthodologie combinée sur les données de charge en temps-réel du système électrique français, et l'ensemble de données australiennes pour l'Etat de la Nouvelle-Galles du Sud. Les résultats de ce système ont montré de meilleures performances en comparaison avec un algorithme de référence, nommé AWNN, et développé en 2016 par M. Rana et al.

Mots clés : Charge électrique, prévision, équilibre offre/demande, combinaison de prévision, techniques adaptatives

Abstract : This work falls under the topic of electric load forecasting. Despite numerous studies in the literature published since the 1960s, this problem is so far considered as difficult because of the nonlinear nature of the electric load time series and the complex relationship with several factors. Thus, satisfactory load forecasts represent an important challenge for researchers, where each author tries to improve the forecast accuracy and compares the performance obtained by its own model with the performance of other models from the literature. Our contribution is therefore in this context for the improvement of forecasting performance. It focuses mainly on the development of methods that can ensure the satisfactory predictions for the two horizons of very short and short-term forecasting.

Nevertheless, we have made a first study for the two time horizons of medium- and long-term forecasting, where the aim is to evaluate a set of existing methods. The results of this study showed that it is definitely advisable to use a set of methods for improving the prediction accuracy. This idea was the key element in a second study devoted to the short-term forecast of electricity consumption in Algeria.

The approach for short-term forecasting consists of using an adaptive hybrid two stage methodology, where the key idea in this instance is to provide a preliminary forecast and to try then to improve its accuracy through the use of secondary forecasting system. The results of this study showed better performance compared to a prediction method derived from the literature, which is currently known under the name of the Holt-Winters-Taylor method.

The approach of very short-term forecasting, in their turn, consists of exploiting this idea of using several methods. However, for forecasts as short as few minutes in advance, there will be no time to attempt a forecast improvement. As a result, it is necessary to look for the best performance from the first attempt. Therefore, our idea for this case is to use a forecast combination and draw the best performance from a set of primary methods. We applied this combined methodology on the real-time load data from the French power system, and on the Australian data for the state of the New South Wales. The results of this system have shown better results in comparison with a reference algorithm from literature, named AWNN and developed in 2016 by M. Rana et al.

Keywords : Electric load, forecast, supply/demand balance, forecast combination, adaptive techniques

Liste des Abréviations

AEMO	Opérateur du marché de l'énergie en Australie
AHES	Lissage exponentiel adaptatif de Holt
ANFIS	Système d'inférence neuro-flou adaptatif
APE	Erreur absolue en pourcentage
AR	Modèle autoregressif
ARIMA	Modèle autorégressif à moyenne mobile intégré
ARMA	Modèle autorégressif à moyenne mobile
AWNN	Modèle de prévision basé sur la transformée en ondelettes et les réseaux de neurones
BPNN	Perceptron multicouche à rétro-propagation de l'erreur
CWT	Transformée en ondelettes continue
DWT	Transformée en ondelettes discrète
EEMD	Décomposition en mode empirique d'ensemble
FCM	Algorithme des c-moyennes floues
FIS	Système d'inférence floue
FKW-ANFIS	Modèle de prévision secondaire composé de WPD, FCM, k-NN, WPD et ANFIS
FKW-BPNN	Modèle de prévision secondaire composé de WPD, FCM, k-NN, WPD et BPNN
FKW-NARX	Modèle de prévision secondaire composé de WPD, FCM, k-NN, WPD et NARX
FT	Transformée de Fourier
HES	Lissage exponentiel de Holt
HFCM-TM	Méthode de combinaison de prévision basée sur la moyenne tronquée et le filtre Hampel
HFCM-WAM	Méthode de combinaison de prévision basée sur la moyenne pondérée et le filtre Hampel
HW	Lissage exponentiel de Holt-Winters
HWT	Lissage exponentiel à double saisonnalité de Holt-Winters-Taylor
IMF	Composants de fonction du mode intrinsèque

k-NN	Algorithme des k-plus proches voisins
LTLF	Prévision à long-terme de la charge électrique
MA	Modèle à moyenne mobile
MAD	Écart absolu moyen
MAPE	Erreur absolue moyenne en pourcentage
MSE	Erreur quadratique moyenne
MTLF	Prévision à moyen-terme de la charge électrique
NARX	Réseau de neurones non linéaire autorégressif avec entrée exogène
OSE	Opérateur du système électrique algérien
P-ANFIS	Modèle ANFIS pour assurer une prévision primaire
P-BPNN	Modèle BPNN pour assurer une prévision primaire
P-NARX	Modèle NARX pour assurer une prévision primaire
PSO	Optimisation par essaim de particules
RMSE	Racine carrée de l'erreur quadratique moyenne
RTE	Gestionnaire du réseau de transport d'électricité de France
S-ANFIS	Modèle ANFIS pour assurer une prévision secondaire
S-BPNN	Modèle BPNN pour assurer une prévision secondaire
S-NARX	Modèle NARX pour assurer une prévision secondaire
SARIMA	Modèle saisonnier autorégressif à moyenne mobile intégré
SFCM	Méthode de combinaison de prévision basée sur la moyenne arithmétique
STFT	Transformée de Fourier à court-terme
STLF	Prévision à court-terme de la charge électrique
VSTLF	Prévision à très court-terme de la charge électrique
WPD	Décomposition en paquets d'ondelettes
WT	Transformée en ondelettes

Table des figures

1.1	Constitution d'un réseau électrique	2
1.2	Equilibre entre la production et la consommation d'énergie électrique	3
1.3	Placement des centrales de production sur la courbe de charge française : 01/06/2016 [1]	4
1.4	Courbe de charge française : 01 Juin 2016	5
1.5	Les différents horizons de temps pour la prévision de la charge électrique	10
1.6	Facteurs d'influence possibles pour la prévision à très court et court-terme de la charge	11
2.1	Notion de la courbe de charge électrique	17
2.2	Cycle journalier la consommation française d'électricité [45]	18
2.3	Courbe de charge hebdomadaire en France : du 08 au 14 Janvier 2011 [45]	19
2.4	Courbe de charge hebdomadaire en France : du 04 au 10 Juin 2011 [45]	20
2.5	Comparaison entre la consommation hebdomadaire des quatre saisons [45]	21
2.6	Consommation d'électricité en France pendant l'année 2011 [45]	21
2.7	Quelques exemples de séries chronologiques [47]	23
2.8	Avant (en haut) et après (en bas) le filtrage de valeurs aberrantes basée sur deux jours de données de charge britannique avec une résolution de 4 secondes [54]	28
2.9	Classification des données	29
2.10	Classement des données	29
2.11	Une approche parallèle pour classer les données de la charge selon l'heure de la journée [59]	30
2.12	Transformée de Fourier d'un signal de série temporelle	31
2.13	Transformée de Fourier à court-terme d'un signal de série temporelle	32
2.14	Différentes méthodes pour assurer les prévisions de la charge électrique	34
2.15	Prévision et intervalle de prévision des taux de défaillance des entreprises [63]	37
3.1	Données mensuelles de la charge de pointe en Algérie : de janvier 2010 à décembre 2014	40
3.2	Evolution de la charge de pointe annuelle en Algérie : 2001-2014	41
3.3	Modèle d'un neurone artificiel	44
3.4	Fonctions d'activation couramment utilisées [71]	45
3.5	Exemple d'un réseau de type perceptron multicouche	46
3.6	Représentation de « Grands » dans le cas booléen et flou	48
3.7	Processus d'un système flou	49
3.8	Architecture d'ANFIS pour deux règles	52

3.9	Fonction d'appartenance couramment utilisées : (a)- sigmoïde, (b)- triangulaire, (c)- forme de cloche	53
3.10	Fonction d'appartenance couramment utilisées : (a)- gaussienne, (b)- trapézoïdale	54
3.11	Algorithme d'apprentissage d'ANFIS pour l'ajustement de la production des paramètres des règles	56
3.12	Résultats des méthodes de prévision à moyen-terme : 2014	57
3.13	Résultats des méthodes de prévision à moyen-terme : 2015	57
3.14	Résultats des méthodes de prévision à long-terme : 2010-2014	58
3.15	Résultats des méthodes de prévision à long-terme : 2015-2020	59
3.16	Les prévisions combinées de SARIMA-ANFIS tracées conjointement avec l'intervalle prévu pour l'année 2015	60
3.17	Les prévisions combinées de ARIMA-ANFIS tracées conjointement avec l'intervalle prévu pour la période entre 2015 et 2020	60
3.18	SARIMA-ANFIS : Erreur absolue en pourcentage pour la période Jan-Mai 2015	62
3.19	SARIMA-ANFIS : Prévision et intervalle de prévision pour la période Jan-Mai 2015	62
4.1	La demande de pointe d'électricité du 01 Janvier 2012 au 30 Novembre 2014	65
4.2	Démarche suivie pour optimiser les paramètres de la méthode de HWT	67
4.3	Regroupement c-moyennes floues : résultats pour le 30 Novembre 2014	68
4.4	Algorithme de classification par c-moyennes floues	69
4.5	Algorithme des k-plus proches voisins	71
4.6	Algorithme des k-plus proches voisins : résultats pour le 30 Novembre 2014	71
4.7	Dilatation et translation d'une ondelette mère	72
4.8	Ondelettes mères couramment utilisées	73
4.9	Arbre en paquets d'ondelettes de trois niveaux	74
4.10	Résultats de WPD pour la prévision de la charge de pointe du 30 Novembre 2014	75
4.11	Architecture du réseau neuronal non linéaire autorégressif avec entrée exogène	78
4.12	Structure générale des méthodes proposées	79
4.13	Résultats de FKW-ANFIS pour quatre mois différents : Janvier, Avril, Juillet et Novembre 2014	79
4.14	Résultats en MAPE de HWT et FKW-ANFIS pour chaque jour de la semaine	81
5.1	Les données de la demande d'électricité du système électrique français : 01 Jan-31 Dec 2014	85
5.2	La pointe journalière d'électricité en France durant quatre semaines différentes	85

5.3	Structure générale de la méthodologie proposée pour la prévision en ligne de charge	89
5.4	Résultats de HFCM-TM pour les données de charge français : du 09 Août au 12 Août 2014 2014	93
5.5	Résultats de HFCM-TM pour les données de charge français : du 13 Août 2014-15 Août 2014	93
5.6	Résultats de HFCM-TM pour les données de charge français : du 10 Janvier 2015 au 13 Janvier 2015	94
5.7	Résultats de HFCM-TM pour les données de charge français : du 14 Janvier 2015 au 16 Janvier 2015	94
5.8	Les paramètres de lissage exponentiel de AHES le long d'une semaine en été et une semaine en hiver	96
5.9	Résultats de HFCM-TM et HFCM-WAM pour le processus de prévision du 26 Décembre 2014 à 00 :00	98
5.10	Répartition de APE pour AHES, SFCM, HFCM-WAM, et HFCM-TM pour les charges normales	99
5.11	Répartition de APE pour AHES, SFCM, HFCM-WAM, et HFCM-TM pour les charges anormales	99
5.12	Charge prédite par HFCM-TM tracée en fonction de la charge réelle de la base de données australienne	100
5.13	Outil développé pour la prédiction en ligne de la demande d'électricité	102

Liste des tableaux

2.1	Valeurs manquantes dans une base de données utilisée pour prévoir le pic journalier de la charge électrique en Algérie	29
3.1	Avantages et inconvénients de la logique floue et les réseaux de neurones	50
3.2	Performances des modèles proposés pour MTLF	59
3.3	Performances des modèles proposés pour LTLF	60
3.4	SARIMA-ANFIS : Résultats pour la prévision des pics mensuels pour l'année 2015	61
3.5	ARIMA-ANFIS : Résultats pour la prévision des pics annuels pour la période entre 2015-2020	61
3.6	SARIMA-ANFIS : Prévision et intervalle de prévision pour la période Jan-Mai 2015	63
4.1	Les principaux éléments, paramètres, et entrées pour les différents modèles	76
4.2	L'erreur relative (MAPE) de toutes les méthodes proposées	77
4.3	Résultats obtenus lorsque HWT et FKW-ANFIS sont utilisés pour prévoir la charge de pointe des jours fériés en 2014	81
5.1	Les principaux éléments et paramètres des différentes méthodes considérées	87
5.2	L'erreur relative (MAPE) pour la prévision de la consommation d'électricité en France pendant deux semaines	95
5.3	L'erreur relative (MAPE) pour les jours fériés de 2014 en France	97
5.4	Résultats de prévision pour le 26 Décembre 2014, à 00 :00	98
5.5	L'erreur relative (MAPE) pour les données de charge de la Nouvelle-Galles du Sud	101
5.6	Le temps de calcul de toutes les méthodes considérées dans quatre cas de prévision différents	103

Table des matières

Liste des Abréviations	vi
Introduction générale	xiv
1 Problématique et objectifs	1
1.1 Introduction	1
1.2 Équilibre de l'offre et de la demande	1
1.2.1 Ajustement de l'offre	2
1.2.2 Ajustement de la demande	3
1.3 Etat de l'art	4
1.3.1 Modèles d'utilisation finale et modèles économétriques	4
1.3.2 Méthodes classiques et techniques d'intelligence artificielle	6
1.3.3 Méthodes de combinaison de prévision	7
1.3.4 Prévision des charges atypiques et prévision en ligne	8
1.4 Domaine d'intérêt	10
1.5 Facteurs influant sur la charge électrique	11
1.5.1 Catégorie des consommateurs	12
1.5.2 Facteurs météorologiques	12
1.5.3 Évolution récente de la charge	12
1.5.4 Jour de la semaine	13
1.5.5 Heure de la journée	13
1.5.6 Effets saisonniers	13
1.5.7 Événements spéciaux	13
1.5.8 Perturbations aléatoires	13
1.6 Problématique	13
1.7 Objectif de la thèse	14
1.8 Conclusion	15
2 Généralités et notions de base	16
2.1 Introduction	16
2.2 Notions de base sur la charge électrique	16
2.3 Notions sur les séries temporelles	20
2.3.1 Définition d'une série temporelle	20
2.3.2 Composantes d'une série temporelle	22
2.3.3 Modélisation d'une série temporelle	24
2.3.4 Méthodes pour la prévision dans une série chronologique	24
2.4 Analyse préliminaire des données de charge électrique	26
2.4.1 Prétraitement des données	27
2.4.2 Classement et classification des données de charge électrique	28
2.4.3 Décomposition des signaux de charge	31

2.5	Prévision de la charge électrique	33
2.6	Mesures des erreurs de prévision	33
2.6.1	Erreur absolue en pourcentage	35
2.6.2	Déviati on absolue moyenne	35
2.6.3	Erreur quadratique moyenne	35
2.6.4	Racine carrée de l'erreur quadratique moyenne	36
2.6.5	Erreur absolue moyenne en pourcentage	36
2.7	Intervalle de prévision	36
2.8	Conclusion	38
3	Évaluation des méthodes de prévision à moyen et à long-terme de la charge électrique	39
3.1	Introduction	39
3.2	Données de charge utilisées	39
3.3	Méthodes proposées pour la prévision des pics de charge	40
3.3.1	Lissage exponentiel de Holt	41
3.3.2	Modèle autorégressif à moyenne mobile intégré	42
3.3.3	Lissage exponentiel de Holt-Winters	42
3.3.4	Modèle saisonnier autorégressif à moyenne mobile intégré	43
3.3.5	Modèle de réseau de neurones artificiel à rétro-propagation	43
3.3.6	Système d'inférence neuro-flou adaptatif	47
3.4	Application pour la prévision à moyen et à long-terme de la charge électrique	58
3.5	Conclusion	63
4	Méthodologie hybride adaptative à deux étapes dédiée à la prévision de la pointe journalière	64
4.1	Introduction	64
4.2	Série chronologique de la demande de pointe journalière d'électricité en Algérie	64
4.3	Méthodologie proposée pour la prévision de la pointe journalière	66
4.3.1	Méthode de Holt-Winters-Taylor	66
4.3.2	Classification par la méthode des c-moyennes floues	67
4.3.3	Algorithme des k-plus proches voisins	70
4.3.4	Décomposition en paquets d'ondelettes	72
4.3.5	Système d'inférence neuro-flou adaptatif	75
4.3.6	Résultat final pour la prévision de la pointe journalière	75
4.4	Evaluation de la méthodologie proposée pour la prévision des pics quotidiens de charge	78
4.5	Conclusion	82
5	Prévision en ligne basée sur une méthodologie efficace de combinaison de prédictions	83
5.1	Introduction	83

5.2	Analyse et traitement des séries chronologiques de la demande d'électricité	83
5.3	Méthodologie proposée pour la prévision en ligne de la charge électrique	86
5.3.1	Modèles primaires de prévision	88
5.3.2	Une procédure d'identification de valeurs aberrantes pour éviter les grandes erreurs de prévision	91
5.3.3	Estimation finale de la charge électrique	92
5.4	Application de HFCM-TM pour la prévision à très court-terme de la charge électrique	92
5.4.1	Évaluation des prévisions de la consommation française d'électricité	95
5.4.2	Évaluation des prévisions de la consommation australienne d'électricité	99
5.4.3	Présentation de l'outil proposé pour la prévision en ligne de charge	100
5.5	Conclusion	104
6	Conclusions et perspectives	106
6.1	Conclusions	106
6.2	Perspectives	108
	Bibliographie	110

Introduction générale

La demande d'énergie électrique dans les domaines industriel, commercial et résidentiel ne cesse d'augmenter de manière significative. Pour veiller à ce que l'énergie soit exploitable par tous les utilisateurs, les compagnies d'électricité doivent faire face à un défi important qui consiste à adapter en permanence leur niveau de production à celui de consommation. Cela présente un intérêt particulier pour les réseaux électriques modernes où de nombreux pays ont opté pour la déréglementation et la libéralisation du marché de l'énergie. L'amélioration des stratégies d'exploitation, de planification et de gestion des charges est devenue donc un objectif fondamental dans le cadre de l'avènement de charges atypiques au sein des systèmes électriques en constante évolution multiformes.

Prendre de bonnes décisions en termes de gestion d'un réseau électrique n'est guère possible si l'on ne dispose pas des techniques efficaces permettant de faire face au caractère difficilement stockable de l'énergie électrique et de mettre en disposition une source de service continu, de prix bas et d'énergie fiable. En particulier, des outils permettant l'étude et la prévision de la consommation d'électricité sont indispensables pour garantir l'équilibre offre/demande et l'exploitation sûre du réseau électrique. Offrant une estimation avancée pour une répartition optimale de l'énergie électrique est un défi important non pas seulement pour les transactions de l'énergie sur les marchés concurrentiels de l'électricité, mais aussi pour les pays en développement à ressources limitées où la capacité d'importation et d'exportation d'électricité est limitée. Des modèles précis de prévision de charge sont nécessaires pour une variété d'horizons temporels. Ceux-ci pourraient être énumérés comme suit : à très court-terme de quelques minutes à une heure, à court-terme de quelques heures à une semaine, à moyen-terme de quelques semaines à un an et à long-terme pour plusieurs années.

La prévision à long-terme de la charge électrique (LTLF) est indispensable pour permettre aux compagnies d'électricité de disposer d'une vision prospective afin de prendre des décisions importantes concernant les investissements nécessaires pour la planification d'un nouveau réseau et les extensions des réseaux électriques existants. Ces types de prévisions sont souvent de nature complexe en raison des incertitudes futures telles que les facteurs politiques, situation économiques, ...etc.

La prévision à moyen-terme de la charge électrique (MTLF) joue un rôle capital pour prendre des décisions importantes concernant la planification de la maintenance et l'évaluation de la sécurité de tout système de service public. Les prédictions exactes à moyen-terme sont indispensables pour trouver des périodes durant lesquelles des opérations d'entretien de lignes ou de postes pourront être réalisées. Ces opérations immobilisent le réseau puisqu'elles conduisent à mettre l'un de ses éléments hors service pour plusieurs semaines. Il faut donc garantir que l'exploitation restera suffisamment sûre pendant les périodes à venir. La prévision à moyen-terme fournit également un élément fondamental d'informations qui seront utilisées pour

les négociations de contrats bilatéraux et le développement des stratégies rentable pour l'achat de carburant.

La prévision à court-terme de la charge électrique (STLF) est un facteur essentiel pour assurer l'équilibre physique global entre l'offre et la demande à l'échelle d'un pays. Cette prévision est généralement effectuée pour prévoir le minimum et le maximum de la demande quotidienne d'électricité, l'énergie totale, ainsi que la courbe de charge journalière. Elle servira également comme une base pour l'engagement des unités de production, la coordination hydro-thermique, l'évaluation de l'échange, l'écoulement de charge, l'analyse de contingence et comme un élément important dans la composition de l'expédition et le processus de tarification économique.

La prévision à très court-terme de la charge électrique (VSTLF) est une tâche essentielle pour la gestion en temps réel de l'ensemble du réseau électrique, le contrôle de fréquence, ainsi que pour les fonctions de dispatching économiques. La moindre erreur dans la prédiction pourrait sérieusement affecter les niveaux de production, en particulier pour les grands services publics d'électricité. Des prévisions précises facilitent le problème de la production et la gestion de charge dans une large mesure. Ces prévisions intégrées avec les informations sur les transactions programmées, les exigences en matière de coûts de production, le marché au comptant et les réserves imposées par un opérateur de réseau indépendant sont utilisées pour déterminer la meilleure stratégie pour les ressources de service.

Sans réplication, la prévision de la charge électrique est une fonction essentielle dans les activités opérationnelles quotidiennes des compagnies d'électricité. La prévision des charges futures est également importante pour la planification du réseau et le développement des infrastructures. Donc, il est nécessaire de prédire la charge future d'une manière à déterminer une estimation de la quantité de puissance nécessaire et de se préparer adéquatement pour la production. C'est dans cette optique que cette thèse est orientée. Elle s'intéresse, en effet au développement des techniques et méthodologies permettant de prévoir efficacement la charge électrique des différents horizons temporels.

À ce propos, le manuscrit est constitué de cinq grands chapitres, précédés d'une introduction générale en plus d'une conclusion générale et des perspectives. Dans le premier chapitre nous allons présenter un bref aperçu sur l'importance de la prévision de charge des systèmes électriques, l'état de l'art du problème considéré, ainsi que les différents objectifs de la thèse. Le deuxième chapitre est consacré à une présentation de quelques notions de base sur les variations de la consommation d'électricité et les différentes méthodes et démarches à suivre pour assurer les prévisions satisfaisantes de la charge électrique. Le troisième chapitre utilise les observations de la demande d'électricité en Algérie pour évaluer des méthodes de prévision à moyen et à long-terme de la charge électrique. Le quatrième chapitre décrit la prévision des pics quotidiens de la charge électrique en Algérie en utilisant une méthodologie hybride adaptative à deux étapes. Cette méthode consiste à assurer une prévision primaire de la charge électrique, et d'essayer ensuite d'améliorer la performance en utilisant un modèle secondaire de prévision. Dans le dernier chapitre, nous allons présenter une nouvelle méthodologie combinée pour assurer les prévisions en ligne et à très court-

terme de la demande d'électricité. L'idée de base pour cette étude consiste à utiliser un système de prévision qui a la capacité d'obtenir une bonne précision de prévision, assurer un temps de calcul faible et éviter les grandes erreurs de prédiction. Cette méthode sera également accompagnée d'un outil de prévision convenable, dont le but est de faciliter la tâche de prédiction à l'utilisateur et d'assurer des prédictions satisfaisantes et adaptatives quelle que soit la situation exposée.

Problématique et objectifs

1.1 Introduction

Avec la croissance vertigineuse des réseaux électriques et de l'augmentation de leur complexité, de nombreux facteurs sont devenus influents sur la production d'énergie électrique, la demande ou la gestion de charge. La prévision de charge est l'un des facteurs critiques pour le fonctionnement économique des systèmes électriques. Elle aide les compagnies d'électricité à mieux gérer les risques et réduire les coûts opérationnels. La prévision des charges futures est également importante pour la planification du réseau et le développement des infrastructures. Donc, il est nécessaire de prédire la charge future d'une manière à déterminer une estimation de la quantité de puissance nécessaire et de se préparer adéquatement pour la production. Ce chapitre résume les difficultés rencontrées dans le problème de prévision de charge, les facteurs influençant sur la consommation, l'état de l'art du problème et l'objectif de la thèse.

1.2 Équilibre de l'offre et de la demande

Le système électrique se définit comme étant l'ensemble indissociable assurant la production, le transport et la distribution d'électricité aux collectivités, entreprises et particuliers. Le système électrique est composé de quatre maillons : les moyens de production, le réseau de transport, les réseaux de distribution et les sites de consommation (cf. Figure 1.1). L'électricité est produite dans des centrales électriques (thermiques classiques, hydrauliques, photovoltaïque, éolienne...), qui peuvent fournir de l'électricité à des millions de personnes. Ensuite, elle est transportée et acheminée sur de grandes distances vers les consommateurs. Le transport et la distribution de l'électricité se font à l'aide d'un réseau électrique de lignes aériennes et souterraines, dans le but d'alimenter tous les consommateurs (usines, maisons, ...etc). Sachant que l'énergie électrique est non stockable, l'électricité est immédiatement acheminée vers l'endroit où elle sera consommée. Peu importe la quantité demandée, le fournisseur d'électricité doit y répondre instantanément et de manière efficace aux variations de la demande.

Comme il est montré dans la Figure 1.2, il est indispensable d'avoir un équilibre entre la production et la consommation à tout instant pour l'ensemble du réseau électrique, tout en assurant le maintien de la fréquence dans les plages contractuelles,

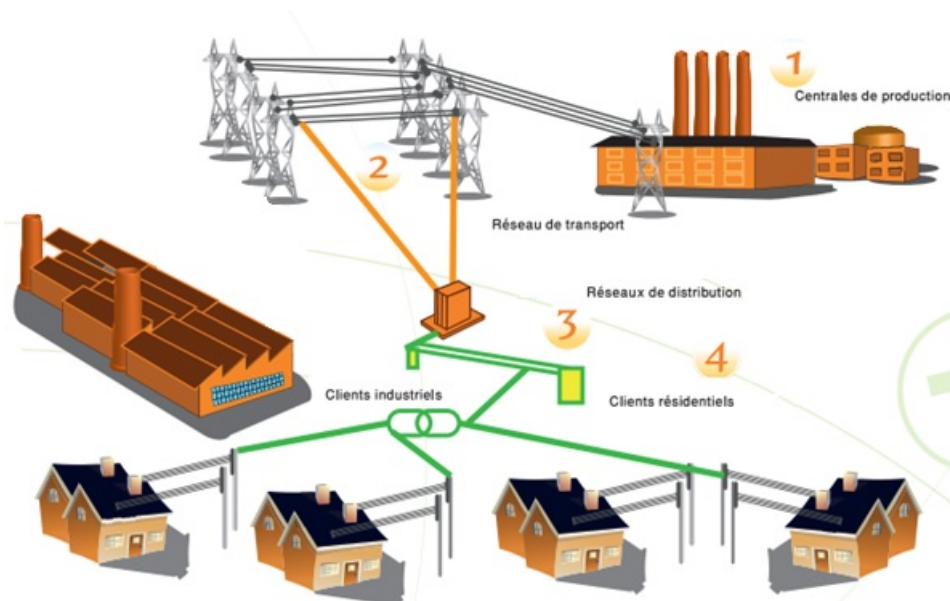


FIGURE 1.1 – Constitution d'un réseau électrique

la gestion de la réserve du parc de production et, éventuellement, les échanges internationaux.

En conséquence, toute la chaîne de production, transport et distribution doit permettre une adaptation permanente entre l'offre et la demande d'électricité. En pratique, cette adaptation constante peut résulter de 2 types d'ajustements : ajustement de l'offre et ajustement de la demande

1.2.1 Ajustement de l'offre

L'ajustement de l'offre consiste, à court-terme, à adapter minute par minute la production des centrales disponibles et, à plus long-terme, à construire de nouvelles capacités de production et de distribution. Les compagnies d'électricité prévoient trois classes de centrales de génération :

- les centrales de base qui débitent leur pleine capacité en tout temps. Les centrales nucléaires et les centrales à production fatale (éoliennes, solaire, hydrauliques sur les fleuves) sont particulièrement aptes à remplir ce rôle. L'ensemble de ces centrales fournit une grande partie de la puissance demandée par les consommateurs.
- les centrales intermédiaires de puissance moyenne qui peuvent réagir rapidement aux fluctuations de la demande.
- les centrales de pointe qui ne débitent leur pleine capacité que pendant de courtes périodes. Les centrales de pointe doivent être mises en marche dans un délai très court ; elles utilisent donc des turbines à combustion, centrales hydroélectriques à réservoir et centrales de pompage-turbinage.

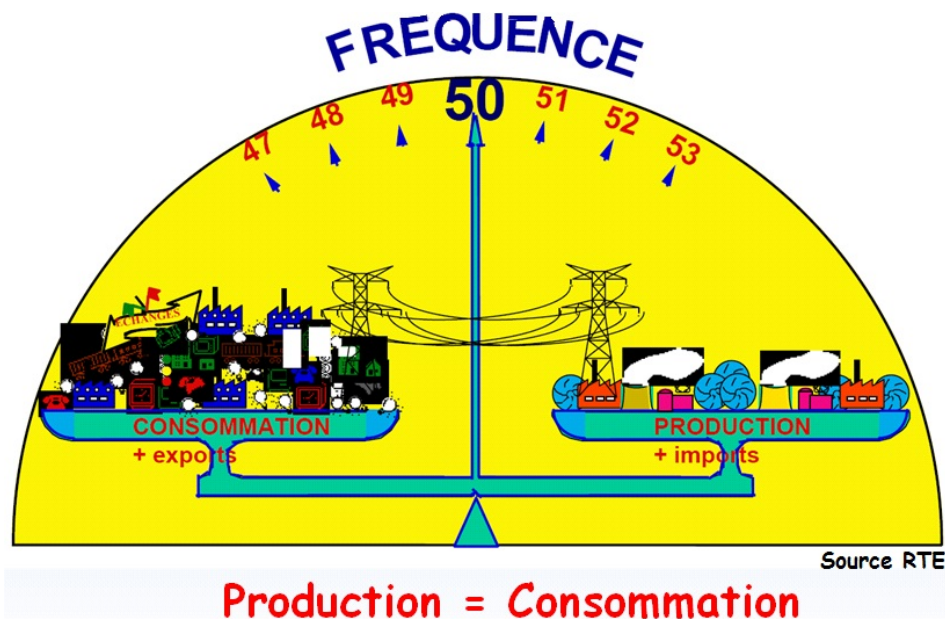


FIGURE 1.2 – Équilibre entre la production et la consommation d'énergie électrique

En France par exemple, le gestionnaire du réseau de transport d'électricité (RTE) effectue le placement des centrales sur la courbe de charge, qui consiste, pour chaque quart d'heure, à empiler les centrales par ordre de prix croissant, jusqu'à la couverture complète de la demande électrique. La Figure 1.3 donne un exemple réel de placement sur la courbe de charge pour un jour ouvrable typique. La production totale ne couvre pas seulement la consommation finale française, mais aussi les exportations et le pompage. À la base du diagramme sont placées les centrales éoliennes, solaire, bioénergie ; ensuite vient le nucléaire, dont le coût marginal est très bas ; puis viennent les centrales thermiques à gaz et les centrales hydrauliques.

1.2.2 Ajustement de la demande

Le système électrique est, en quelque sorte, piloté par la consommation. Seule une prévision rigoureuse permet à l'opérateur du système électrique de déterminer les moyens de production nécessaires pour satisfaire la demande des consommateurs et assurer l'équilibre globale pour les différents horizons temporels.

Dans le cas d'une prédiction à court-terme, la prévision d'une courbe de charge est généralement effectuée la veille pour le lendemain, et réactualisée en cours de la journée envisagée. A titre d'exemple, la Figure 1.4 reprend un tel diagramme journalier pour la consommation française d'électricité. RTE ajuste à tout moment les volumes de production aux besoins en électricité des consommateurs. Chaque jour, RTE effectue une prévision primaire de consommation pour le lendemain, et la corrige encore une fois pendant la journée envisagée. Ces prévisions sont constituées de 96 valeurs de puissance, correspondant aux 96 quarts d'heure d'une journée. Elles

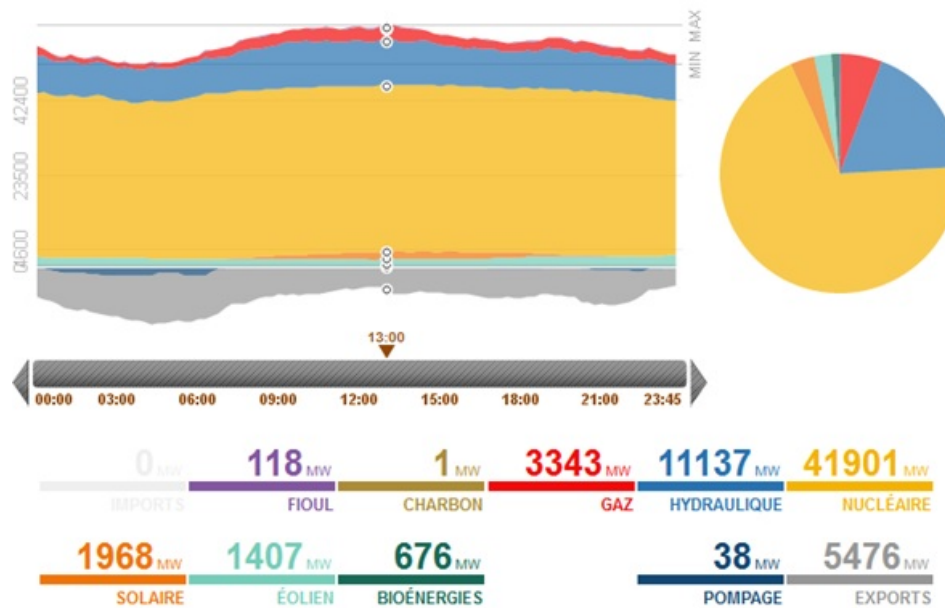


FIGURE 1.3 – Placement des centrales de production sur la courbe de charge française : 01/06/2016 [1]

servent de références pour assurer que les programmes de production prévus par les différents fournisseurs d'électricité permettent de satisfaire à l'équilibre de l'offre et de la demande.

1.3 Etat de l'art

Dans la bibliographie, on peut trouver un large éventail de méthodes pour la prévision de la demande d'électricité. La classification est basée sur certaines caractéristiques, comme par exemple le type de données à fournir au modèle, le temps de calcul informatique requis, ainsi que l'horizon temporel à envisager. Les modèles d'utilisation finale, les modèles économétriques, et leurs combinaisons font partie des premières méthodes utilisées pour la prévision à moyen et à long-terme de la charge électrique ; tandis que les techniques d'intelligence artificielle, les méthodes hybrides et les combinaisons de prévision représentent aujourd'hui un axe de recherche très important pour les deux domaines de prévision à très court et court-terme de la charge électrique.

1.3.1 Modèles d'utilisation finale et modèles économétriques

L'approche d'utilisation finale estime directement la consommation d'électricité en utilisant des informations détaillées telles que les descriptions des appareils utilisés par les clients, les dimensions des maisons, l'âge de la clientèle, les changements

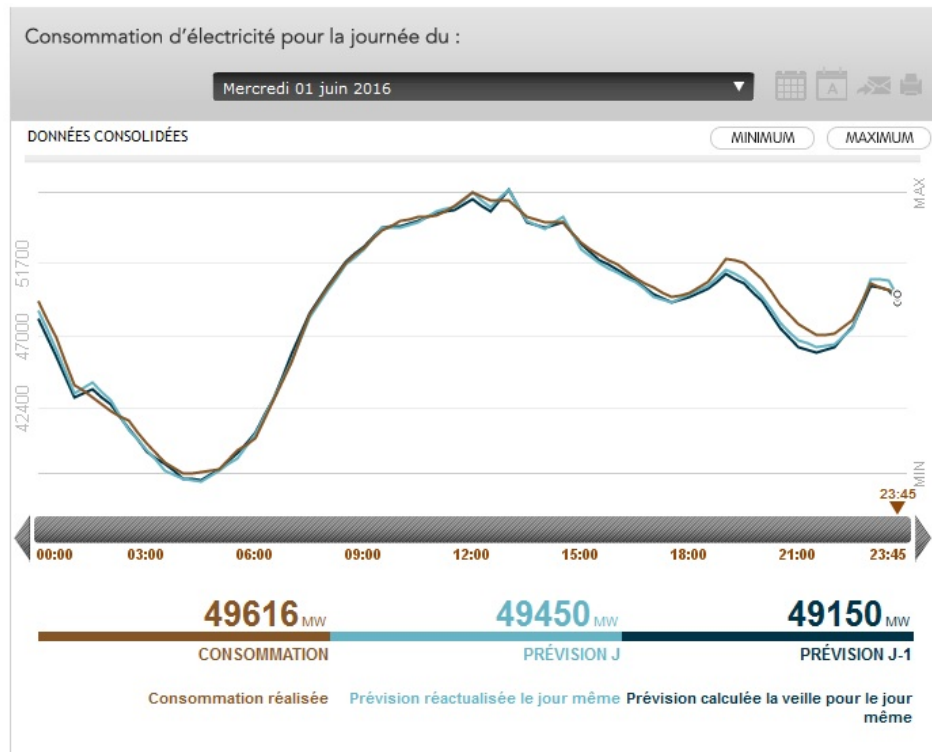


FIGURE 1.4 – Courbe de charge française : 01 Juin 2016

technologiques, le comportement des clients, et la dynamique des populations [2]. L'idée de base de l'analyse de l'utilisation finale consiste à supposer que la demande d'électricité dépend de ce qu'elle est utilisée pour (l'éclairage, le refroidissement, le chauffage, la réfrigération, ...etc). Par exemple, en étudiant les données historiques de la quantité d'électricité utilisée pour chaque appareil électrique dans une maison, le nombre projeté d'appareils dans chaque maison, et le nombre projeté de maisons ; une estimation de la quantité l'électricité nécessaire pour faire fonctionner tous les appareils dans une zone géographique au cours d'une période à venir pourra être déterminée. En utilisant des techniques similaires pour l'électricité utilisée dans le commerce et l'industrie, puis en ajoutant les totaux pour les secteurs résidentiels, commerciaux et industriels ; une prévision totale de la demande d'électricité peut être dérivée. Les modèles d'utilisation finale se concentrent donc sur les différentes utilisations de l'électricité dans le secteur résidentiel, commercial et industriel. Idéalement, cette approche est très précise. Cependant, il est sensible à la qualité des données utilisées. Par exemple, dans ce procédé, la distribution de l'âge de l'équipement est importante pour certains types d'appareils. Les prévisions d'utilisation finale nécessitent moins de données historiques, mais plus d'informations sur les clients et leurs équipements.

L'approche économétrique utilise des facteurs économiques tels que le revenu

par habitant, les niveaux d'emploi et le prix de l'électricité pour prévoir la demande d'électricité. Les changements démographiques, le développement économique, la construction industrielle et le développement technologique sont également utilisés pour la prévision à long-terme de la charge électrique. Cette démarche utilise des équations mathématiques complexes pour montrer les relations entre la demande d'électricité et les facteurs qui influent sur cette demande. Par exemple, on peut montrer à l'aide d'une équation comment la demande d'électricité dans le passé a réagi à la croissance de la population, les changements de prix, ...etc. Pour chaque facteur influençant, l'équation peut montrer si le facteur a entraîné une augmentation ou une diminution de la demande d'électricité, ainsi que l'importance (en pour cent) de l'augmentation ou de la diminution. L'équation est ensuite testée et affinée afin de s'assurer une représentation des relations passées aussi fiable que possible. Une fois cela fait, les valeurs projetées des facteurs affectant la demande (population, revenu, prix) sont mis dans l'équation pour faire les prévisions. Une procédure similaire est suivie pour toutes les équations du modèle [3]. L'un des premiers modèles économétriques c'est celui développé par la commission du commerce de l'État de l'Iowa, et utilisé pour prévoir la charge horaire (y compris les pics) pour les services publics de l'Iowa [4]. Le modèle est composé par 24 équations, une pour chaque heure de la journée où des variables indépendantes tels que le prix d'électricité et le revenu ont été considéré. Le modèle a été pratiquement utilisé pour prévoir la charge pour un horizon temporel de quelques mois à l'avance.

Cependant, l'analyse de l'utilisation finale et les méthodes économétriques nécessitent une grande quantité d'informations pertinentes aux appareils, clients, économie, ...etc. Leurs applications sont complexes et nécessitent une participation d'un expert. De plus ces informations sont souvent pas disponibles pour le grand public et ce genre de prévisions est généralement conservé et prendra en charge exclusivement par les compagnies de service public. On peut citer, à titre d'exemples, les modèles à utilisation finale REEPS [5], LBL [6], MEDEE [7] et DELMOD [8]. Ces modèles ont été développés pour prévoir la consommation résidentielle dans les États-Unis et l'Autriche, et ont été mis en œuvre également dans de nombreux pays, y compris l'Australie.

1.3.2 Méthodes classiques et techniques d'intelligence artificielle

Afin de simplifier les prévisions à moyen et à long-terme, de les rendre plus précises, et d'éviter l'utilisation d'une grande quantité d'informations détaillées, une variété de modèles basés sur des techniques statistiques classiques ou d'intelligence artificielle a été élaborée par les chercheurs dans le domaine de la prévision de la charge électrique. Les techniques statistiques incluent les modèles de régression [9], les modèles de série temporels de Box-Jenkins[10], et le filtre de kalman [11]. Ces approches classiques se subdivisent encore en deux catégories : l'une se concentre exclusivement sur les propriétés temporelles de la variable étudiée (méthodes univariées) tandis que l'autre considère la relation entre la variable étudiée, son passé et d'autres variables explicatives . L'avantage de ces modèles est la simplicité de

modélisation. Cependant, les techniques univariées telle que la méthode de séries temporelles ARIMA (processus autorégressif à moyenne mobile intégré) sont parfois imprécises, car ils ignorent des effets exogènes importants. Même si les modèles classiques à plusieurs variables telles que la régression expriment la charge électrique en fonction des facteurs exogènes, y compris les variables sociales et économiques, le processus de modélisation linéaire de ces méthodes fait face à une grande difficulté pour déterminer empiriquement la relation correcte qui existe entre la demande d'électricité et les autres entrées explicatives. Au cours des années récentes, l'avènement de nouvelles techniques de l'intelligence artificielle a permis de modéliser automatiquement les relations non linéaires et complexes qui existent entre les données de charge électrique et les différents facteurs influençant sur la consommation d'électricité. Par conséquent, de nombreuses études ont porté ces dernières années sur l'application des techniques d'intelligence artificielle pour la prévision à moyen et à long-terme de la charge électrique. Ces approches comprennent principalement les réseaux de neurones [12], les systèmes flous [13], les systèmes neuro-flous [14], et les systèmes experts [15].

En vue du rôle vital des prévisions pour les prochaines minutes ou le jour suivant dans les opérations quotidiennes des compagnies d'électricité, beaucoup d'attention a été attirée dans ces dernières années pour améliorer la précision des prévisions à très court-terme et à court-terme de la charge électrique. Par conséquent, une grande variété de techniques a été étudiée dans la littérature pour ces deux horizons temporels au fil de ces dernières années, tant avec les outils classiques qu'avec les méthodes de l'intelligence artificielle. Parmi ces modèles, on peut citer les méthodes de lissage exponentiel [16], les processus de Box-Jenkins [17], l'approche de régression [18], les approches basées sur les réseaux de neurones [19], la logique floue [20], les systèmes neuro-flous [21], et les approches hybrides [22], [23], [24], [25], et [26].

1.3.3 Méthodes de combinaison de prévision

Cependant, ce n'est pas une tâche facile d'assurer les prédictions précises de la charge, car le comportement de la demande d'électricité est complexe et la série chronologique de la charge est non-stationnaire dans la moyenne et la variance, contenant des termes de tendance, de multiples cycles de fluctuations saisonnières, ainsi qu'un comportement de variation aléatoire. En fait, il n'y a pas de méthode de prévision qui donne le meilleur résultat pour tous les temps : certaines approches peuvent performer très bien dans certaines circonstances, et ne refléteraient pas les variations de la charge dans d'autres [27]. Par conséquent, la méthodologie de combinaison de prévision, également appelée la prévision d'ensemble, est un outil approprié qui a été appliqué avec succès dans de nombreuses disciplines pour améliorer la précision des prévisions [28].

Les nouvelles études de recherche sur le domaine de la prévision de charge montrent que la stratégie de combiner différents modèles est très utile pour améliorer les performances de prévision. La façon la plus simple de combiner les prévisions consiste à attribuer un poids égal à tous les prévisions individuelles. Cette méthode

fonctionne raisonnablement bien quand nous sommes incertains de ce qui est la meilleure méthode. Donc, en pratique, on peut apporter des améliorations importantes dans la performance de prévision de charge en utilisant tout simplement la moyenne simple des prévisions. Dans l'étude donnée par James W. Taylor [29], plusieurs méthodes de prévision ont été évaluées en utilisant les données britanniques de la demande d'électricité, et les résultats ont montré que la moyenne simple des prévisions issues de l'approche fondée sur les conditions météorologiques et la méthode de lissage exponentiel à double saisonnalité de Holt-Winters-Taylor (HWT) a donné lieu à des prévisions qui ont surpassé les performances de toutes les autres méthodes au-delà d'une heure d'avance. Une autre méthode populaire pour combiner les prévisions est la pondération inégale en fonction de la performance des prévisions passées des modèles. Dans l'approche proposée par J. Wang et al. [30], la charge à court-terme est obtenue en combinant les prévisions générées par le modèle saisonnier autorégressif à moyenne mobile intégré (SARIMA) et le modèle de réseau de neurones (ANN), et les poids de la combinaison sont déterminés en utilisant l'approche variance-covariance. Les résultats de cette étude ont montré une amélioration importante de la précision de prévision par rapport à d'autres méthodes. Un modèle combiné pour la prévision à court-terme de charge basé sur la décomposition en mode empirique d'ensemble (EEMD) et l'optimisation par essaim de particules (PSO) est proposé par Z.Liu et al. [31]. Tout d'abord, la séquence de charge est décomposée en un nombre limité de composants de fonction du mode intrinsèque (IMF) et un reste par EEMD. Ensuite, les futures séquences sont prédites avec des modèles de prévision appropriés. PSO est proposée et utilisée pour optimiser les pondérations de combinaisons. Les résultats numériques indiquent que le modèle proposé dans cette étude obtient un taux de précision des prévisions plus élevé en comparaison avec d'autres techniques. Un certain nombre de recherches récentes ont également suggéré l'utilisation de la combinaison des prévisions comme un outil prometteur pour améliorer les performances des prévisions de la demande d'électricité [32], [33], [34].

1.3.4 Prévision des charges atypiques et prévision en ligne

Néanmoins, l'inconvénient des modèles combinés cités ci-dessus est la non-consideration des grandes erreurs de prévision. En fait, la plupart des méthodes existantes dans la littérature de la prévision de charge électrique se concentrent seulement sur l'évaluation de la performance dans des conditions normales et typiques, et ignorent souvent l'effet de la charge anormale qui se produit au cours des jours spéciaux (tels que les jours fériés). Par exemple, si un modèle combiné construit par une combinaison pondérée est mis en œuvre pour prévoir la charge d'un jour férié particulier, une mauvaise estimation pourrait être obtenue au cas où une grande erreur de prévision est obtenue par la méthode individuelle qui avait un poids plus élevé. Cette mauvaise estimation peut être attribuée à des comportements de charge des jours atypiques qui se diffèrent par rapport à ceux des jours normaux. De plus, pour un jour spécial donné, le comportement de la charge peut varier d'une année à

l'autre, selon le jour et la saison de l'année où la charge anormale est survenue. Il est donc plus compliqué et plus difficile d'établir un modèle précis pour la prévision des charges anormales par rapport au cas de la construction d'un modèle de prévision des charges normales.

Peu d'études ont été menées pour développer des techniques précises de prévision des charges atypiques. Certains d'entre eux accordent plus d'attention à la qualité des données et la rigueur de l'analyse des séries chronologiques de charge. Certains autres papiers visent à fournir la sortie appropriée d'une combinaison de prévisions. Alors que d'autres travaux, comme l'étude réalisée par M. Ghofrani et al. [35], se concentrent sur la considération à la fois de la combinaison optimale de prévisions et la capture de l'information la plus pertinente de la série chronologique de charge comme deux éléments clés pour effectuer la tâche de l'amélioration de l'exactitude des prévisions durant les conditions atypiques de charge. L'une des méthodes existantes dans la littérature qui mettent l'accent sur la qualité des données pour prédire la charge atypique est la méthode de lissage exponentiel à triple saisonnalité développée par S. Arora et J. W. Taylor dans [36]. La contribution méthodologique de ce travail est de montrer comment la méthode de lissage exponentiel peut être adaptée pour modéliser la charge pour les jours spéciaux, lorsqu'elle est utilisée en conjonction avec une approche basée sur des règles. La méthode à base de règles proposée par les auteurs a été capable de modéliser la charge normale et anormale dans un cadre unifié. Cependant, ce modèle est trop compliqué, nécessite une longue chaîne de valeurs de charge, et une intervention fréquente de la part du prévisionniste est pratiquement essentielle pour restructurer le modèle de prévision. Dans une autre étude, un travail de recherche visant à améliorer la prévision à court-terme pour les conditions anormale de charge a été présenté [37]. La méthode de cette étude utilise des modèles de réseaux neurones artificiels pour fournir la courbe de charge prévue et des modèles d'inférence floue pour prévoir les charges maximales et minimales de la journée spéciale. Les charges de la journée spéciale sont prédites en combinant les résultats de l'ANN et de la méthode d'inférence floue. L'avantage de cette structure hybride était d'utiliser les avantages des deux, c'est-à-dire la capacité de généralisation de l'ANN et la capacité d'inférence floue pour manipuler et formaliser l'expérience et la connaissance des prévisionnistes. Les résultats de la recherche ont montré que la méthode de prévision hybride pourrait apporter une amélioration considérable de l'exactitude des prévisions pour les jours spéciaux. Dans Réf. [38] M. Ghayekhloo et al ont présentés une structure hybride de prévision de charge avec un algorithme de prétraitement des données pour améliorer la précision de prévision. Le procédé de prétraitement des données est une combinaison de la sélection d'entrée, de la décomposition en ondelettes et d'une nouvelle normalisation pour fournir les entrées les plus appropriées aux algorithmes des réseaux de neurones bayésiens. L'algorithme génétique est utilisé pour optimiser les coefficients de pondération des différentes composantes de prévision. La performance et l'exactitude de cette méthode de prévision à court-terme a été évaluée à l'aide des données de charge de la Nouvelle-Angleterre. Les résultats de cette étude ont montré une amélioration significative de l'exactitude des prévisions par rapport aux autres

techniques existantes.

Nous aimerions cependant souligner que la plupart des méthodes existantes dans la littérature de la prévision de la charge électrique ont été mis en œuvre hors ligne en utilisant des bases de données simulées. Cependant, la bonne utilisation de tout algorithme de prévision dans la pratique exige non seulement la prise en compte d'une précision raisonnable, mais exige également l'inclusion de beaucoup d'autres paramètres, tels que la capacité d'adaptation en ligne, la simplicité d'utilisation et le temps de calcul faible.

1.4 Domaine d'intérêt

Comme nous l'avons détaillé précédemment, la prévision de charge se décompose en quatre catégories : prévision à long-terme, prévision à moyen-terme, prévision à court-terme, et prévision à très court-terme. Ainsi, chaque catégorie a ses propres méthodes ou techniques de prévision. Cependant, les chercheurs ont toujours trouvés des difficultés pour assurer une précision satisfaisante pour un horizon aussi long que plusieurs mois ou plusieurs années à l'avance. Par conséquent, tout modèle de prévision à moyen ou à long-terme de la charge est imprécis par nature. Cette imprécision est attribuée principalement à la difficulté d'obtenir des informations fiables sur les facteurs climatiques, économiques et politiques ; qui pourraient influencer sur la consommation d'électricité durant les prochaines mois ou années à venir. D'un autre côté, l'incertitude peut être diminuée à un grand degré dans le cas de prévision à très court-terme et à court-terme de la demande d'électricité, où des données comme la température, le type du jour, le jour de la semaine, et la saison sont facilement accessibles et exploitables. C'est pour cela que beaucoup de travaux ont été effectués dans la littérature de prévision à très court et à court-terme de la charge en comparaison avec les deux autres horizons temporels.

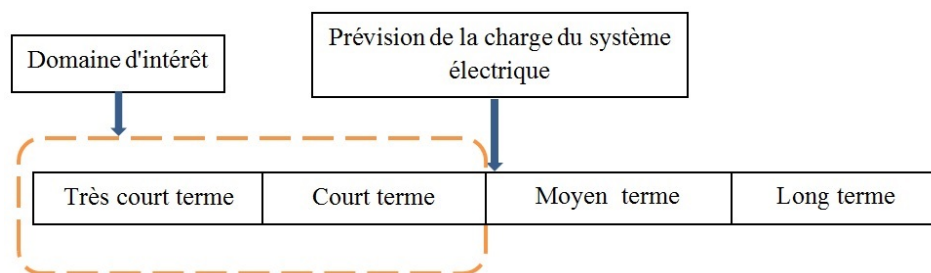


FIGURE 1.5 – Les différents horizons de temps pour la prévision de la charge électrique

Cela sera encore respecté dans notre thèse, où nous donnerons beaucoup d'attention pour la prévision à très court et à court-terme de la charge électrique. La Figure 1.5 tente de clarifier le domaine d'intérêt de cette recherche. Cependant, la plupart des travaux effectués dans la littérature de ces deux derniers horizons

temporels s'intéressent seulement sur la prévision pour des conditions de charge normales, et les charges atypiques (tel que le cas des jours fériés) sont souvent ignorées. Contrairement à ces travaux, on s'intéresse dans notre thèse au développement des méthodes permettant d'assurer des prévisions satisfaisantes pour à la fois les conditions normales et anormales de la charge électrique.

1.5 Facteurs influant sur la charge électrique

La tendance de charge est différente d'un moment à un autre et la consommation d'électricité est influencée par plusieurs facteurs. Cependant, les facteurs affectant la charge à différents horizons temporels ne sont pas nécessairement les mêmes. La variation de la charge à court-terme dépend fortement des conditions climatiques et les facteurs de temps comme l'heure de la journée, le jour de la semaine ...etc. La charge de moyen à long-terme est déterminée par des facteurs tels que la croissance de la population, les revenus du capital, les facteurs démographiques, le produit intérieur brut ...etc. La plupart des compagnies de services publics utilisent cette dépendance distinctive dans le choix des variables d'entrée pour le modèle de prévision.

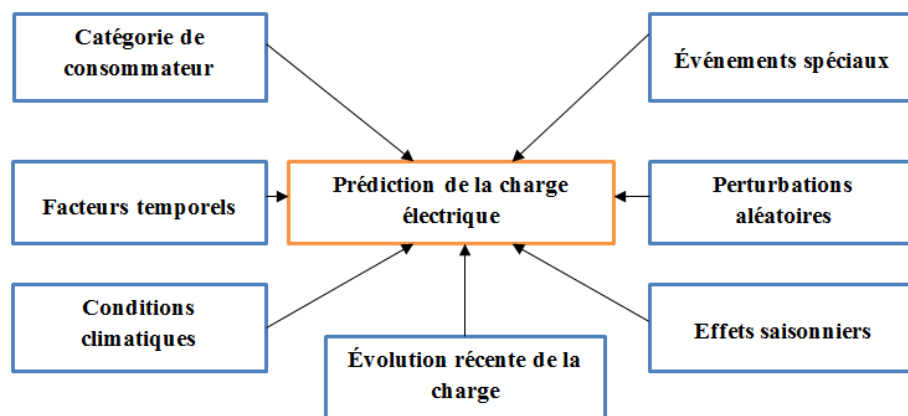


FIGURE 1.6 – Facteurs d'influence possibles pour la prévision à très court et court-terme de la charge

Certains des facteurs mis en évidence dans la Figure 1.6 pourraient être considérés pour la prévision à court-terme et à très court-terme de la charge électrique. Cependant, les conditions météorologiques pourraient être considérées comme un facteur important à introduire dans certains modèles de prévision à court-terme pour améliorer la précision de prévision [39], tandis que la tendance récente de la charge électrique et l'historique de la consommation d'électricité seraient largement suffisants pour le cas de prévision à très court-terme de la charge électrique [40].

1.5.1 Catégorie des consommateurs

L'évolution de la consommation d'électricité est unique pour les clients qui appartiennent à la même catégorie. Normalement, les clients de type industriel consomment une grande partie de l'électricité. Les charge résidentielles sont normalement les plus petits et souvent les plus complexes à déterminer en raison de différentes routines d'utilisation d'électricité des consommateurs [41].

1.5.2 Facteurs météorologiques

Les données météorologiques sont les premiers éléments influençant la consommation d'électricité. Ainsi, les variations cycliques annuelles de la consommation a comme origine l'évolution des conditions météorologiques tout au long de l'année. Parmi ces facteurs, la température et la nébulosité sont fortement considérées dans les modèles de prévision réalisés par le gestionnaire du réseau de transport d'électricité de France.

- La température : La prise en compte des hypothèses de températures et de ses variations permet de quantifier la modification de consommation induite par l'utilisation du chauffage électrique en hiver ou de la climatisation en été. A titre d'exemple, RTE estime actuellement qu'en hiver, une variation moyenne d'un degré Celsius sur l'ensemble du territoire français peut entraîner une variation d'environ 2500 MW de la consommation à la pointe (soit l'équivalent de la consommation moyenne en hiver d'environ 2 millions de foyers). Il est estimé aussi que le gradient de température lié à la climatisation vaut environ 400 MW/°C [42].
- La nébulosité : La nébulosité représente le taux de couverture nuageuse. Cette grandeur s'exprime en Octa et varie de 0 à 8 (0 correspond à un ciel complètement dégagé et 8 à un ciel couvert). La nébulosité a une influence sur l'utilisation de l'éclairage, mais aussi sur le chauffage en modifiant les effets du rayonnement solaire dans les habitations. Une variation moyenne d'un Octa sur la France entraîne une variation d'environ 800 MW de la consommation [42].

1.5.3 Évolution récente de la charge

L'évolution récente de la charge est principalement le support de la prévision. Cependant, l'un des gros problèmes pour la construction d'un modèle de prédiction est de déterminer le nombre optimal des instances de consommation avant la période à prédire. Habituellement, l'ensemble de données est déterminé en prenant en considération une analyse de corrélation. Cependant, l'analyse effectuée par Santos [43], rapporte que l'utilisation d'une longue chaîne de valeurs de charge n'offre aucun type d'avantage dans la conception du modèle de prévision. Pour la prédiction de courte durée, des valeurs de charge de deux semaines proches pourraient être suffisantes pour concevoir une structure de base d'un modèle de prévision avec une précision raisonnable.

1.5.4 Jour de la semaine

En règle générale, chaque jour de la semaine a sa tendance unique de la demande d'électricité. En Algérie par exemple, la charge est légèrement stable dans les jours ouvrables, et la consommation dans les week-ends est très basse par rapport aux autres jours de la semaine. Parfois, les jours de la semaine pourraient être choisis comme intrants dans un modèle de prévision.

1.5.5 Heure de la journée

La charge a une caractéristique cyclique. Contrairement à la charge industrielle qui est légèrement stable, la charge résidentielle se change considérablement au cours de la journée. Les pics pour une charge résidentielle sont fréquemment observés dans le matin et dans le soir.

1.5.6 Effets saisonniers

Les changements saisonniers ont une influence directe sur les charges du système. Pour garantir la fourniture et la fiabilité, la plupart des entreprises de services publics considèrent le plus mauvais cas des scénarios possible (effets de variation saisonnière). La variation saisonnière aide les entreprises de service public pour déterminer le pic d'été ou d'hiver.

1.5.7 Événements spéciaux

La demande associée à certain événement pourrait influencer de manière significative la charge planifiée du système. Par exemple, la coupe du monde de football et ces activités liées peuvent influencer grandement la charge dans le système d'alimentation. Bien souvent, tels événements sont épineux à prendre en compte pour un modèle de prévision de charge.

1.5.8 Perturbations aléatoires

Contrairement aux consommateurs résidentiels, un léger changement dans les charges industrielles provoque une variation de charge importante. Par exemple, si on considère l'impact de la mise en route ou l'arrêt des charges importantes telles que les fonderies, les souffleries, ou d'une usine d'acier ; sans conteste, le changement sur la charge n'est pas négligeable.

1.6 Problématique

Malgré les nombreux travaux dans la littérature de la prévision de charge électrique publiés depuis les années 1960, ce problème est encore considéré comme difficile pour les chercheurs en raison du caractère fortement non linéaire de la série de charge électrique et la relation complexe avec les facteurs économiques, climatiques,

les modes d'utilisation, le jour de la semaine, l'heure de la journée, ...etc. Avec la croissance vertigineuse des réseaux électriques et de l'augmentation de leur complexité, l'estimation de la charge future avec des données historiques est devenue de plus en plus difficile.

Le problème le plus difficile est la prévision précise pour des conditions atypiques de charge, telles que le cas des variations rapides de la charge qui se produisent durant les jours atypiques. La prévision précise pour les jours atypiques constitue un aspect essentiel dans la prévision de charge des systèmes électriques. La moindre erreur dans la prévision pourrait gravement affecter les niveaux de production, en particulier pour les grandes compagnies d'électricité. Ainsi, les prévisions précises sont nécessaires pour la sécurité du système électrique et sa fiabilité globale. Cependant, le profil de charge (forme de la courbe de charge intra-journalière) diffère pour différents types de jour spécial. En outre, pour une journée spéciale donnée, comme un jour férié particulier, le profil peut varier d'une année à l'autre, en fonction du jour de la semaine et la saison de l'année où la charge atypique a été produite. En raison de ces complexités, la prévision de charge dans des conditions anormales a été largement omise dans la littérature de la prédiction de la charge électrique, et a été laissée donc à l'expérience et le jugement du contrôleur central du réseau électrique. Les quelques méthodes existantes dans la littérature pour la prévision des charges atypiques sont caractérisées par un processus de modélisation trop compliqué qui nécessite une longue chaîne de valeurs de charge, et une intervention fréquente de la part du prévisionniste est pratiquement essentielle pour restructurer le modèle de prévision. En outre, nous soulignons également que la plupart des méthodes de prévision dans la littérature ont été mises en œuvre hors ligne à l'aide de l'utilisation de données simulées. En fait, la bonne utilisation d'un algorithme de prévision de charge dans la pratique exige non seulement une précision raisonnable, mais nécessite également la considération de nombreux autres paramètres, comme la capacité d'adaptation en ligne, la simplicité d'utilisation et l'efficacité de calcul. Pour surmonter ces limitations, une approche qui peut assurer un temps de calcul faible et permet l'incorporation automatique d'une expertise antérieure des profils de charge est extrêmement nécessaire pour assurer les prévisions satisfaisantes durant à la fois les conditions normales et spéciales de la charge électrique.

1.7 Objectif de la thèse

Comme mentionné précédemment, l'objectif de notre travail consiste à assurer les prévisions précises de la charge électrique pour les différents horizons temporels. Cependant, au vu du rôle vital des prévisions pour les prochaines minutes ou le jour suivant dans les opérations quotidiennes des compagnies d'électricité, nous nous donnons plus d'attention à développer des modèles adaptatifs et intelligentes pour les deux cas de prévision à très court-terme et à court-terme de la charge électrique. Notre vision pour le cas de prévision à court-terme est d'utiliser un système adaptatif à deux étages. Le but dans ce modèle est d'assurer une première prévi-

sion, est d'essayer ensuite d'améliorer la précision lorsque des variables tels que les valeurs réelles de température pour le jour à prévoir seront disponibles. Pour le cas de prévision à très court-terme il n'y aura pas de temps pour tenter une amélioration de prévision. Donc, il fallait chercher la meilleure performance dès un premier temps. Par conséquent, notre idée pour ce cas de prévision est d'utiliser une combinaison de prévisions et de tirer la meilleure performance à partir d'un ensemble de méthodes primaires. En fait, nous avons opté sur l'application d'une méthode combinée, puisque à notre connaissance, la littérature de prévision de charge connaît un vrai manque de travaux qui discutent l'utilisation des méthodes combinées pour assurer des prévisions où le délai d'exécution est aussi court que quelques minutes à venir. En outre, contrairement à la plupart des travaux existants dans la littérature, où les modèles de prévision ont été mis en œuvre à travers une simulation des données de charge électrique, nous donnons beaucoup d'attention dans notre thèse au développement d'un outil convenable pour assurer la tâche de la prévision online de la consommation d'électricité. En plus, les deux modèles proposés pour VSTLF et STLF doivent assurer des prédictions satisfaisantes non seulement pour des conditions normales de la charge, mais aussi pour des conditions atypiques telles que le cas des variations rapides de la consommation d'électricité qui se produisent pendant les jours fériés. Cela doit être effectué tout en respectant un temps de calcul court et une faible intervention de la part de l'utilisateur des modèles de prévision.

1.8 Conclusion

Dans ce chapitre nous avons présenté un bref aperçu sur l'importance de la prévision de charge des systèmes électriques, les différents facteurs influençant sur la consommation d'électricité, ainsi que l'état de l'art du problème considéré. Nous avons également annoncé que l'objectif de la thèse est d'assurer des prédictions précises pour les différents horizons temporels. Cependant, au vu du rôle vitale des prévisions pour les prochaines minutes ou le jour suivant dans les opérations quotidiennes des compagnies d'électricité, nous nous donnons plus d'attention à développer des modèles adaptatifs et intelligents pour les deux cas de prévision à très court-terme et à court-terme de la charge électrique. Notre vision pour le premier cas est d'utiliser un système adaptatif à deux étages, où nous donnons dans une première étape une prévision primaire et nous essayons ensuite d'améliorer la précision lorsque des variables telles que les valeurs réelles de température pour le jour à prévoir seront disponibles. Pour le cas de prévision à très court-terme, l'idée de base est d'utiliser une combinaison de prévisions et de tirer la meilleure performance à partir d'un ensemble de méthodes primaires. Cependant, avant de mettre en pratique ces idées de base et d'entamer le développement des modèles de prévision appropriés, nous donnons dans le chapitre suivant des notions de base et des généralités liées au problème de la prévision de la charge électrique.

Généralités et notions de base

2.1 Introduction

La prévision de la consommation d'électricité est l'une des missions essentielles pour un fonctionnement économique et fiable du réseau électrique. La qualité de cette prévision contribue à garantir l'équilibre offre-demande d'électricité à tout instant ainsi que la maîtrise des flux sur le réseau. Elle a donc une influence directe sur la sûreté de fonctionnement du système électrique. Avant d'entamer le développement d'un modèle de prévision de qualité, il est tout d'abord important de comprendre le problème et savoir les différents paramètres pouvant influencer sur la puissance demandée par les consommateurs. A cet effet, ce chapitre présente des notions de base sur la consommation d'électricité, les différentes variations de la charge électrique, le pré-traitement des données de charge et les différentes méthodes et démarches à suivre pour assurer des prévisions satisfaisantes de la charge électrique.

2.2 Notions de base sur la charge électrique

La charge d'un système électrique, dans le contexte de puissance, est la somme des charges de tous les consommateurs à un instant donné. Cette consommation totale subit de grandes fluctuations : la charge varie d'une manière continue dès qu'un résident allume ou éteint la lumière, un industriel met en marche un moteur de grande puissance, une vague de froid ou de chaleur est enregistrée, ...etc. Cet échantillon de variation de charge est appelé la courbe de charge électrique. Nous pouvons avoir une courbe de charge journalière, mensuelle, ou annuelle. La Figure 2.1 inclut un exemple de courbe de charge journalière pour un jour typique en France. Cette courbe est le reflet, au quotidien, de modes de vie et d'habitudes des différents consommateurs. D'après la courbe de charge présentée dans cette figure, on peut remarquer qu'il y a peu de consommation la nuit et par conséquent peu de production. Le matin, lorsque tout le monde se réveille, les lumières s'allument un peu partout, ainsi que toutes sortes d'appareils électriques (ordinateur, téléviseur, ...etc) sont mis en service. L'activité dans les secteurs de transport et d'agriculture se met peu à peu en route et les usines démarrent intensivement leur production. Ces activités permettent donc d'atteindre une pointe de consommation électrique (la puissance demandée la plus élevée) qui se produit généralement pendant la période de soir (ou l'après-midi en été). Il faut donc augmenter la production pour répondre

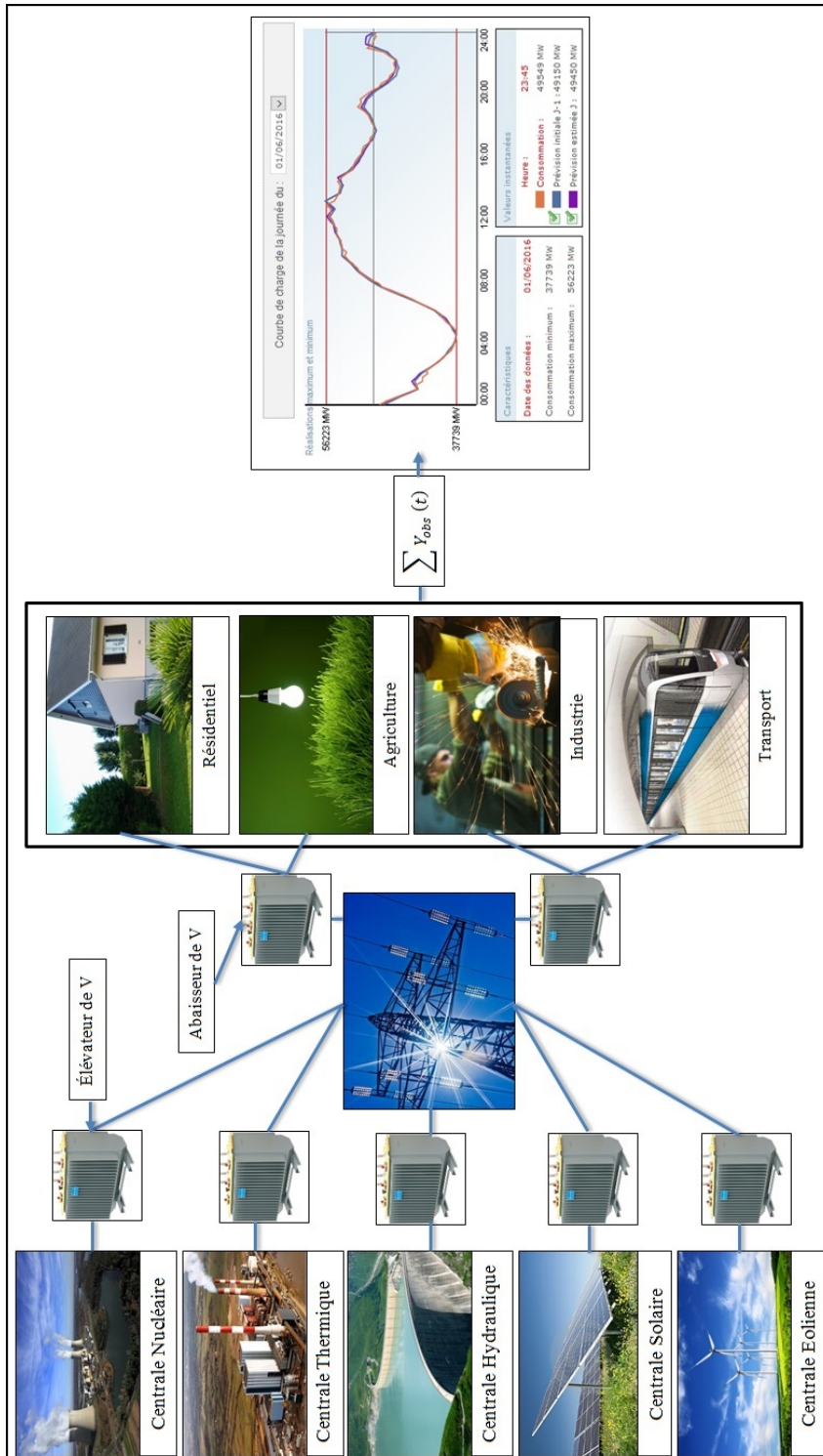


FIGURE 2.1 – Notion de la courbe de charge électrique

à cette évolution de la demande énergétique. Lorsque la nuit tombe, la charge du réseau diminue à nouveau sensiblement, ainsi que la production.

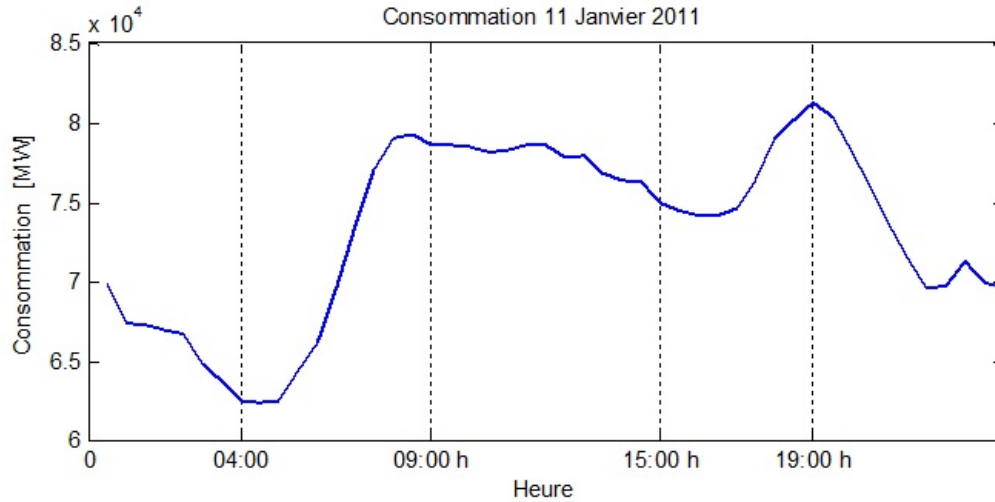


FIGURE 2.2 – Cycle journalier la consommation française d'électricité [45]

Par ailleurs, la demande globale d'électricité suit un cycle journalier, et également des cycles hebdomadaires et annuels. Ces trois cycles se distinguent notamment sur une combinaison de facteurs, dont la succession des jours de la semaine, le climat, la saison, l'activité économique et les habitudes de vie quotidienne des ménages. En France, à titre d'exemple, on peut identifier trois cycles temporels de la consommation nationale d'électricité :

- un cycle journalier, avec deux valeurs de pointe.
- un cycle hebdomadaire, avec 5 jours ouvrables globalement stables et le week-end (samedi et dimanche) où la consommation diminue (Figure 2.3 et Figure 2.4).
- un cycle annuel, avec une pointe de consommation en hiver et un creux en été (Figure 2.5 et Figure 2.6).

Le profil de la consommation sur une journée, comme il est montré par la Figure 2.2, est caractérisé par deux valeurs importantes : la pointe du matin et la pointe du soir. Le maximum de consommation est atteint à la pointe du soir vers 13h00 en été et vers 19h00 en hiver [42]. La prévision précise de la pointe quotidienne de la demande d'électricité (pic de charge électrique) est une tâche très importante pour les processus de prise de décision dans le système électrique à l'échelle d'un pays, car les conséquences d'une sous-estimation ou une surestimation peut augmenter significativement le coût d'exploitation. La surestimation de la future charge électrique peut provoquer le démarrage d'un grand nombre d'unités de production et conduit à une réserve redondante de l'énergie électrique. En revanche, la sous-estimation de la charge provoque une défaillance dans la fourniture d'une quantité suffisante de

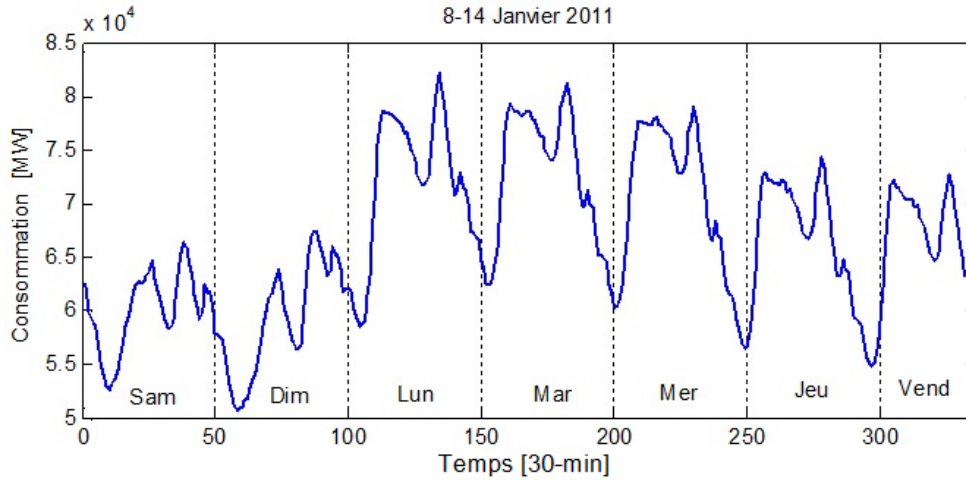


FIGURE 2.3 – Courbe de charge hebdomadaire en France : du 08 au 14 Janvier 2011 [45]

l'énergie électrique, ce qui peut conduire à l'interruption de l'approvisionnement et de coupures de courant pendant les périodes de la demande de pointe. Si l'ensemble du système électrique national devait arrêter à cause d'un black-out, ce pourrait prendre des heures, peut-être même des jours pour rétablir [44].

La Figure 2.4 présente la consommation française d'électricité pendant une semaine de Juin. Elle montre que la puissance demandée pendant le week-end est différente par rapport à la charge des jours ouvrables. Pour des saisons différentes, à partir des données de la Figure 2.5 et de la Figure 2.6, nous pouvons observer que la consommation électrique maximale se produit au cours de la saison d'hiver, mais la saison d'été connaît une diminution significative en terme de consommation d'électricité. Pour une bonne prédiction de la charge électrique, la construction d'un modèle doit prendre en compte les cycles de variations temporelles ainsi que les divers facteurs influençant sur la consommation d'électricité (tels que les fluctuations météorologiques). Cependant, ces cycles de variation temporelle ne sont pas nécessairement les mêmes pour les différents pays et régions : si la diminution de la demande d'électricité reliée au weekend est censée se produire pendant les deux jours de samedi et de dimanche en France, cette diminution est reliée au jour de vendredi en Algérie ; la consommation électrique maximale est censée se produire au cours de la saison d'hiver en France et en Australie, alors que le pic annuel de consommation d'électricité est censé être réalisé pendant la saison d'été en Algérie. Pour une bonne prédiction de la charge électrique, il faut donc faire appel à une analyse approfondie des données de charge électrique et comprendre l'évolution de la consommation d'électricité au cours du temps. Cette analyse rentre donc dans le cadre de l'étude des séries temporelles. Une discipline qui sera discutée plus en profondeur dans la section suivante.

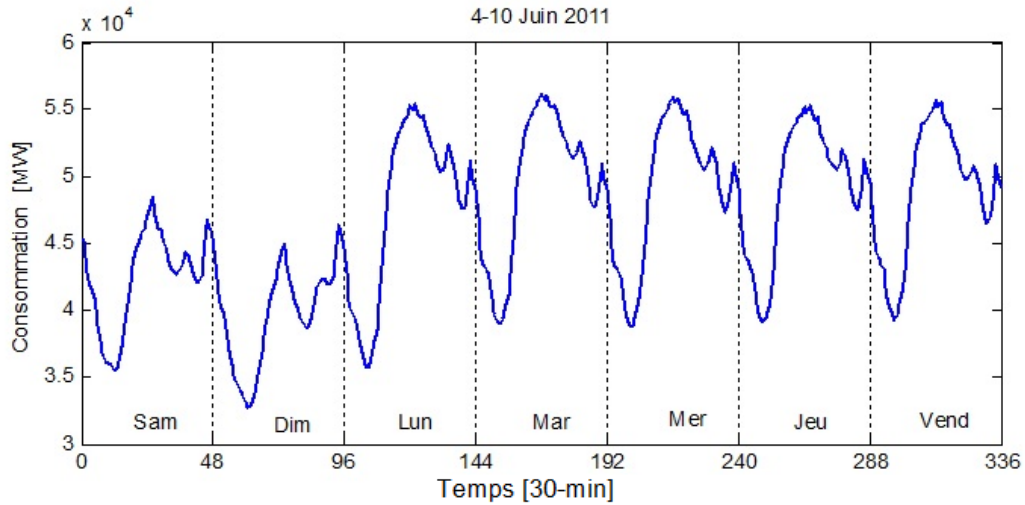


FIGURE 2.4 – Courbe de charge hebdomadaire en France : du 04 au 10 Juin 2011 [45]

2.3 Notions sur les séries temporelles

2.3.1 Définition d'une série temporelle

Une série temporelle, ou série chronologique, est une succession de valeurs numériques représentant l'évolution d'une grandeur spécifique au cours du temps. On appelle donc une série chronologique toute suite d'observations X_t , indexées par un ensemble T (espace des temps), avec $t = 1, 2, \dots, n$. Le nombre n est appelé la longueur de la série. L'indice temps peut être selon les cas l'heure, le jour, le mois, l'année ... etc. De telles suites de variables aléatoires peuvent être exprimées mathématiquement afin d'en comprendre le passé et expliquer les variations observées, de prédire les valeurs futures X_{t+h} ($h = 1, 2, 3, \dots$) à partir des valeurs passées X_1, X_2, \dots, X_t , et d'étudier le lien avec d'autres séries.

L'importance de ce domaine est illustrée par les nombreux domaines d'application, dont les buts poursuivis sont multiples (e.g. description, filtrage, modélisation) :

- Économie : estimation des grandeurs économiques
- Finance : évolution des cours d'action et les taux d'intérêt
- Démographie : réalisation des scénarios démographiques prospectifs
- Météorologie : analyse statistique des observations climatiques
- Médecine : analyse d'électrocardiogrammes
- Géophysique : analyse des enregistrements sismiques
- Théorie du signal : Traitement des signaux bruités
- Traitement d'images : analyse d'images satellites et médicales
- Énergie : prévision de la charge électrique

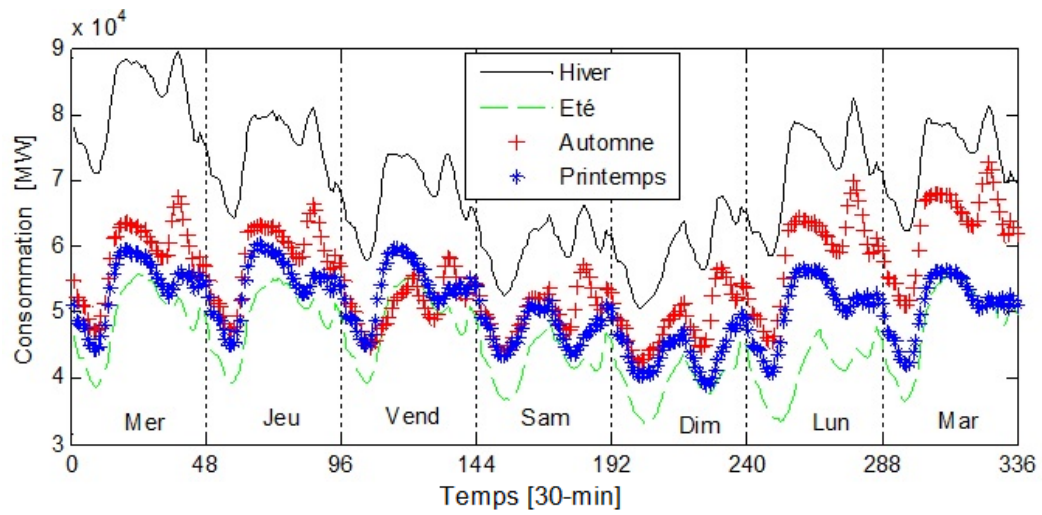


FIGURE 2.5 – Comparaison entre la consommation hebdomadaire des quatre saisons [45]

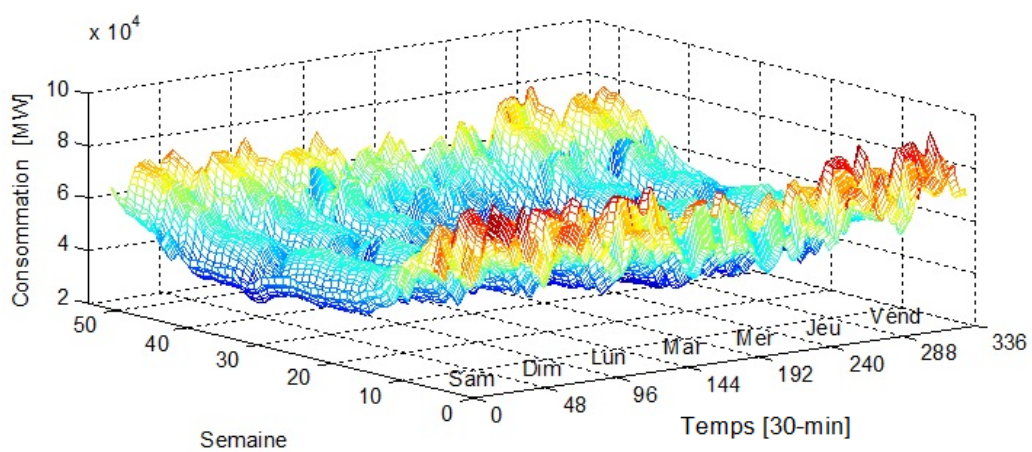


FIGURE 2.6 – Consommation d'électricité en France pendant l'année 2011 [45]

2.3.2 Composantes d'une série temporelle

Une règle générale en statistique consiste à représenter graphiquement les observations avant d'effectuer le moindre calcul. L'observation de ce graphique est souvent une aide à l'analyse de la série chronologique et permet de se faire une vision sur les différentes composantes constitutantes, ce qui permet ensuite de bien choisir la démarche ou la méthode la plus adaptée pour modéliser la série chronologique. On peut compter trois composantes principales : une composante fondamentale qui représente l'évolution générale du phénomène étudié, un mouvement saisonnier représentant les variations saisonnières, et une composante aléatoire représentant des variations accidentelles imprévisibles [46].

2.3.2.1 Tendance

La composante fondamentale (f_t) ou tendance traduit l'évolution vers le haut ou vers le bas du niveau moyen du phénomène dans le temps. La tendance générale d'une série chronologique définie par les deux extrémités X_1 et X_n est l'évolution globale de la variable en considération en passant d'un niveau f_1 lié à l'extrémité X_1 à un niveau f_n lié à l'extrémité X_n . Les tendances peuvent être linéaires ou non linéaires. Le patron d'une tendance linéaire peut être représenté par une ligne droite. Celui d'une tendance non linéaire pourra être associé par exemple à une fonction quadratique ou à une courbe exponentielle.

2.3.2.2 Composante saisonnière

La composante saisonnière ou mouvement saisonnier représente des effets périodiques de période connue s qui se reproduisent de façon plus ou moins identique d'une période sur l'autre. La chronique correspondante est notée S_t , avec $t = 1, 2, \dots, n$. Elle est généralement supposée rigoureusement périodique : $S_{t+s} = S_t$. Plusieurs composantes saisonnières peuvent éventuellement être présentes dans une même série. Ce serait par exemple le cas lors de l'étude de la demande horaire d'électricité à partir de données journalières relatives à plusieurs années d'observations. La première composante prendrait en compte la variation liée à l'heure de la journée, la seconde composante est liée au jour de la semaine, et la troisième composante prendrait en compte les variations liées aux mois de l'année.

2.3.2.3 Composante résiduelle

La composante résiduelle ou variations accidentelles est la partie non structurée du phénomène. Elle est modélisée par une suite de variables aléatoires ε_t , $t = 1, 2, \dots, n$, centrées, non corrélées et de même variance (bruit blanc).

La Figure 2.7 montre différentes séries chronologiques, avec quelques caractéristiques aussi différentes. On peut remarquer que la consommation des ménages en Allemagne et le produit intérieur brut en France semblent avoir augmenté régulièrement. Le taux de chômage en Allemagne semble avoir globalement augmenté depuis

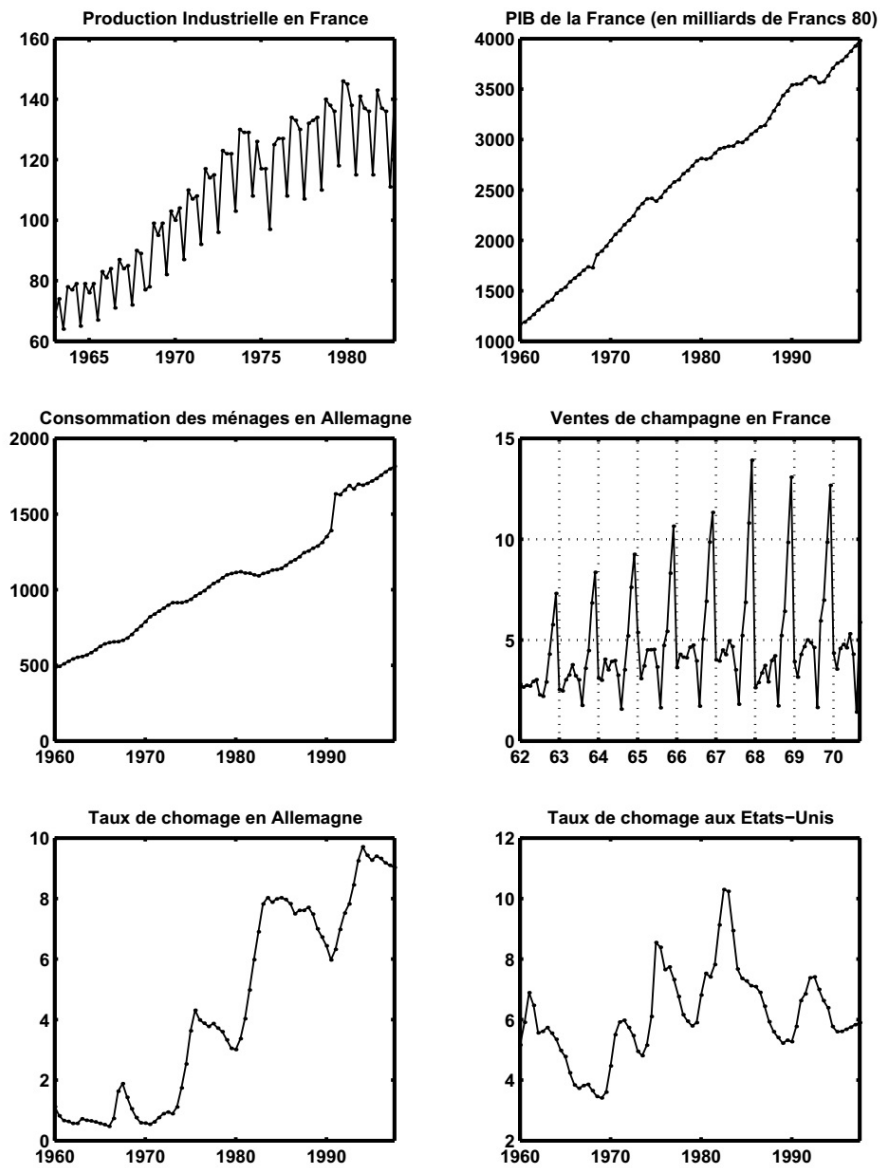


FIGURE 2.7 – Quelques exemples de séries chronologiques [47]

1960, mais avec une alternance de baisses et de hausses soudaines. Le taux de chômage des États-Unis ne semble pas évoluer globalement, mais présente également cette alternance de baisses et de hausses. Les ventes de champagnes, tout comme la production industrielle en France semblent exhiber un caractère saisonnier. D'autre part, les variations de ces deux dernières séries ont une amplitude qui semble augmenter au cours du temps. Cependant, toutes les séries présentées dans la figure ont un aspect aléatoire avec des fluctuations irrégulières qui ont parfois une amplitude anormalement élevée.

2.3.3 Modélisation d'une série temporelle

La modélisation d'une série chronologique consiste à écrire une équation précisant la façon dont les composantes s'articulent les unes par rapport aux autres pour décrire la série observée. Il existe principalement trois schémas de composition [48] :

- Schéma additif : $Y_t = f_t + S_t + \varepsilon_t$. Graphiquement, l'amplitude des variations saisonnières est constante autour de la tendance
- Schéma multiplicatif : $Y_t = f_t \times S_t (1 + \varepsilon_t)$. Graphiquement, l'amplitude des variations saisonnières varie.
- Schéma mixte : $Y_t = f_t \times S_t + \varepsilon_t$.

où f_t est la tendance ; S_t représente la composante saisonnière de différentes périodes selon qu'il s'agit de données trimestrielles, mensuelles, hebdomadaires ou journalières par exemples ; ε_t représente la variation aléatoire due à de nombreuses causes pas forcément bien identifiées, mais de répercussion limitée.

2.3.4 Méthodes pour la prévision dans une série chronologique

Les modèles utilisés pour prévoir les valeurs futures d'une série chronologique sont généralement classés en deux catégories : méthodes déterministes, et méthodes stochastiques. Le processus de modélisation de la première classe, à l'image des méthodes de lissage exponentiel, consiste à supposer que la série chronologique est parfaitement prévisible à l'aide de ces valeurs passées. Les modèles stochastiques sont de même type que les modèles déterministes puisque les valeurs passées sont aussi considérées comme des éléments utiles pour la prévision, mais possèdent également une structure aléatoire de corrélation non nulle.

2.3.4.1 Méthodes de lissage exponentiel

Les méthodes de prévision par lissage exponentiel modélisent les composantes d'une série qui est facile à observer sur un diagramme de série chronologique. Cette approche décompose les données en leurs composantes, puis étend dans le temps les estimations des composantes pour fournir des prévisions. On distingue principalement trois modèles de lissage exponentiel : lissage exponentiel simple, lissage double (on dit aussi lissage exponentiel de Holt), et la méthode de Holt-Winters.

Lissage exponentiel simple

Le lissage exponentiel simple est utilisé pour lisser les séries chronologiques discrètes. L'efficacité de cet algorithme pour ce cas peut être attribuée à sa simplicité et sa précision raisonnable. Soit une série chronologique observée. Formellement, l'équation de lissage exponentiel simple prend la forme [49] :

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{y}_t \quad (2.1)$$

Où y_t est la valeur actuelle connue de la série à l'instant t , \hat{y}_t représente la valeur prévue de la variable y pour l'instant t , \hat{y}_{t+1} est la valeur prévue pour l'instant $t + 1$, et α est la constante de lissage.

La constante de lissage α est un nombre choisi entre zéro et un, $0 \leq \alpha \leq 1$. Plus elle tend vers 1, plus la prévision sera influencée par les observations récentes. Plus α tend vers 0, plus la prévision est influencée par les observations éloignées dans le temps [49]. Pour assurer la prévision à l'instant $t + 1$, nous devons utiliser la valeur prévue de la variable y à l'instant t , ainsi que la valeur actuelle connue de la série à l'instant t . La relation générale du lissage exponentiel simple peut être écrite donc comme suit :

$$\hat{y}_{t+1} = \alpha \sum_{k=0}^{t-1} (1 - \alpha)^k y_{t-k} \quad (2.2)$$

Lissage exponentiel de Holt et de Holt-Winters

Lorsque la série présente une tendance linéaire, et n'est soumise à aucune variation saisonnière, on peut effectuer les prévisions à l'aide du lissage exponentiel de Holt. Dans le cas d'une série soumise à des variations saisonnières, on utilise souvent le modèle de Holt-Winters. Ces deux méthodes seront expliquées avec plus de détails dans le troisième chapitre.

2.3.4.2 Méthodes pour processus stochastiques

Les modèles stochastiques font aussi appel aux schémas présents dans les données. Il n'est cependant pas toujours facile d'identifier visiblement les composantes d'un diagramme de série chronologique. Pour faciliter la tâche, les processus stochastiques utilisent des fonctions de différentiation, d'autocorrélation et d'autocorrélation partielle pour aider à identifier un modèle acceptable.

Modèle autoregressif (AR)

Les modèles autoregressifs ont été introduits par Yule en 1927 [50]. On prend en compte une dépendance linéaire du processus à son propre passé :

$$AR(p) : X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t \quad (2.3)$$

Où $p \in \mathbb{N}$ est l'ordre du processus, $\varphi_1, \dots, \varphi_p$ sont des constantes réelles et $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc.

Donc, un modèle autorégressif est une régression linéaire simple de la valeur courante de la série contre une ou plusieurs valeurs antérieures de la série.

Modèle à moyenne mobile (MA)

Les modèles à moyennes mobiles ont été également introduits en 1927, par Slutsky [51]. Un processus moyenne mobile est construit à partir de l'idée que l'observation au temps t s'explique linéairement par les observations d'un bruit blanc :

$$MA(p) : X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2.4)$$

Où $q \in \mathbb{N}$ est fixe et $\theta_1, \dots, \theta_q$ sont des constantes réelles.

Modèle autorégressif à moyenne mobile (ARMA)

Développés par Box & Jenkins en 1970 [52], les modèles autorégressifs à moyennes mobiles sont une combinaison d'un modèle autorégressif et un modèle de moyennes mobiles :

$$ARMA(p, q) : X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2.5)$$

Modèles ARIMA et SARIMA

Le modèle autorégressif à moyenne mobile intégré (ARIMA) et le modèle saisonnier autorégressif à moyenne mobile intégré (SARIMA) ont été développés afin de pouvoir modéliser un grand nombre de phénomènes réels qui présentent des tendances et/ou des saisonnalités. On applique en fait des modèles ARMA à des séries dites différenciées. Par exemple, pour un ARIMA d'ordre 1, on suppose que $X_t - X_{t-1}$ est un ARMA (on note en général $\Delta X_t = X_t - X_{t-1}$) en utilisant les opérateurs identité et rétrograde $BX_t = X_{t-1}$. Donc, un processus ARIMA est un processus ARMA intégré. En outre, un processus SARIMA est un processus ARMA intégré avec une composante saisonnière. Les modèles ARIMA et SARIMA seront expliquées avec plus de détails dans le troisième chapitre.

2.4 Analyse préliminaire des données de chargé électrique

Nous avons vu dans la section précédente que l'analyse de la série chronologique du phénomène en étude, à savoir la demande de charge électrique, est une étape importante afin de bien choisir une méthode de modélisation statistique appropriée. Cependant, le prétraitement, le regroupement et la décomposition des données de

charge électrique sont des phases très importantes afin d'améliorer la performance des méthodes de l'intelligence artificielle.

2.4.1 Prétraitement des données

La qualité des données est une problématique majeure dans de nombreuses applications, y compris le domaine de prévision, car la qualité de prévision, surtout pour les méthodes d'intelligence artificielle, est définie par la qualité des données d'entrée utilisées dans le processus de traitement. Déterminer des démarches et des techniques de prétraitement des données représente une phase importante afin de bien préparer les données pour une exploitation et une extraction efficace de l'information, puisque la qualité des résultats finaux est largement conditionnée par le soin porté à cette première étape. Cette phase de prétraitement des données est généralement effectuée en faisant appel à une étude de détection de valeurs aberrantes et de valeurs manquantes afin de construire une base de données complète et filtrée.

Valeurs aberrantes

Une valeur aberrante est une valeur qui diffère de façon significative de la tendance globale des autres observations quand on observe un ensemble de données ayant des caractéristiques communes. Avant l'utilisation d'une méthode de prévision, il est important d'éliminer toute observation qui peut apparaître incompatible avec le reste des données. Ces valeurs aberrantes peuvent être dues par une raison physique connue, exemple le mauvais étalonnage de l'appareil de mesure, erreur de saisie ou de calcul, comme elles peuvent refléter une raison non connue. Pour repérer les données aberrantes, un contrôle de cohérence sur les données est généralement effectué en utilisant des algorithmes de filtrage tels que le filtrage par médiane mobile [53]. Cependant, ces filtres ne sont pas efficaces d'une manière absolue, donc l'usage du bon sens et de l'expérience est le plus sûr pour décider si une valeur extrême est une valeur aberrante et d'essayer ensuite de la filtrer.

La Figure 2.8 présente un exemple de valeurs aberrantes qui peuvent être rencontrées au moment du prétraitement des données de charge électrique. Des données de charge, très élevées ou très faibles, en tout cas très différents de la grande majorité des valeurs, doivent conduire à s'interroger sur la validité de ces données, c'est pour cette raison que la réflexion sur la cohérence des données est essentielle.

Données manquantes

Les données manquantes sont des cellules apparaissent comme vierges (ne contiennent aucune valeur) dans la base de données d'une variable. Le non considération de ces valeurs manquantes peut aboutir à des résultats erronés et entraîner une perte de performance dans les approches de prévision car l'information disponible est inférieure à celle initialement prévue. Un autre problème est que les hypothèses effectuées en aval de nombreuses procédures statistiques sont basées sur des observations complètes et que les valeurs manquantes peuvent compliquer la théorie requise.

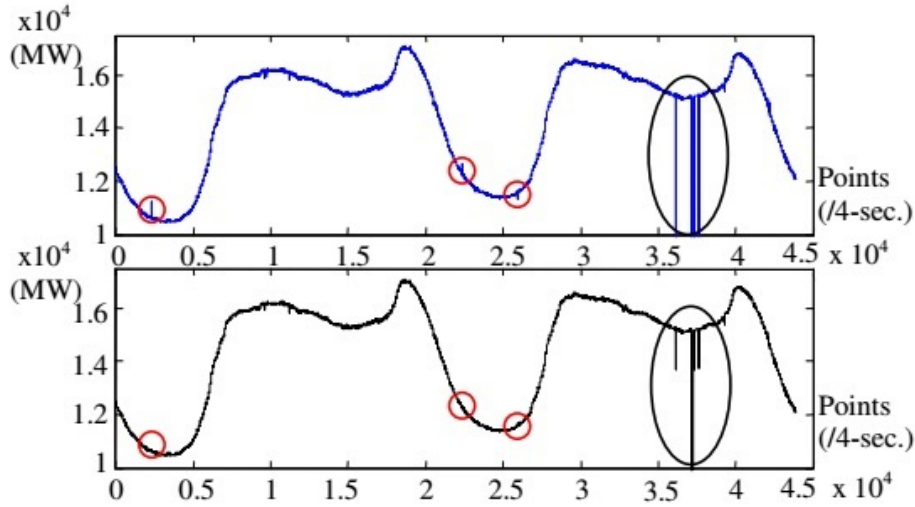


FIGURE 2.8 – Avant (en haut) et après (en bas) le filtrage de valeurs aberrantes basée sur deux jours de données de charge britannique avec une résolution de 4 secondes [54]

Il est important donc de remplacer toutes les données manquantes d'une variable afin d'améliorer la précision et d'aborder de nombreux problèmes. Ce remplacement est généralement effectué en remplaçant chacune des valeurs manquantes par la valeur moyenne de toute la série, la valeur prévue obtenue par régression, ou encore par le biais de l'utilisation des valeurs des plus proches voisins.

Le Tableau 2.1 présente une partie d'une base de données utilisée pour prévoir les pics de consommation d'électricité en Algérie. Les valeurs de charge pour le 07 Décembre 2014, montrées en gras dans le tableau, sont reliées en fait à des données manquantes, dont les valeurs ont été remplacées par les valeurs correspondant pour le même type de jour de la semaine qui précède.

2.4.2 Classement et classification des données de charge électrique

Le classement et la classification sont des disciplines reliées de près ou de loin à plusieurs domaines, dont le but est de comprendre et d'organiser des objets qu'on souhaite traiter en catégorie. On peut grouper les méthodes classificatoires en deux grandes familles : supervisée (classement) et non supervisée (classification). Dans un classement (classification en anglais) on affecte les objets à des groupes préétablis, où un nouvel individu est rangé dans une classe en fonction de ses caractéristiques, utilisées comme variables explicatives. La classification (clustering en anglais) est, en quelque sorte, le travail préliminaire au classement, savoir la recherche des classes "naturelles" dans le domaine étudié. Les groupes sont constitués donc au cours de l'analyse où les individus ayant des caractéristiques similaires sont regroupés.

Tableau 2.1 – Valeurs manquantes dans une base de données utilisée pour prévoir le pic journalier de la charge électrique en Algérie

Jour	Pointe du matin	Pointe du Soir	Température maximale	Température minimale
	[MW]	[MW]	[°C]	[°C]
30/11/2014	6821	8011	19	14
01/12/2014	7022	8261	17	12
02/12/2014	7051	8282	16	12
03/12/2014	7053	8342	20	10
04/12/2014	6996	8238	19	09
05/12/2014	6283	8007	17	10
06/12/2014	6739	8492	17	08
07/12/2014	6821	8011	14	08

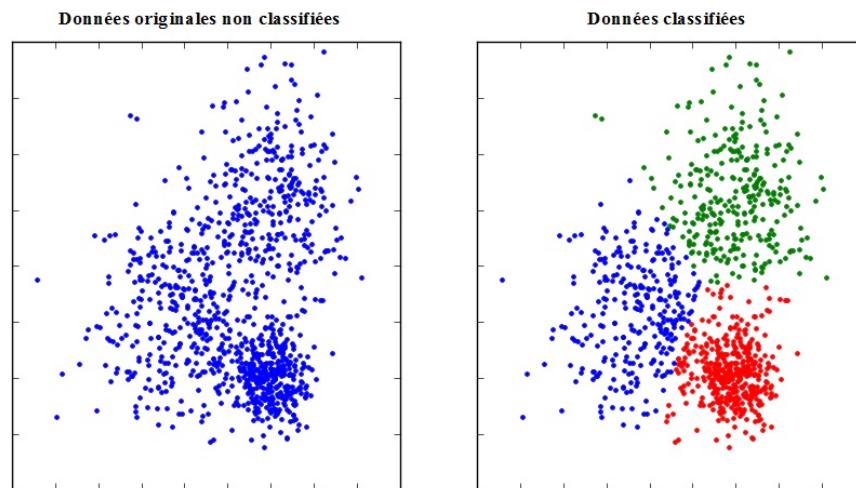


FIGURE 2.9 – Classification des données

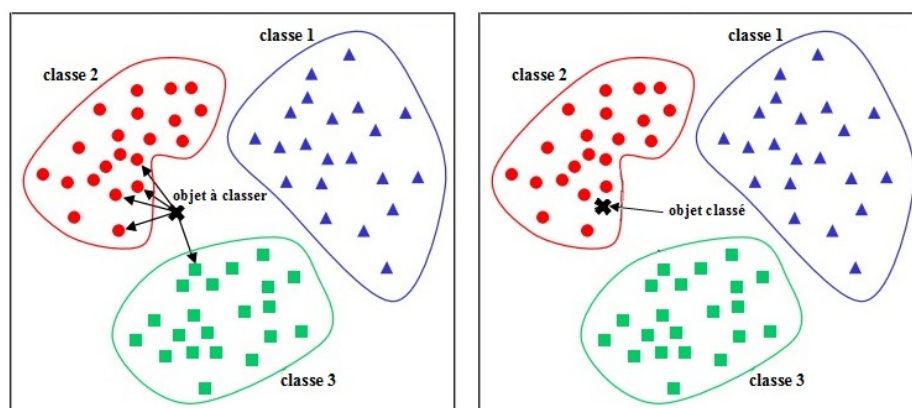


FIGURE 2.10 – Classement des données

Les deux notions de classification et de classement sont indiquées respectivement, par la Figure 2.9 et la Figure 2.10. Ces opérations ont pour but de diminuer la dimension de l'espace d'entrée ou encore de séparer les éléments d'une population d'effectif important, afin d'organiser les objets en groupes (clusters) dont les membres ont en commun diverses propriétés.

Pour le domaine de la prévision de charge, des techniques d'exploration et de classification de données ont été largement utilisées pour identifier des données de charge avec des comportements similaires de la demande dans une base de données historiques. Pour la prévision à court-terme, les données de charge électrique peuvent être segmentées selon la saison, le type de jour, et l'heure de la journée [55].

La classification des données de charge selon les saisons (été, automne, hiver, printemps) a l'avantage de réduire les données à utiliser. Une analyse plus poussée des données de charge peut servir également à identifier le type de jour et à réduire donc la quantité d'information qu'un modèle doit incorporer. De telles approches ont été mises en œuvre avec succès par plusieurs auteurs [55], [56], [57], [58].

La classification selon la saison peut être déterminée par différentes approches, citons deux types :

- algorithme des k-moyennes (k-means)
- regroupement en c-moyennes floues (FCM)

La sélection du type de jour peut être aussi déterminée par différentes approches, citons deux types :

- algorithme des k-plus proches voisins
- cartes auto-organisées de Kohonen

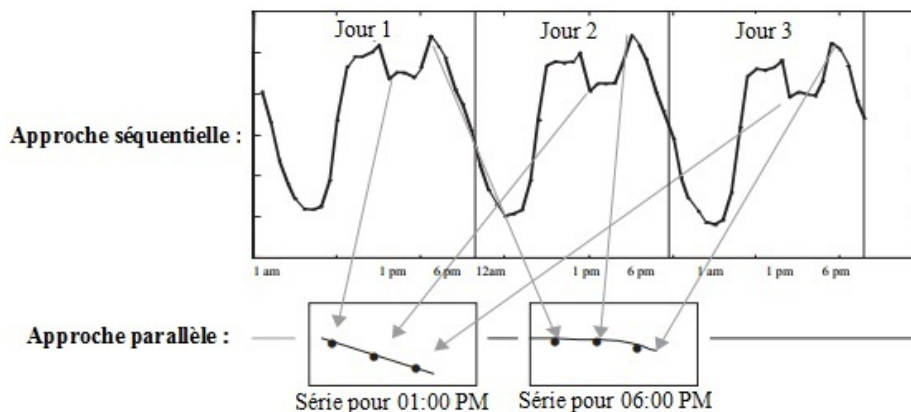


FIGURE 2.11 – Une approche parallèle pour classifier les données de la charge selon l'heure de la journée [59]

La sélection par heure du jour est une autre méthode pour classifier les données de la charge électrique. En comparaison avec les modèles séquentiels qui consistent à modéliser la charge horaire comme une série unique, il est parfois utile de modéliser chaque heure de la journée par une série séparée. Cette modélisation séparée est

connue sous le nom de modélisation parallèle, dont un exemple est montré par la Figure 2.11. L'avantage de cette approche apparaît surtout dans le cas de la prévision à très court-terme de la charge électrique, puisque l'approche parallèle permet de diminuer la quantité des données à utiliser, améliorer la qualité des données et diminuer le temps d'exécution de la méthode de prévision.

2.4.3 Décomposition des signaux de charge

Nous avons vu dans le cas de l'analyse et modélisation des séries chronologiques que l'évolution d'une grandeur spécifique dépend seulement d'un seul variable qui est le temps. Mais cette démarche d'analyse n'est pas toujours suffisante, car des informations utiles peuvent être cachées dans la partie fréquentiel du signal de la grandeur en examen. Il existe plusieurs démarches pour présenter le comportement fréquentiel d'un signal, mais la méthode la plus populairement utiliser pour décomposer les séries de charge est sans aucun doute, la décomposition par la transformée en ondelettes [60]. A cet égard, nous donnons dans les paragraphes suivants des notions de base sur l'analyse par ondelette et les travaux de base qui ont abouti à cette technique.



FIGURE 2.12 – Transformée de Fourier d'un signal de série temporelle

Transformée de Fourier

La transformée de Fourier (FT) est sans doute la transformation mathématique de départ pour extraire des informations utiles cachées dans un signal. Comme le montre la Figure 2.12, l'analyse de Fourier est utilisée pour transformer un signal de son domaine temporel à son domaine de fréquence. Cette transformation est importante car elle fournit le contenu fréquentiel d'un signal. Néanmoins, la FT a un inconvénient majeur qui est le problème de la localisation. Les informations de temps sont perdues lorsque le signal de série temporelle est transformé dans le domaine fréquentiel. Cela signifie que nous ne pouvons plus connaître quand un événement particulier survient. Ainsi FT est plus adaptée aux signaux stationnaires où le temps n'a pas de préoccupation avec le signal désiré [45].



FIGURE 2.13 – Transformée de Fourier à court-terme d’un signal de série temporelle

Transformée de Fourier à court-terme

Pour corriger la déficience de la FT, la transformée de Fourier à court-terme (STFT) est une technique qui permet de mettre en œuvre une technique de fenêtrage supplémentaire pour analyser seulement une petite partie du signal dans un délai particulier. La Figure 2.13 illustre la STFT d’un signal de série temporelle. D’après cette figure, on peut observer que le signal transformé est en fonction à la fois du temps et de la fréquence. Le signal contenu dans la fenêtre donnée est également segmenté et les composantes de fréquence sont extraites à partir de chacun de ces segments du signal.

Transformée en ondelettes

Bien que la technique de STFT soit capable de fournir quelques informations de temps dans le signal transformé, elle a une limitation d’une taille de fenêtre fixe pour toutes les fréquences. L’information est obtenue seulement avec une précision limitée déterminée par la taille de fenêtre, cela n’est pas souhaitable car il cause des problèmes liés aux résolutions temps-fréquence. Pour résoudre ce problème, une taille de fenêtre variable est requise. Ceci conduit à la mise en œuvre de la transformée en ondelettes.

En 1989, Mallat S.G a proposé un procédé révolutionnaire [61], l’analyse multi-résolution par ondelettes, qui permet de décomposer un signal à plusieurs échelles (résolutions) et de le reconstruire à partir des éléments de cette décomposition. La décomposition du signal en ondelettes est capable de fournir les informations de temps et de fréquence simultanément et donc une représentation temps-fréquence du signal plus riches [62]. Le fait que l’analyse par ondelettes permet une meilleure représentation dans le plan temps-fréquence lui donne beaucoup d’avantages :

- Pour des signaux physiques présentant des variations très rapides et des discontinuités, l’analyse en ondelettes est adaptée car l’ondelette va détecter ces variations et les analyser.
- La localisation en temps est précieuse pour beaucoup d’applications.
- La transformée en ondelette peut représenter complètement et efficacement un signal quelconque avec peu de coefficients.

En raison de la grande utilité de l'analyse par ondelette, nous avons opté d'utiliser cette méthode de décomposition de signaux dans notre modèle de prédiction à court-terme de la charge électrique. De ce fait, nous revenons plus tard sur une description plus détaillée sur cette méthode de décomposition.

2.5 Prédiction de la charge électrique

Après une analyse de la série chronologique de la charge électrique, une étape de classification et de classement de données historiques et une possible décomposition de signaux de la demande d'électricité, il vient finalement le moment pour choisir le modèle (on utilise aussi le terme technique, méthode, approche, ou méthodologie) de prédiction et de tenter d'assurer les prévisions satisfaisantes de la consommation d'électricité. Le terme modèle est ici pris au sens très large. Ce modèle pourra être une régression, un lissage exponentiel, une méthode de l'intelligence artificielle, ou n'importe quelle technique prévisionnelle.

Les performances de tout modèle de prédiction dépendent fortement de la présence d'historiques suffisamment fiables et importants, de l'expérience de l'utilisateur, et de l'horizon de prédiction. Le choix et l'ajustement de modèle de prédiction représente en effet une étape potentielle qui doit être déterminée après une analyse préliminaire des différentes informations et jugements de la part du prévisionniste. Par exemple, il est conseillé d'utiliser une technique de lissage exponentiel double lorsque les données historiques contiennent une tendance. En outre, si les données comprennent un terme de saisonnalité, on peut utiliser la méthode de Holt-Winters (HW) ou encore un modèle saisonnier de SARIMA. Lorsque des facteurs exogènes sont disponibles, on peut utiliser des techniques de l'intelligence artificielle pour bien modéliser la relation existante entre ces facteurs explicatifs et la charge électrique.

Néanmoins, quel que soit la méthode sélectionnée, il est totalement impossible que cette méthode permettra le meilleur résultat pour tout le temps : certaines méthodes performant bien dans certaines situations et n'arrivent pas à suivre les variations de la charge électrique dans d'autres cas. Il est conseillé donc d'utiliser une combinaison de plusieurs techniques pour diminuer l'incertitude et améliorer la précision de prédiction. En effet, la tendance actuelle dans la recherche en prédiction de la charge électrique se focalise sur l'utilisation des « méthodologies » hybrides et combinées permettant de réaliser des systèmes adaptatifs capables d'apprendre des relations complexes en milieu incertain. La Figure 2.14 résume les différentes démarches ou méthodes qui peuvent être utilisés pour assurer la prédiction de la charge électrique.

2.6 Mesures des erreurs de prédiction

Il est rare que l'on réussisse à prédire exactement la demande de charge électrique. L'erreur de prédiction est la différence entre la demande prévue et la demande réelle. On peut s'attendre à ce qu'un modèle de prédiction génère des erreurs, mais

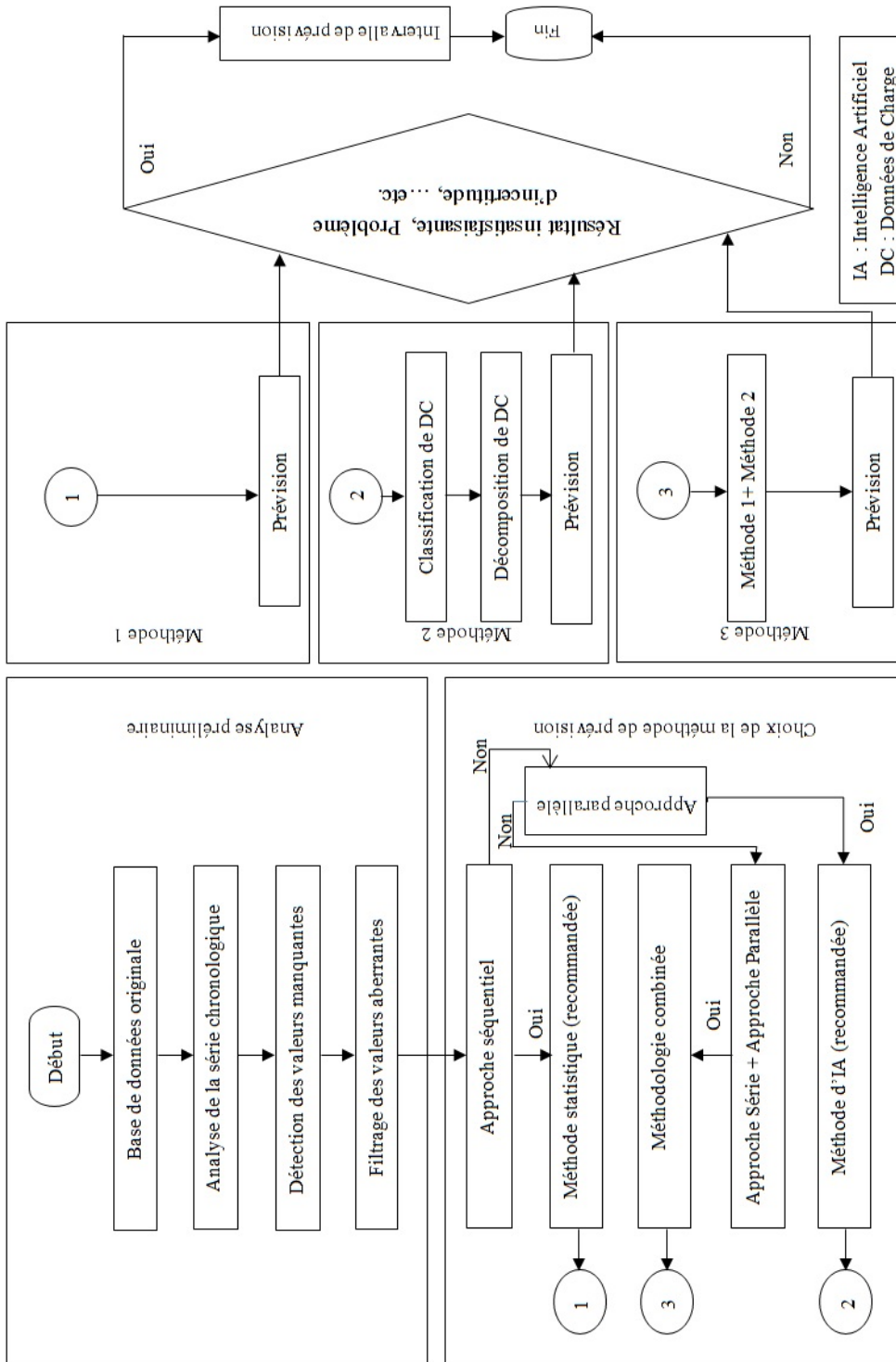


FIGURE 2.14 – Différentes méthodes pour assurer les prévisions de la charge électrique

un modèle fera une sous-estimation de la demande aussi souvent que des surestimations. Le calcul de l'erreur moyenne faite par un modèle prévisionnel nous donne une mesure de l'exactitude avec laquelle ce modèle adhère au patron des données historiques. Il existe plusieurs moyens pour mesurer la qualité des prévisions, mais les plus populairement utilisés sont les cinq mesures suivantes :

- APE : Erreur absolue en pourcentage
- MAD : Déviation absolue moyenne
- MSE : Erreur quadratique moyenne
- RMSE : Racine carrée de l'erreur quadratique moyenne
- MAPE : Erreur absolue moyenne en pourcentage

2.6.1 Erreur absolue en pourcentage

Calculer un pourcentage d'erreur nous permet de comparer une valeur réelle à une valeur théorique. L'erreur absolue en pourcentage est comme son nom l'indique s'agit de la différence entre la valeur prévue et la valeur réelle exprimée comme un pourcentage de la valeur réelle (par exemple, la valeur prévue est 10% plus loin de la valeur réelle). Ce pourcentage permet donc de savoir à quel point la valeur prévue était proche (ou éloignée) de la valeur réelle. Tout ce dont on est besoin pour apprendre à calculer un pourcentage d'erreur est une valeur prévue et une valeur réelle.

$$APE = \frac{|\hat{y}_t - y_t|}{y_t} \times 100 \quad (2.6)$$

Où y_t représente la demande réelle à la période t , et \hat{y}_t la prévision de la demande faite pour la période t .

2.6.2 Déviation absolue moyenne

La déviation absolue moyenne (MAD) est la moyenne des erreurs faites par le modèle de prévision sur une période de temps. L'équation suivante montre comment est calculé la MAD :

$$MAD = \frac{\sum_{t=1}^n |\hat{y}_t - y_t|}{n} \quad (2.7)$$

Où y_t représente la demande réelle à la période t , \hat{y}_t la prévision de la demande faite pour la période t , et n est le nombre total de cas de prévision.

2.6.3 Erreur quadratique moyenne

Avoir des petites erreurs, en-dessous et au-dessus des valeurs réelles de la charge électrique et qui s'annulent les unes les autres, est probablement le mieux que l'on puisse souhaiter. L'effet de petites erreurs de prédiction n'est ordinairement pas très grave. En revanche, les grandes erreurs de prévision peuvent être difficiles à surmonter. En conséquence, une mesure permettant de pénaliser les grandes erreurs

plus que les petites peut être désirable. L'une des mesures pour pénaliser les grandes erreurs de prévision est le calcul de l'erreur quadratique moyenne. Le MSE multiplie chaque erreur par elle-même (le carré de l'erreur), donnant ainsi un poids plus grand aux grandes erreurs qu'aux petites erreurs.

$$MSE = \frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n} \quad (2.8)$$

Où y_t représente la demande réelle à la période t , \hat{y}_t la prévision de la demande faite pour la période t , et n est le nombre total de cas de prévision.

2.6.4 Racine carrée de l'erreur quadratique moyenne

Une autre mesure pour calculer les erreurs de prévision est donnée par la racine carrée de l'erreur quadratique moyenne :

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2} \quad (2.9)$$

Cette mesure pénalise également les erreurs importantes, et elle est d'une grande utilité surtout pour comparer plusieurs estimateurs. Si on veut comparer deux estimateurs, l'estimateur le plus efficace est simplement celui qui a la valeur d'erreur la plus petite.

2.6.5 Erreur absolue moyenne en pourcentage

En plus du calcul de la racine carrée de l'erreur quadratique moyenne, il est aussi important de se faire une idée du modèle en utilisant une erreur relative. L'erreur relative faite par un modèle de prévision est mesurée par l'erreur absolue moyenne en pourcentage. L'équation suivante montre comment la MAPE est calculée :

$$MAPE = \left(\frac{100}{n} \right) \sum_{t=1}^n \frac{|\hat{y}_t - y_t|}{y_t} \quad (2.10)$$

2.7 Intervalle de prévision

Au lieu de donner une valeur unique à la charge à prévoir, il peut être intéressant de donner un intervalle en indiquant la probabilité que la valeur de la charge électrique soit comprise. L'intervalle de prévision fournit l'information sur les prévisions individuelles de la variable dépendante (consommation d'électricité). C'est-à-dire qu'un intervalle de prévision pour une valeur prévue de la variable dépendante nous donne une étendue de valeurs autour desquelles une observation supplémentaire de la variable dépendante peut être attendue (avec un niveau de certitude donné).

L'utilisation d'un intervalle de prévision est très utile pour les deux cas de prévision à moyen et à long-terme de la charge électrique, dans la mesure où une grande

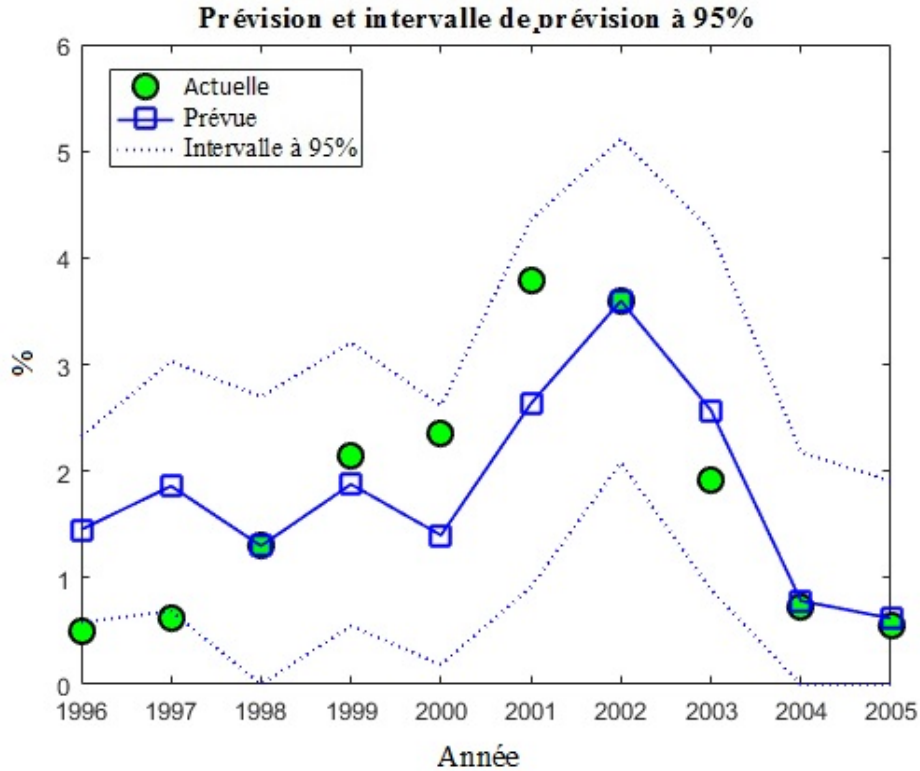


FIGURE 2.15 – Prévision et intervalle de prévision des taux de défaillance des entreprises [63]

incertitude est liée à ces deux horizons temporels. L'utilisation d'un intervalle de prévision est également importante dans des disciplines où l'incertitude représente un facteur important à envisager. La Figure 2.15 représente une courbe traduisant l'évolution des défaillances d'entreprises, tracées avec les deux intervalles de prévision de 95%. Ces deux intervalles peuvent être calculés en faisant appel à deux éléments clés : (1) un indicateur de performance, et (2) les valeurs prévues de la variable en étude :

$$\text{Partie supérieure} = \hat{y}_t + 1.96\sqrt{(e)^2} \quad (2.11)$$

$$\text{Partie inférieure} = \hat{y}_t - 1.96\sqrt{(e)^2} \quad (2.12)$$

Où y_t représente la valeur réelle de la variable dépendante (défaillances d'entreprises dans l'exemple de la Figure 2.15) à la période t ; \hat{y}_t représente la valeur prévue de la variable dépendante pour la période t ; e est un indicateur de l'erreur de prédiction, qui peut être calculé à partir de la performance antérieure du modèle de prévision.

2.8 Conclusion

Dans ce chapitre, une analyse de la variation de la charge électrique et une description sur les différentes méthodes à suivre pour assurer les prévisions satisfaisantes ont été effectuées. Nous avons vu que les performances de tout modèle de prévision dépendent fortement de la présence d'historiques suffisamment fiables et importants, de l'expérience de l'utilisateur et de l'horizon de prévision. Le choix et l'ajustement de modèle de prévision représente en effet une étape potentielle qui doit être déterminée après une analyse préliminaire des différentes informations et jugements de la part du prévisionniste. Pour avoir une idée sur les performances des différentes méthodes de prévision, nous examinerons dans le chapitre suivant une évaluation pratique de plusieurs modèles, dont le but est d'assurer les prévisions à moyen et à long-terme de la charge électrique.

Évaluation des méthodes de prévision à moyen et à long-terme de la charge électrique

3.1 Introduction

Les prévisions précises à moyen et à long-terme de la charge électrique ont un effet significatif pour prendre des décisions sur l'expansion du réseau électrique et la planification de future capacité de production. Ce chapitre utilise les observations de la demande d'électricité en Algérie pour évaluer des méthodes de prévision à moyen et à long-terme de la charge électrique. Nous considérons des méthodes conçues pour capturer la tendance et la saisonnalité dans les données de consommation d'électricité, ainsi que des techniques d'intelligence computationnelle. En vue de la grande incertitude reliée à ce genre de prévisions, nous voulons aussi de prévoir un intervalle dans lequel les futures valeurs de charge peuvent être comprises.

3.2 Données de charge utilisées

Dans cette étude, nous considérons deux séries de charge électrique : l'une est pour la prévision à moyen-terme, et l'autre est utilisée pour la prévision à long-terme. Ces données ont été collectées à partir du site web de l'opérateur du système électrique (OSE) algérien [64]. La première série, comme le montre la Figure 3.1 se compose de cinq années d'observations mensuelles pour la demande d'électricité en Algérie : de Janvier 2010 à Décembre 2014. Cette série montre un cycle saisonnier annuel d'une durée de douze périodes. Cependant, la consommation d'électricité atteint ses plus hauts niveaux en hiver, quand tous les habitants sont tous ensemble à la maison et utilisent en même temps, l'éclairage, appareils de chauffage électrique et autres appareils ménagers (téléviseurs, ordinateurs, ... , etc). En été, les pics sont causés principalement par l'augmentation d'utilisation des appareils de climatisation ces dernières années. La deuxième série en considération enregistre les pics de consommation annuels en Algérie le long de quatorze ans : de 2001 à 2014. Comme le montre la Figure 3.2, la consommation d'électricité au niveau national augmente année après année en raison de l'évolution démographique et économique du pays.

Le pic national annuel a été multiplié par environ 2.28 entre 2001 et 2014 et augmente aussi de 4791 MW à 10927 MW, ce qui représente un taux de croissance annuel moyen égal à 6.59%.

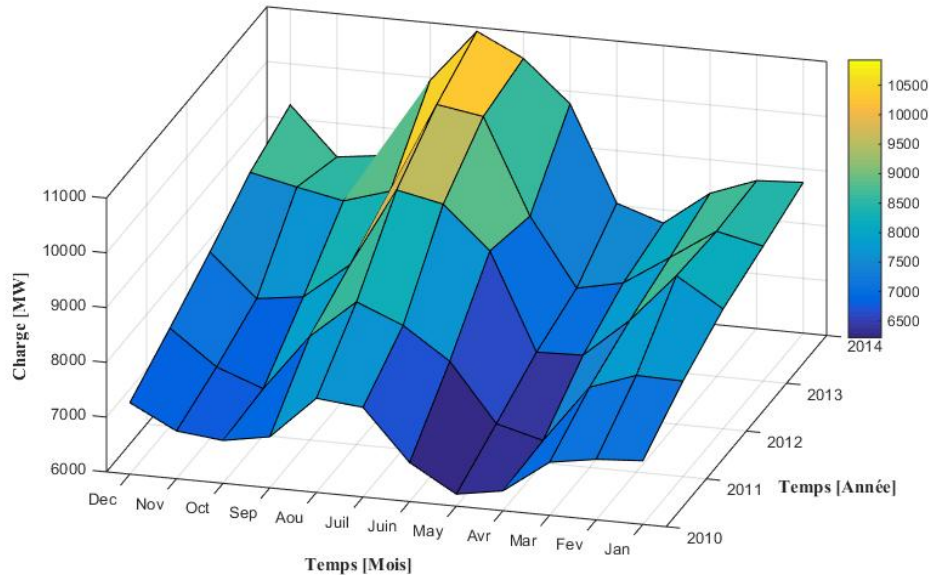


FIGURE 3.1 – Données mensuelles de la charge de pointe en Algérie : de janvier 2010 à décembre 2014

3.3 Méthodes proposées pour la prévision des pics de charge

Les modèles proposés dans cette étude sont conçus pour assurer des prévisions pour un horizon temporel qui peut être de douze mois à venir dans le cas de prévision à moyen-terme, et qui peut atteindre plusieurs années à venir pour les méthodes de prévision à long-terme de la charge électrique. Pour le cas de la prévision à moyen-terme, nous avons opté à utiliser le lissage exponentiel de Holt-Winters (HW) et le processus saisonnier autorégressif à moyenne mobile intégré (SARIMA), en raison de la grande capacité de ces deux approches pour modéliser la saisonnalité contenue dans la série chronologique de la demande de charge. Pour le cas de la prévision à long-terme, nous avons opté à utiliser le lissage exponentiel de Holt (HES) et le processus autorégressif à moyenne mobile intégré (ARIMA), en tant que techniques efficaces pour modéliser la tendance et l'évolution globale des données historiques de la consommation d'électricité. Cependant, nous avons opté à utiliser les modèles de perceptron multicouche à rétro-propagation de l'erreur (BPNN) et le système d'inférence neuro-flou adaptatif (ANFIS) pour à la fois la prévision à moyen et à long-terme de la charge électrique.

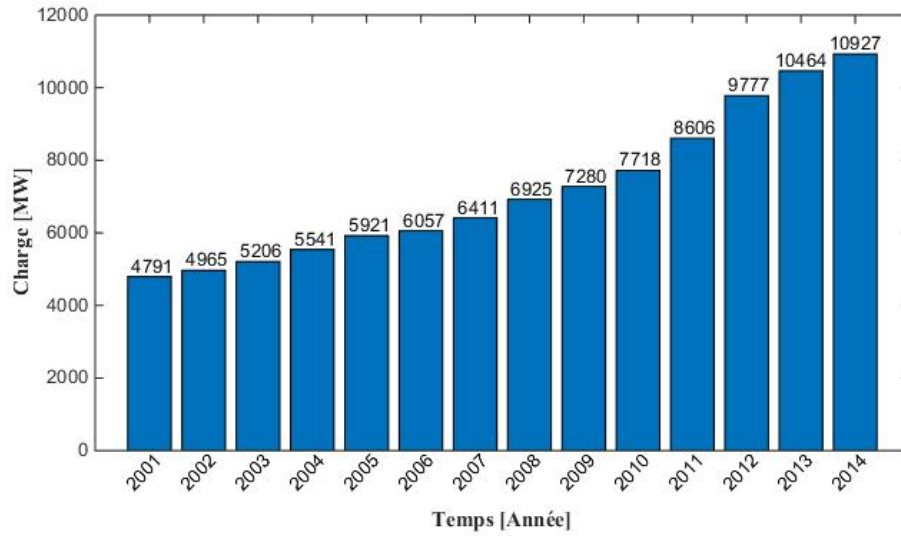


FIGURE 3.2 – Evolution de la charge de pointe annuelle en Algérie : 2001-2014

3.3.1 Lissage exponentiel de Holt

Les approches de lissage exponentiel sont des techniques relativement simples mais robustes qui ont été largement utilisées pour plusieurs applications commerciales tel que le contrôle des stocks [65]. Principalement, il existe trois techniques de lissage exponentiel nommées : lissage exponentiel simple [66], lissage exponentiel de Holt [67], et la méthode de Holt-Winters [68]. La méthode de lissage exponentiel simple, nécessite peu de calcul, et elle est utilisée lorsque les données historiques ne disposent pas d'une variation périodique ni de tendance [69]. La méthode de Holt, également connue comme méthode de lissage exponentiel double, est utilisée pour des séries chronologiques qui contiennent une tendance. Pour les séries chronologiques avec des variations saisonnières, le lissage exponentiel de Holt-Winters est utile car il permet de capturer à la fois la tendance et la saisonnalité dans les données historiques. La formulation pour le lissage exponentiel de Holt est donnée dans les expressions suivantes :

$$S_t = \alpha y_t + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (3.1)$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)(T_{t-1}) \quad (3.2)$$

$$\hat{y}_t(k) = S_t + kT_t \quad (3.3)$$

S_t et T_t représentent respectivement le niveau et la pente à l'instant t ; α et γ sont les paramètres de lissage; y_t est la valeur réelle de la série chronologique à la période t ; et $\hat{y}_t(k)$ est la prévision k -étape d'avance fait à partir de l'origine de prévision t .

Cependant, les paramètres de toute technique de lissage exponentiel ont une grande influence sur la précision des prévisions. Afin d'améliorer la performance, l'optimisation des paramètres de la méthode de Holt a considérablement attiré notre attention. Dans ce chapitre, les paramètres de la méthode de HES sont estimés en minimisant l'erreur de prévision de la pointe annuelle de la demande en Algérie sur la période entre 2004 et 2009. Nous avons généré dans une première étape 2000 vecteurs de paramètres à partir d'un générateur de nombres aléatoires entre 0 et 1. Pour chacun des vecteurs, nous avons ensuite évalué les prévisions sur la période entre 2004 et 2009. Le seul produisant la plus faible valeur d'erreur absolue moyenne en pourcentage a été choisi comme vecteur de paramètre final. Les valeurs optimisées des paramètres de la méthode sont les suivants : $\alpha = 0,701$, et $\gamma = 0,971$.

3.3.2 Modèle autorégressif à moyenne mobile intégré

La classe des modèles ARIMA, popularisée et formalisée par Box et Jenkins [41], est l'une des approches les plus populairement utilisées pour la prévision et la modélisation des séries chronologiques. Dans un modèle ARIMA, la valeur future d'une variable est censée être une combinaison linéaire des valeurs passées et des erreurs passées, exprimée comme suit :

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \dots \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3.4)$$

Le modèle est généralement considéré comme un modèle $ARIMA(p, d, q)$ où les paramètres p , d , et q sont des nombres entiers non négatifs qui se réfèrent respectivement à l'ordre du polynôme autorégressif (AR), d'intégration (I), et du polynôme de moyenne mobile (MA) du modèle. Un modèle $ARIMA(p, d, q)$ peut être représenté comme suit :

$$\phi_p(B) \nabla^d y_t = C_0 + \theta_q(B) \varepsilon_t \quad (3.5)$$

Où y_t est la valeur réelle à l'instant, ε_t est l'erreur à l'instant t , ϕ et θ sont des coefficients, B est un opérateur de retard et C_0 est une constante non nulle. Le modèle ARIMA adopté dans ce chapitre peut être décrit comme $ARIMA(3, 1, 1)$.

3.3.3 Lissage exponentiel de Holt-Winters

Pour les séries chronologiques avec des variations saisonnières, la méthode de Holt-Winters est très utile car elle permet de capturer à la fois la tendance et la saisonnalité dans les données historiques. La formulation de HW est donnée dans les expressions suivantes :

$$S_t = \alpha(y_t/D_{t-s}) + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (3.6)$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)(T_{t-1}) \quad (3.7)$$

$$D_t = \delta(y_t/S_t) + (1 - \delta)D_{t-s} \quad (3.8)$$

$$\hat{y}_t(k) = (S_t + kT_t)D_{t-s+k} \quad (3.9)$$

S_t et T_t , sont le niveau et la pente à l'instant t ; D_t est l'indice saisonnier pour le cycle annuel s d'une durée de douze périodes; α , γ , et δ sont les paramètres de lissage; y_t est la valeur réelle de la série temporelle pour la période t ; et $\hat{y}_t(k)$ est la prévision k -étape à l'avance fait de l'origine de prévision t .

Dans ce chapitre, les paramètres de la méthode de HW sont estimés en minimisant la valeur de MAPE pour une seule étape de prévision à l'avance. Nous avons généré dans une première étape 2000 vecteurs de paramètres à partir d'un générateur de nombres aléatoires entre 0 et 1. Pour chacun des vecteurs, nous avons ensuite évalué la MAPE pour la prévision de la demande de pointe mensuelle sur la période comprise entre 2012 et 2013. Le seul produisant la valeur de MAPE la plus basse a été choisi comme vecteur de paramètre final. Les valeurs optimisées des paramètres de la méthode sont les suivants : $\alpha = 0,093$; $\gamma = 0,253$; et $\delta = 0,748$.

3.3.4 Modèle saisonnier autorégressif à moyenne mobile intégré

Le processus saisonnier autorégressif à moyenne mobile intégré est une extension du processus autorégressif à moyenne mobile intégré qui a été mis en place pour modéliser des séries chronologiques avec tendances, comportement saisonnier et d'autres caractéristiques non stationnaires. Dans le cas de MTLF, la composante de la saisonnalité vient du cycle saisonnier annuel S d'une durée de douze périodes (douze mois). Un SARIMA, noté SARIMA $(p, d, q) \times (P, D, Q)_S$, peut être déclaré sous la forme suivante :

$$\varphi_p(B) \phi_P(B^S) (1 - B)^d (1 - B^S)^D y_t = \theta_q(B) \Theta_Q(B^S) \varepsilon_t \quad (3.10)$$

Où y_t et ε_t sont respectivement la valeur réelle de la charge et l'erreur aléatoire à l'instant t . B est l'opérateur de retard qui satisfait : $B^t y_t = y_{t-1}$. $\phi_P(B^S)$, $\Theta_Q(B^S)$ et $(1 - B^S)^D$ sont correspondants aux processus autorégressif, polynôme moyenne mobile et ordre de différenciation pour les composants saisonniers. Alors que $\varphi_p(B)$, $\theta_q(B)$ et $(1 - B)^d$ sont correspondants aux processus autorégressif, moyenne mobile et intégration pour la composante non saisonnière. S est la durée de la saison. Le modèle SARIMA adopté dans le présent chapitre peut être décrit comme SARIMA $(1, 1, 1) \times (1, 1, 1)_{12}$

3.3.5 Modèle de réseau de neurones artificiel à rétro-propagation

3.3.5.1 Rappel sur les réseaux de neurones

Les modèles de réseaux de neurones artificiels sont, à l'origine, une imitation du fonctionnement du cerveau. Le cerveau d'humain contient environ dix milliards de neurones et chacun est connecté à environ 10000 autres. Il fait la somme de toutes les informations qu'il reçoit et émet un signal si cette somme est suffisamment élevée [70]. L'objectif des réseaux de neurones artificiels est donc de modéliser le

fonctionnement des neurones réels, mais aussi de permettre un apprentissage. Un neurone artificiel peut être représenté comme dans la Figure 3.3.

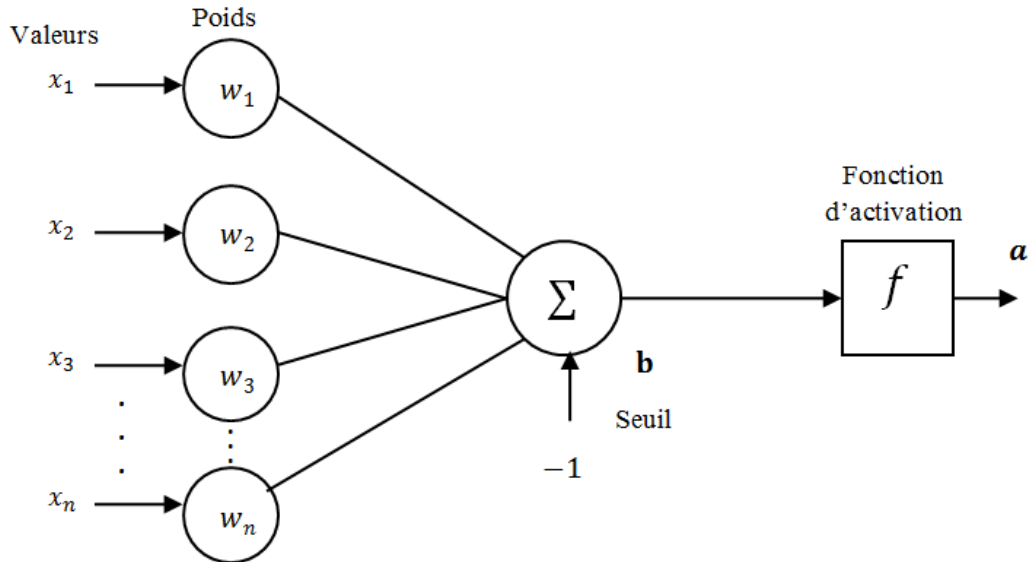


FIGURE 3.3 – Modèle d'un neurone artificiel

Le neurone calcule la somme de ses entrées (x_1, x_2, \dots, x_n) , pondérées par les poids synaptiques (w_1, w_2, \dots, w_n) , la compare à un seuil b . Ensuite cette valeur passe à travers la fonction d'activation f pour produire sa sortie. Le résultat après transfert est la sortie du neurone. La sortie peut être présentée par la formule suivante :

$$a = f \left(\sum_{j=1}^n w_j x_j - b_j \right) \quad (3.11)$$

Il existe de nombreuses formes possibles pour la fonction d'activation. Les plus courantes sont présentées sur la Figure 3.4.

Dans un réseau de neurones artificiel, les neurones sont regroupés en couches. Habituellement, chaque neurone dans une couche est connecté à tous les neurones dans la couche précédente et de la couche suivante (excepté pour la couche d'entrée et celle de sortie du réseau). L'information donnée à un réseau de neurones est propagée couche par couche de la couche d'entrée à la couche de sortie en passant par soit une ou plusieurs couches intermédiaires (cachées). L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle l'architecture du réseau et les poids des connexions entre les neurones sont modifiés jusqu'à l'obtention du comportement désiré. On distingue deux grandes classes d'algorithmes d'apprentissage : l'apprentissage supervisé et l'apprentissage non supervisé. Les réseaux à apprentissage supervisé sont généralement destinés à reproduire un processus quelconque dont on connaît seulement un comportement de référence précis (quelques

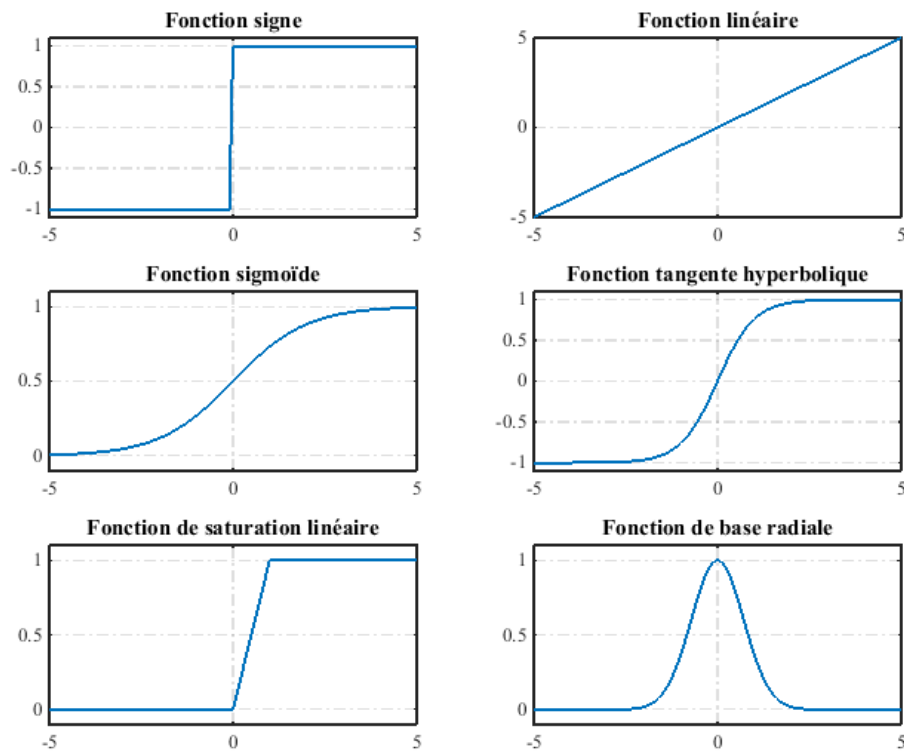


FIGURE 3.4 – Fonctions d’activation couramment utilisées [71]

entrées et les sorties correspondantes). Ainsi, le réseau va se modifier jusqu’à ce qu’il trouve la bonne sortie, c’est-à-dire celle attendue, correspondant à une entrée donnée. Les réseaux à apprentissage non supervisé sont généralement utilisés en classification lorsque les classes auxquelles doivent appartenir les données ne sont pas connues a priori. Les données servant à l’apprentissage sont donc différents dans les deux cas, puisqu’elles consistent en une série de couples (entrée ; sortie désirée correspondante) pour les apprentissages supervisés, tandis que pour les apprentissages non supervisés, les données sont uniquement des entrées et aucune réponse désirée n’est prise en considération.

3.3.5.2 Réseau à rétro-propagation

Il existe beaucoup d’architectures des réseaux de neurones, mais l’architecture la plus populairement utilisée est le perceptron multicouche à rétro-propagation de l’erreur. Le réseau à rétro-propagation est conçu pour qu’il fonctionne comme réseau multicouche, en utilisant la fonction d’apprentissage supervisée. Les réseaux à rétro-propagation possèdent la capacité de calculer le gradient de l’erreur pour chaque neurone, de la dernière couche vers la première, et de modifier donc les poids tant

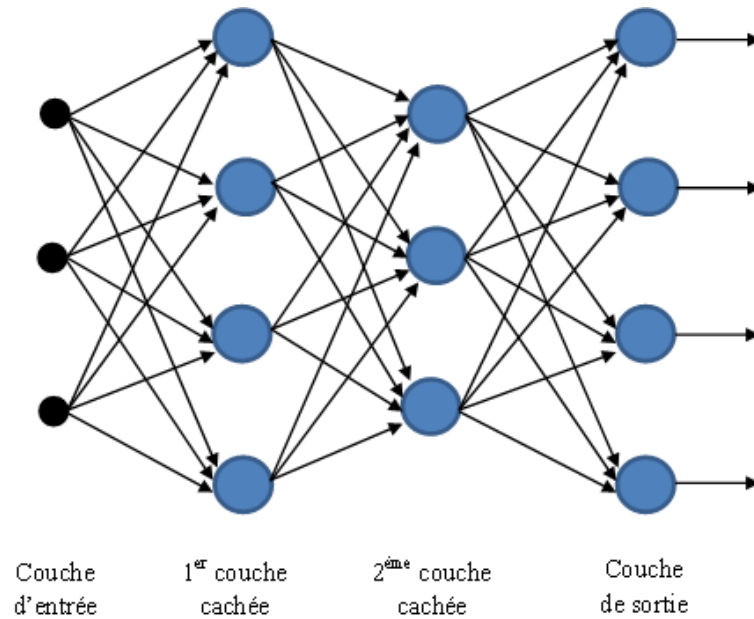


FIGURE 3.5 – Exemple d'un réseau de type perceptron multicouche

que la réponse correspondant à chaque entrée n'est pas assez proche de la réponse souhaitée. Une fois en fonctionnement, ils ont la capacité de généraliser (fournir des sorties satisfaisantes à des entrées que le système n'a jamais vues dans sa phase d'entraînement).

La Figure 3.5 donne l'exemple d'un réseau contenant une couche d'entrée, deux couches cachées et une couche de sortie. La couche d'entrée représente toujours une couche virtuelle associée aux entrées du système. Elle ne contient aucun neurone. Les couches suivantes sont des couches de neurones. Dans l'exemple illustré, il y a 3 entrées, 4 neurones sur la première couche cachée, trois neurones sur la deuxième et quatre neurones sur la couche de sortie. Les sorties des neurones de la dernière couche correspondent toujours aux sorties du système. Dans le cas général, un perceptron multicouche peut posséder un nombre de couches quelconque et un nombre de neurones (ou d'entrées) par couche également quelconque [72].

L'algorithme de rétro-propagation standard se résume à la série d'étapes suivantes :

1. Initialiser tous les poids à de petites valeurs aléatoires ;
2. Normaliser les données d'entraînement ;
3. Permuter aléatoirement les données d'entraînement ;
4. Pour chaque donnée d'entraînement n :
 - (a) Calculer les sorties observées en propageant les entrées vers l'avant ;
 - (b) Ajuster les poids en rétro-propageant l'erreur observée ;

$$w_{ji} = w_{ji}(n-1) + \Delta w_{ji}(n) = w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \quad (3.12)$$

où le "gradient local" est défini par :

$$\delta_j(n) = \begin{cases} e_j(n) y_j(n) [1 - y_j(n)] & \text{Si } j \in \text{couche de sortie} \\ y_j(n) [1 - y_j(n)] \sum_k \delta_k(n) w_{kj}(n) & \text{Si } j \in \text{couche cachée} \end{cases} \quad (3.13)$$

avec $0 \leq \eta \leq 1$ représentant le taux d'apprentissage et $y_i(n)$ représentant soit la sortie du neurone i sur la couche précédente, si celui-ci existe, soit l'entrée i autrement.

Il existe plusieurs algorithmes pour optimiser l'erreur et mettre à jour les paramètres. Néanmoins, l'algorithme le plus populairement utilisé est la méthode de Levenberg-Marquardt. Cet algorithme s'applique au cas particulier où l'erreur e est une erreur quadratique moyenne. En fait, cet algorithme est devenu un standard pour optimiser l'erreur quadratique due à ses propriétés de convergence rapide et de robustesse.

5. Répéter les étapes 3 et 4 jusqu'à un nombre maximum d'itérations ou jusqu'à ce que la racine de l'erreur quadratique moyenne soit inférieure à un certain seuil.

3.3.5.3 Modèle de réseau de neurones proposé

Le modèle de prévision proposé est un perceptron multicouche à rétro-propagation de l'erreur. L'architecture de ce réseau consiste en une couche cachée de trois neurones sigmoïde suivie par une couche de sortie du neurone linéaire.

Les données de charge de pointe pour les quatre années précédentes sont adoptées comme entrées pour le modèle de prévision à long-terme, tandis que le nombre d'entrées est réduit à une seule entrée (les données de la dernière année) dans le cas de prévision à moyen-terme. Dans les deux cas, les données de charge pour l'année suivante sont considérées comme la sortie du réseau. L'algorithme de Levenberg-Marquardt est conçu pour entraîner le réseau et la fonction utilisée pour évaluer la performance est l'erreur quadratique moyenne. Le nombre maximum d'itération pour l'entraînement du réseau est limité à 20 itérations.

3.3.6 Système d'inférence neuro-flou adaptatif

Le système d'inférence neuro-flou adaptatif est un système adaptatif qui combine les avantages du réseau de neurones artificiel et la logique floue en une seule capsule. Avant de détailler l'architecture de ce système neuro-flou et de présenter les modèles ANFIS proposés, nous donnons ci-dessous quelques notions sur la logique floue.

3.3.6.1 Rappel sur la logique floue

La logique floue, développée par Lofti A. Zadeh en 1965 à partir de sa théorie des sous-ensembles flous [73], constitue l'une des meilleures techniques de modélisation des différents phénomènes. La logique floue est aussi appelée « logique linguistique » car ses valeurs de vérité sont des mots du langage courant : « plutôt vrai, presque faux, loin, si loin, près de, grand, petit, ...etc ». La logique floue a pour objectif l'étude de la représentation des connaissances imprécises, des raisonnements approchés et la modélisation des notions vagues du langage naturel pour pallier l'inadéquation de la théorie des ensembles classiques dans ce domaine.

En théorie des ensembles classiques, l'appartenance d'un élément à un sous-ensemble est booléenne. Cependant, les classes d'objets rencontrés dans le monde physique ne possèdent pas de critères d'appartenance bien définis. Les sous-ensembles flous permettent en revanche de connaître le degré d'appartenance d'un élément au sous-ensemble. Un sous-ensemble flou \bar{A} d'un univers de discours U est caractérisé par la fonction suivante :

$$\mu_{\bar{A}} : U \rightarrow [0, 1] \quad (3.14)$$

où $\mu_{\bar{A}}$ est le niveau ou degré d'appartenance d'un élément de l'univers de discours U dans le sous-ensemble flou \bar{A} . On peut définir aussi le sous-ensemble flou \bar{A} dans l'univers de discours U comme suit :

$$\bar{A} = \{(x, \mu_{\bar{A}}) \mid x \in U\} \quad (3.15)$$

avec $\mu_{\bar{A}}(x)$ comme le degré d'appartenance de x dans \bar{A} .

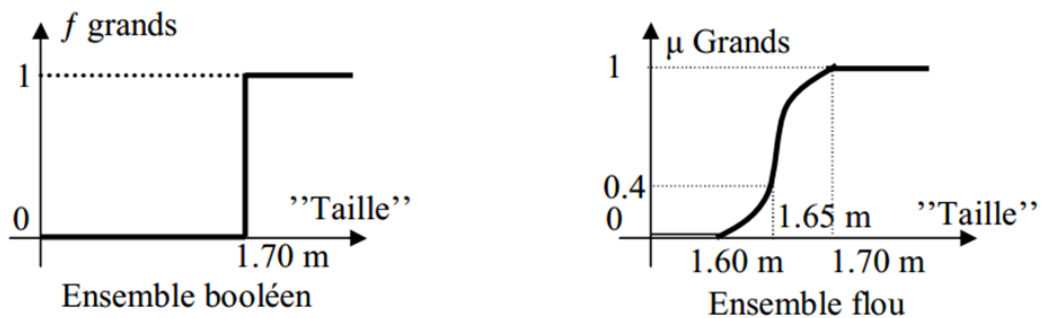


FIGURE 3.6 – Représentation de « Grands » dans le cas booléen et flou

En prend maintenant l'exemple de la Figure 3.6. Cet exemple présente l'univers de discours U défini sur \mathbb{R} et A le sous-ensemble classique pour représenter les tailles supérieures ou égales à 1.70m , où la fonction caractéristique est définie par :

$$\mu_A(x) = \begin{cases} 0, & \text{si } x < 1.7, \\ 1, & \text{si } x \geq 1.7, \end{cases} \quad (3.16)$$

Soit 1.65m la taille de x . Dans le cas de l'ensemble classique, x n'appartient pas à la classe des « Grands ». Dans le cas de l'ensemble flou, x appartient à la classe des « Grands » avec un degré d'appartenance 0.4. Le terme « Taille » est considéré donc comme une variable linguistique dont les valeurs sont des mots (petite, moyenne, grande). Une variable linguistique peut être représentée par un tripler $(x, T(x), U)$ dans lequel x est le nom de la variable linguistique, $T(x)$ l'ensemble des valeurs linguistiques de x et U l'univers de discours. Cependant, si nous avons considéré dans l'exemple qu'une personne mesurant 1.7m appartient au sous-ensemble des personnes de taille grande, la définition du sous-ensemble de la taille « Grande » se changera pour le cas des joueurs du basket-ball de la NBA. Il faut donc évaluer un ensemble de règles afin de transformer les données d'entrée en données de sortie. Cette fonction peut être effectuée avec succès en utilisant un système d'inférence floue (FIS).

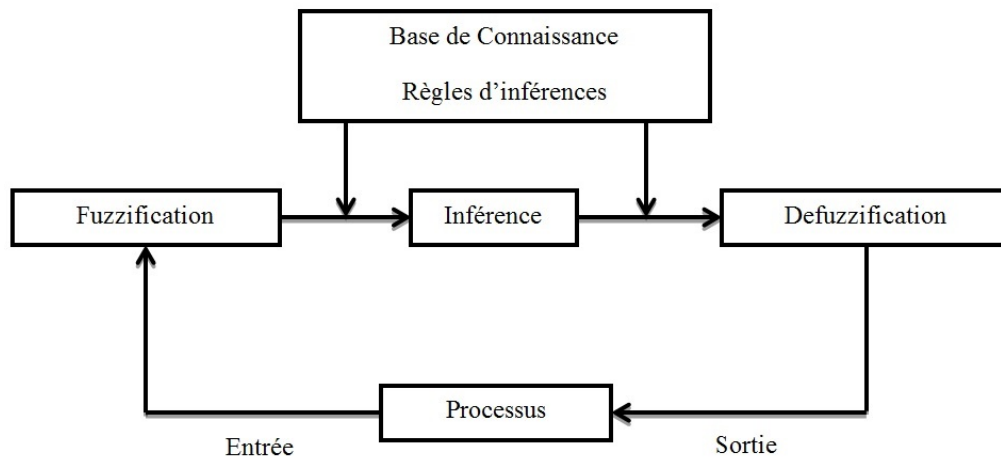


FIGURE 3.7 – Processus d'un système flou

Un système flou est un processus de formulation de relations entre des entrées et des sorties en utilisant les règles de la logique floue. A partir de ces relations, une décision peut être extraite ou une forme peut être découverte. Ce processus met en relief un certain nombre d'outils, que sont : les fonctions d'appartenance, les opérateurs logiques et les règles Si-Alors. De manière classique, le fonctionnement interne d'un système flou repose sur la structure présentée par la Figure 3.7 qui inclut quatre blocs :

- La base de connaissances.
- La fuzzification des variables d'entrée.
- L'inférence à partir d'une base de connaissance.
- La défuzzification.

1. La base de connaissances contient les définitions des fonctions d'appartenance (formes et paramètres) associées aux variables d'entrée/sortie ainsi que l'en-

semble des règles floues.

2. La fuzzification consiste à calculer, pour chaque valeur d'entrée numérique, les degrés d'appartenance aux ensembles flous associés et prédéfinis dans la base de données du système flou. Ce bloc réalise la transformation des entrées réelles en des parties symboliques floues utilisables par le mécanisme d'inférence.
3. La troisième étape est le moteur d'inférence, qui permet de calculer le degré de vérité des différentes règles du système et associer à chacune de ces règles une valeur de sortie. Ces règles floues peuvent être de type Mamdani (Eq. 3.17) comme elles peuvent également s'agir des règle de type Takagi-Sugeno (Eq. 3.18).

$$\text{SI (.....) Alors } Y \text{ est } B, \quad B \text{ valeur floue} \quad (3.17)$$

$$\text{SI (.....) Alors } Y \text{ est } B, \quad B \text{ valeur numérique} \quad (3.18)$$

Dans le modèle de Takagi-Sugeno, la conclusion est une valeur numérique. Dans le cas d'une règle floue de type Mamdani, la conclusion est une valeur floue. Après l'inférence il faut passer donc par une étape de défuzzification afin d'obtenir la valeur réelle de la sortie désirée.

4. La dernière étape est la défuzzification, qui consiste à remplacer l'ensemble des valeurs de sorties des différentes règles résultant de l'inférence par une valeur numérique unique représentative de cet ensemble.

Tableau 3.1 – Avantages et inconvénients de la logique floue et les réseaux de neurones

Réseaux de neurones	Logique floue
Avantages	
- Le modèle mathématique non requis	- Le modèle mathématique non requis
- Exploitation de données numériques	- La connaissance sur les règles est utilisable
- Plusieurs algorithmes d'apprentissage	- Une interprétation et implémentation simple
Inconvénients	
- Boîte noire	- Difficultés dans la construction des règles
- Réapprentissage est souvent obligatoire	- Ne peut pas apprendre
- Connaissance à base de règles non utilisable	- Adaptation difficile

3.3.6.2 Avantages de l'intégration de la logique floue et les réseaux de neurones

L'intégration et l'association de la logique floue et de la théorie des réseaux de neurones artificiels permettent l'exploitation des avantages des deux méthodes : la

prise de décision, la souplesse et la lisibilité de la première et les capacités d'apprentissage à partir de données mesurées de la seconde. Le Tableau 3.1 montre les avantages et les inconvénients de la logique floue et des réseaux de neurones, et résume la contribution des systèmes neuro-flous. En effet, les algorithmes d'apprentissages sont employés dans les modèles ANFIS pour déterminer les paramètres des systèmes flous, ce qui présente une amélioration automatique de ces derniers au moyen des méthodes spécifiques aux réseaux de neurones, mais le système reste toujours interprétable en termes de règles floues, qui est un aspect très important des systèmes neuro-flous [74].

3.3.6.3 Structure de l'ANFIS

Le système d'inférence neuro-flou adaptatif est l'un des systèmes les plus réussis qui combine les avantages du réseau de neurones artificiel et le système d'inférence floue en une seule capsule [75]. Un ANFIS fonctionne en appliquant des règles d'apprentissages neuronaux pour identifier et régler les paramètres et la structure d'un système d'inférence floue de type Takagi-Sugeno. Le réseau est un perceptron multicouche dans lequel chaque nœud (neurone) exécute une fonction particulière sur les signaux entrants. Pour des raisons de représentation, nous nous considérons un système à deux entrées et une sortie qui est aussi un modèle de 1er ordre composé de deux règles suivantes :

$$\text{Règle 1 : Si } x_1 \text{ est } A_1 \text{ et } x_2 \text{ est } B_1, \text{ Alors } y_1 = f_1(x_1, x_2) = a_1x_1 + b_1x_2 + c_1 \quad (3.19)$$

$$\text{Règle 2 : Si } x_1 \text{ est } A_2 \text{ et } x_2 \text{ est } B_2, \text{ Alors } y_2 = f_2(x_1, x_2) = a_2x_1 + b_2x_2 + c_2 \quad (3.20)$$

L'architecture d'ANFIS qui permet de représenter cette base de règles est réalisée par un réseau adaptatif qui contient des nœuds fixes (circulaires) et des nœuds adaptatifs (carrés) comme illustré dans la Figure 3.8.

Chaque nœud carré ou circulaire applique une fonction sur ses signaux d'entrées et pour une couche donnée les nœuds ont tous le même type de fonction. La sortie O_i^k d'un nœud i de la couche k (appelée nœud (i, k)) dépend des signaux provenant de la couche $k - 1$ et des paramètres du nœud (i, k) .

$$O_i^k = f \left(O_1^{k-1} \dots O_{n_{k-1}}^{k-1}, a, b, c, \dots \right) \quad (3.21)$$

Où n_{k-1} est le nombre de nœuds dans la couche $k-1$, et a, b, c sont les paramètres du nœud (i, k) . Il est à noter qu'un nœud circulaire ne possède pas de paramètres.

Couche 1

La première couche d'une architecture de type ANFIS comporte autant de neurones qu'il y a de sous-ensembles flous dans le système d'inférence représenté. Les

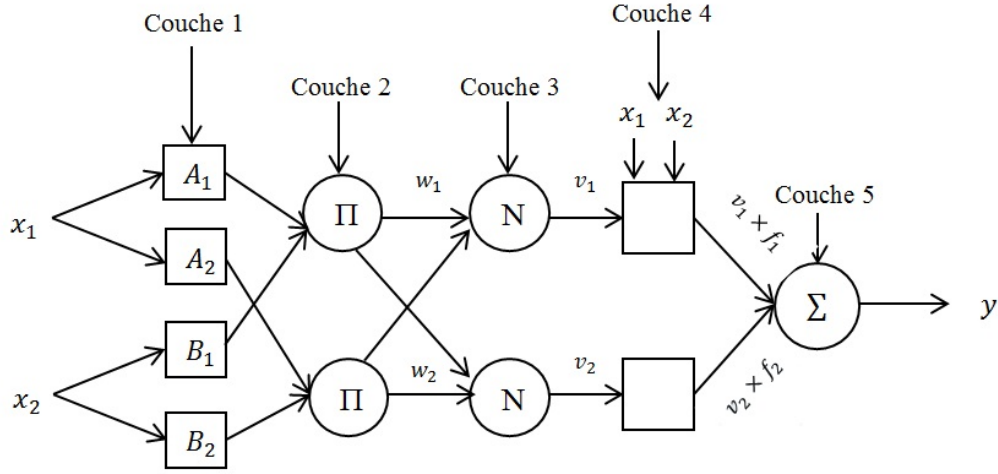


FIGURE 3.8 – Architecture d'ANFIS pour deux règles

nœuds de cette couche sont tous des nœuds adaptatifs. Cette couche réalise la fuzzification des entrées c'est-à-dire qu'elle détermine le degré de vérité d'un sous ensemble flou particulier par sa fonction de transfert :

$$O_i^1 = \mu_{A_i}(x) \quad (3.22)$$

Où x est l'entrée du nœud i , A_i est un sous ensemble flou correspondant au x (variable linguistique) et O_i^1 est le degré d'appartenance de x à A_i . En d'autres termes, O_i^1 est la fonction d'appartenance de A_i , et il spécifie le degré auquel l'entrée donnée satisfait le quantificateur A_i . $\mu_{A_i}(x)$ puisse adopter n'importe quelle fonction d'appartenance floue avec le maximum égal à 1 et le minimum égal à 0. Cependant, les plus couramment utilisées sont les fonctions sigmoïdes, les fonctions d'appartenance en forme de cloche, les fonctions d'appartenance gaussiennes, la fonction en forme de triangle, et la fonction trapézoïdale.

$$\text{Sigmoïdal : } \mu_{A_i}(x) = \frac{1}{1 + e^{-a_i(x-c_i)}} \quad (3.23)$$

$$\text{Gaussienne : } \mu_{A_i}(x) = \exp\left(-\frac{(x-c_i)^2}{2(\sigma_i)^2}\right) \quad (3.24)$$

$$\text{Forme de cloche : } \mu_{A_i}(x) = \frac{1}{1 + \left(\left(\frac{x_1-c_i}{a_i}\right)^2\right)^{b_i}} \quad (3.25)$$

$$\text{Triangulaire : } \mu_{A_i}(x) = \max\left(\min\left(\frac{x-a_i}{b_i-a_i}, \frac{c_i-x}{c_i-b_i}\right), 0\right) \quad (3.26)$$

$$\text{Trapézoïdale : } \mu_{A_i}(x) = \max \left(\min \left(\frac{x - a_i}{b_i - a_i}, 1, \frac{d_i - x}{d_i - c_i} \right), 0 \right) \quad (3.27)$$

Où $\{a, b, c, d, \sigma\}$ est l'ensemble des paramètres. Pendant que les valeurs de ces paramètres changent, les fonctions en formes précédentes changent en conséquence, de ce fait présenter de diverses formes de fonction d'appartenance sur la variable linguistique A_i . Des exemples pour les différentes formes de ces fonctions sont illustrés dans la Figure 3.9 et la Figure 3.10.

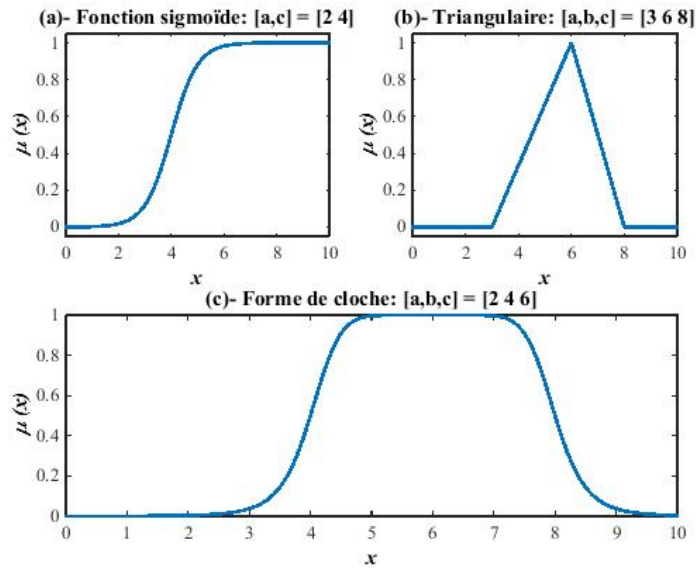


FIGURE 3.9 – Fonction d'appartenance couramment utilisées : (a)- sigmoïde, (b)- triangulaire, (c)- forme de cloche

Couche 2

Les nœuds de cette couche sont tous des nœuds fixes, dont chaque neurone correspond à une règle floue de type Sugeno. Ils reçoivent les signaux de sortie de la couche précédente (le degré de vérité des différents sous-ensembles flous) et envoient leur produit en sortie (le degré de vérité d'une règle floue) :

$$w_i = \mu_{A_i}(x_1) \times \mu_{B_i}(x_2) \quad (3.28)$$

w_i présente le degré de vérité de la règle i .

Couche 3

Chaque neurone dans cette couche est un neurone fixe. L'objectif de cette couche est de calculer le rapport entre le poids de l' $i^{\text{ème}}$ règle et la somme de toutes les poids

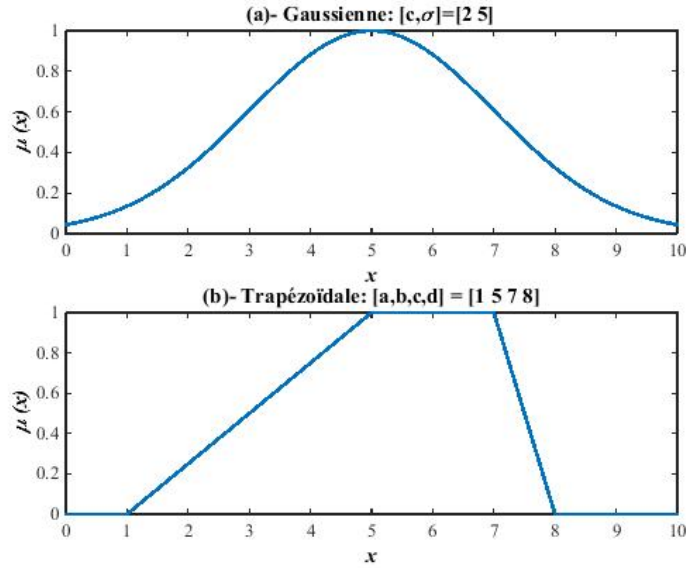


FIGURE 3.10 – Fonction d'appartenance couramment utilisées : (a)- gaussienne, (b)- trapézoïdale

des règles. Cette opération permet donc de calculer le degré de vérité normalisé d'une règle floue donnée :

$$v_i = \frac{w_i}{w_1 + w_2} \quad (3.29)$$

Le résultat à la sortie de chaque nœud représente la contribution de cette règle au résultat final.

Couche 4

Les nœuds dans cette couche sont des nœuds adaptatifs. Chacun de ces nœuds est relié à un neurone de normalisation correspondant et aux entrées initiales du réseau. La sortie d'un nœud i est le conséquent pondéré de la règle, donnée par :

$$O_i^4 = v_i \times f_i = v_i (a_i x_1 + b_i x_2 + c_i), i = 1, 2 \quad (3.30)$$

Où v_i est la sortie de la couche 3, et $\{a_i, b_i, c_i\}$ est l'ensemble des paramètres de sortie de la règle i .

Couche 5

Cette couche comprend un seul neurone circulaire qui effectue la somme des signaux provenant de la couche précédente pour donner la sortie finale du réseau :

$$O_i^5 = \sum_{i=1}^2 v_i \times f_i = \frac{\sum_{i=1}^2 w_i \times f_i}{\sum_{i=1}^2 w_i} \quad (3.31)$$

3.3.6.4 Algorithme d'apprentissage de l'ANFIS

Le système ANFIS applique le mécanisme d'apprentissage des réseaux neurones sur des techniques d'inférences floues. L'apprentissage à partir d'un ensemble de données concerne l'identification des paramètres des prémisses et des conséquences, la structure du réseau étant fixée. D'un autre terme L'ANFIS est un système d'inférence floue dont les paramètres des fonctions d'appartenances et les combinaisons linéaires des paramètres conséquents sont ajustés en utilisant l'algorithme d'apprentissage par rétro-propagation de l'erreur. Jang a proposé d'utiliser une règle hybride d'apprentissage qui combine un algorithme de descente du gradient avec une estimation par moindres carrées.

ANFIS utilise un cycle d'apprentissage de deux passages : le passage en avant et le passage en arrière. Dans le passage avant, les paramètres conséquents sont calculés en utilisant l'algorithme de moindre carré de l'erreur. Dans le passage en arrière, les paramètres des prémisses sont calculés en utilisant l'algorithme de rétro-propagation de l'erreur [69]. L'algorithme pour ajuster la production des paramètres du système d'inférence neuro-flou adaptatif est présenté par la Figure 3.11.

3.3.6.5 ANFIS et maîtrise de l'erreur de prédiction

Beaucoup de paramètres influencent potentiellement sur l'erreur de prédiction du système d'inférence neuro-flou adaptatif. Cependant, la recherche d'un réseau « performant » peut s'orienter suivant les axes cités ci-dessous [76].

- **Structure du réseau.** Concernant le réseau neuro-flou ANFIS, définir une architecture consiste à choisir le nombre de fonctions d'appartenance, leur type (gaussienne, sigmoïde, ...etc) et le nombre de règles floues utilisées. Un compromis doit être trouvé : il n'est pas concevable de construire un système ANFIS trop complexe, la base d'apprentissage étant, elle, limitée.
- **Sélection d'entrées.** La sélection des variables d'entrées est une méthode expérimentale qui vise à déterminer l'ensemble des entrées le plus adéquat pour approximer la relation entrées/sortie. Cette approche est basée sur la construction de groupes d'entrées partiellement disjoints et sur l'évaluation rapide des architectures ainsi obtenues en regard d'un objectif d'apprentissage immédiat (une itération).
- **Base d'apprentissage.** L'ensemble d'apprentissage disponible impacte bien entendu la qualité de généralisation du système ANFIS. Il peut être nécessaire de disposer une base d'apprentissage de taille 5 fois supérieure au nombre de paramètres à optimiser (prémisses et conséquents). Cependant, au-delà de la taille de l'ensemble d'apprentissage, c'est également la pertinence des données qu'il contient qui se révèle critique.

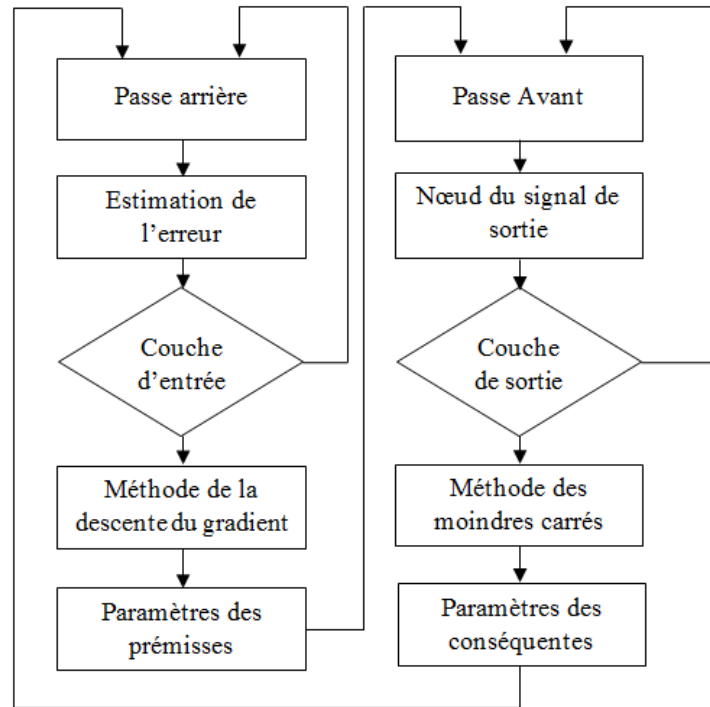


FIGURE 3.11 – Algorithme d'apprentissage d'ANFIS pour l'ajustement de la production des paramètres des règles

- **Algorithme d'apprentissage.** Le critère d'arrêt des algorithmes d'apprentissage est important afin d'éviter les problèmes de « sur- et de sous-apprentissage » (nombre maximum d'itérations ou seuil de l'erreur admise).

3.3.6.6 Modèles ANFIS proposés

La structure du système d'inférence neuro-flou proposé utilise la fonction d'appartenance sigmoïde et vingt époques pour identifier les paramètres du système d'inférence floue et la topologie optimale de l'ANFIS. Le nombre de règles floues est égal à deux ou à seize, selon qu'il s'agit respectivement, d'un cas de prévision à moyen ou à long-terme de la charge électrique. En outre, les données de charge de pointe pour les quatre années précédentes sont adoptées comme entrées pour le modèle de prévision à long-terme, tandis que le nombre d'entrées est réduit à une seule entrée dans le cas de prévision à moyen-terme de la charge électrique.

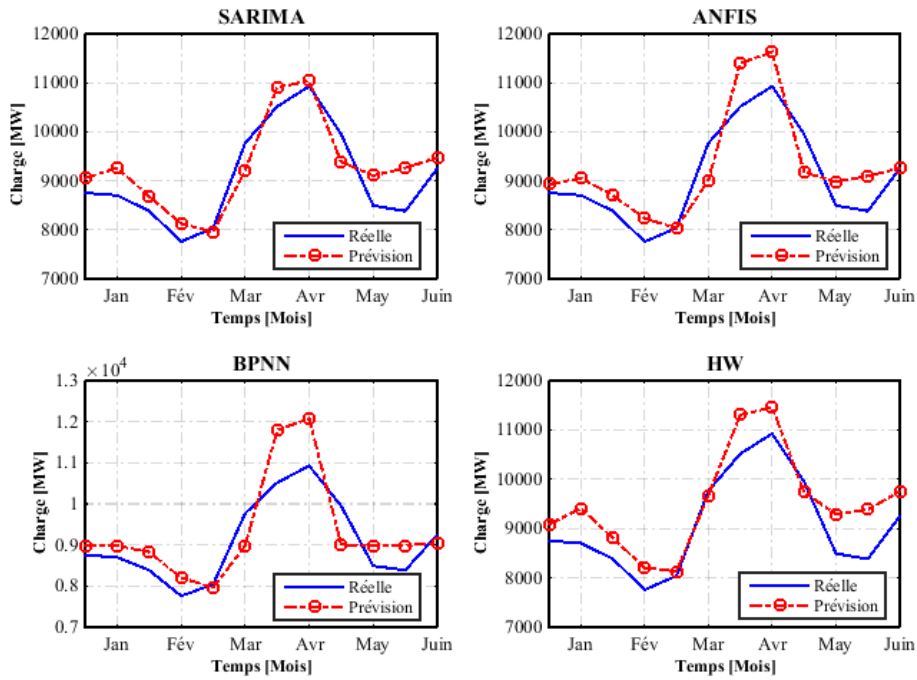


FIGURE 3.12 – Résultats des méthodes de prévision à moyen-terme : 2014

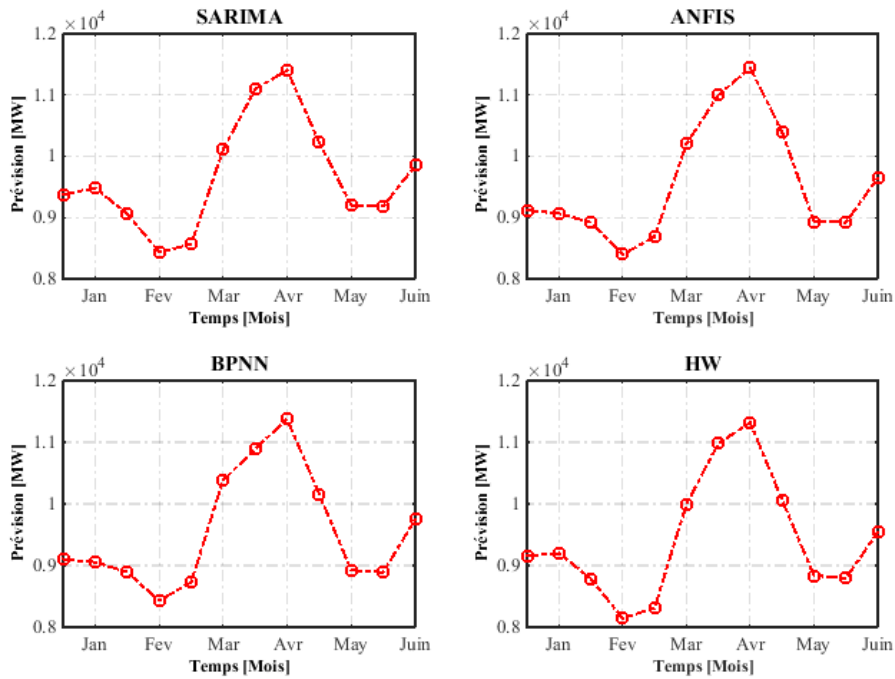


FIGURE 3.13 – Résultats des méthodes de prévision à moyen-terme : 2015

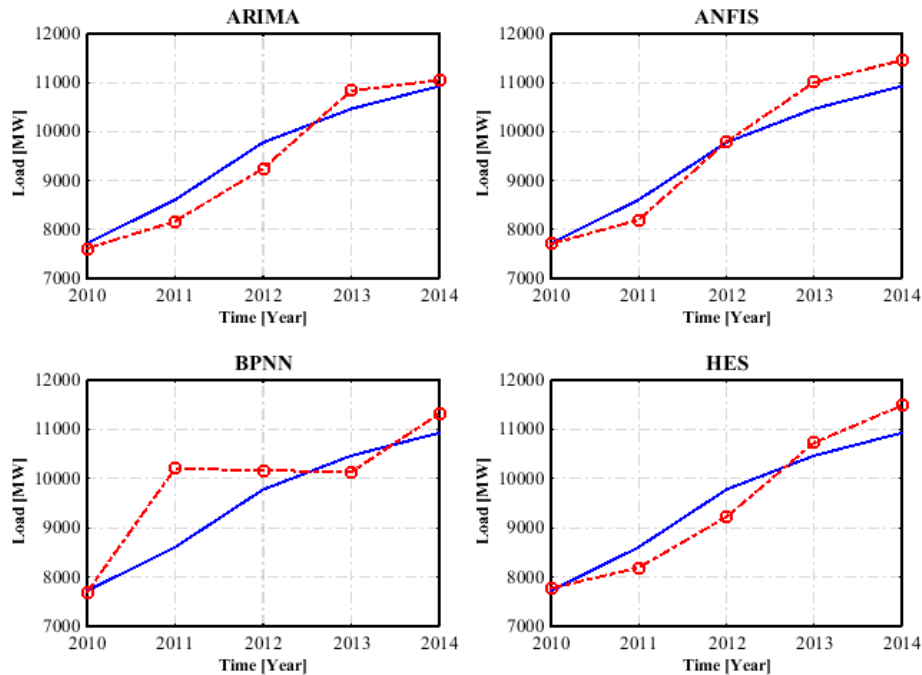


FIGURE 3.14 – Résultats des méthodes de prévision à long-terme : 2010-2014

3.4 Application pour la prévision à moyen et à long-terme de la charge électrique

Dans cette section, nous allons tester les performances des modèles de prédiction à moyen-terme à travers une simulation de la charge de pointe mensuelle en Algérie tout le long de deux années successives : 2014 et 2015. Pour les modèles de prévision à long-terme, nous allons évaluer des prévisions pour une seule étape à l'avance (un an) sur la période entre 2010 et 2014, et pour six années à l'avance pour la période entre 2015 et 2020.

Les résultats des modèles de prévision à moyen et à long-terme sont présentés sur les Figures 3.12-3.13 et les Figure 3.14-3.15, respectivement. Ces figures montrent que les modèles proposés ont prédit avec succès la demande de pointe d'électricité, puisque la forme de courbes de charge prédites suit globalement la forme de courbes de charge réelles. Cependant, on peut observer que la précision du modèle de réseau de neurones artificiel pour le cas de la prévision à long-terme est parfois insatisfaisante. Par exemple, on peut clairement noter une surestimation de la charge de pointe pendant l'année 2011. En outre, le modèle n'a pas été capable de fournir des prévisions satisfaisantes pour plusieurs étapes en avance dans la période entre 2015 et 2020, car la croissance des charges annuelles prévues dans la période citée est relativement faible. Cette baisse en précision est principalement liée au manque d'un

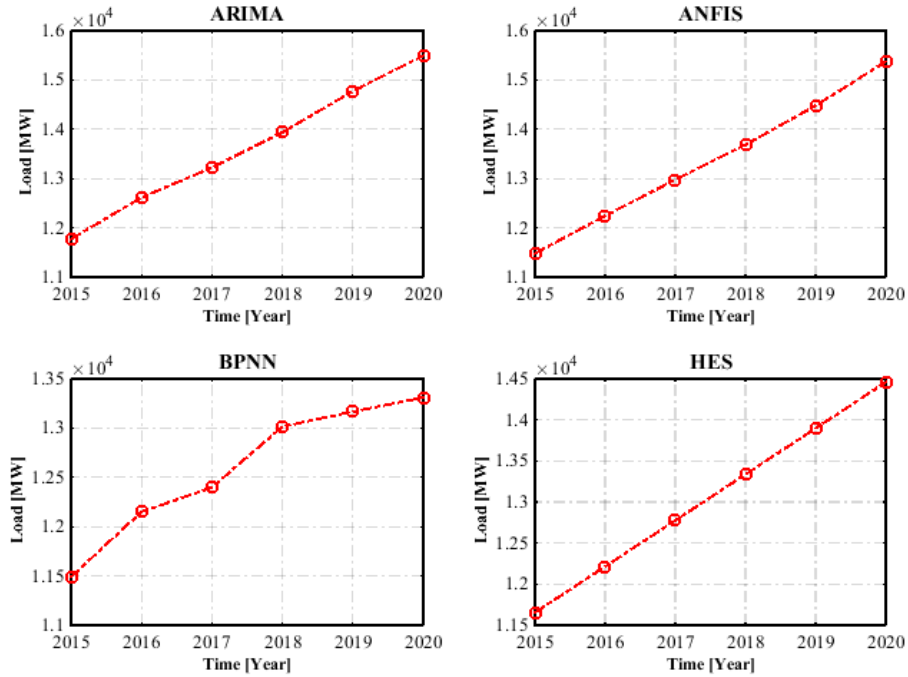


FIGURE 3.15 – Résultats des méthodes de prévision à long-terme : 2015-2020

Tableau 3.2 – Performances des modèles proposés pour MTLF

Précision	Méthode	HW	SARIMA	BPNN	ANFIS
MAPE [%]		05.517	04.620	06.082	05.058
RMSE [MW]		566.96	467.85	685.32	553.05

nombre suffisant de données pour entraîner le réseau. par contre, on peut remarquer que les modèles d’ANFIS ont été capables, avec peu de données, de fournir des résultats satisfaisantes pour à la fois la prévision à moyen et à long-terme de charge électrique. En outre, des méthodes statistiques plus simples telles que le lissage exponentiel double et le modèle ARIMA sembleraient avoir assuré des performances raisonnables.

Cependant, les résultats de prévision détaillés, comme indiqué sur le Tableau 3.2 et le Tableau 3.3, montrent une précision supérieure du modèle basé sur le système d’inférence neuro-flou adaptatif dans le cas de prévision à long-terme de la charge, avec une valeur de MAPE de 3.017%. La meilleure approche pour la prévision à moyen-terme de la charge électrique est le modèle autorégressif saisonnier à moyenne mobile intégrée, avec une valeur de MAPE de 4.620%. En outre, les deux meilleures méthodes pour à la fois la prévision à moyen et à long-terme sont le modèle ANFIS et les approches basées sur les séries temporelles (ARIMA ou SARIMA).

3.4. Application pour la prévision à moyen et à long-terme de la charge électrique 60

Tableau 3.3 – Performances des modèles proposés pour LTLF

Précision	Méthode	HES	ARIMA	BPNN	ANFIS
MAPE [%]		03.770	03.320	05.925	03.017
RMSE [MW]		415.64	357.66	769.36	385.18

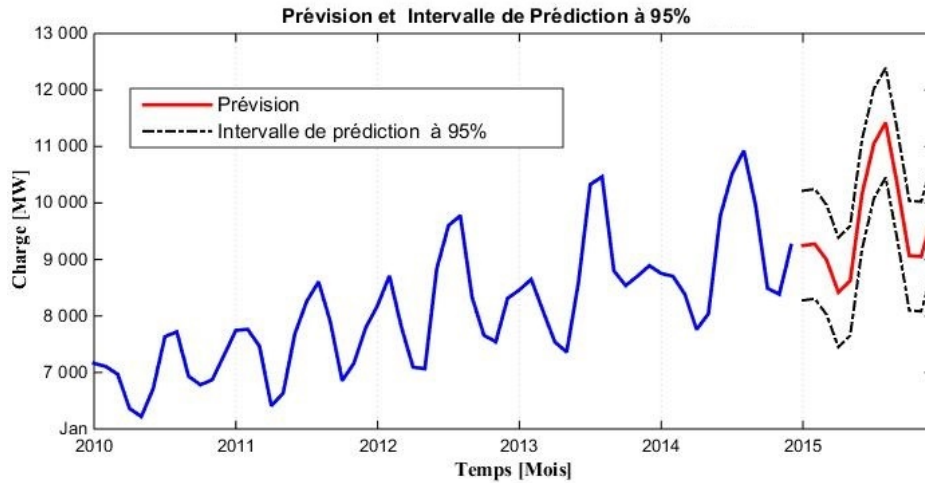


FIGURE 3.16 – Les prévisions combinées de SARIMA-ANFIS tracées conjointement avec l'intervalle prévu pour l'année 2015

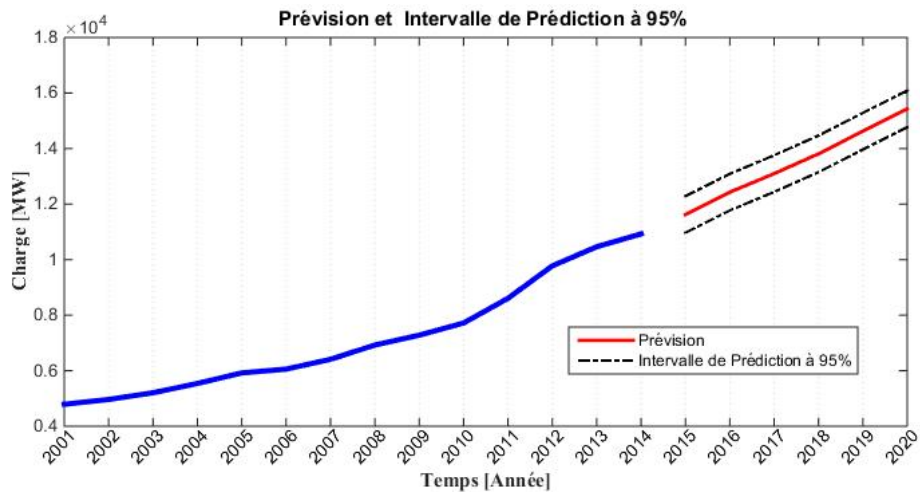


FIGURE 3.17 – Les prévisions combinées de ARIMA-ANFIS tracées conjointement avec l'intervalle prévu pour la période entre 2015 et 2020

3.4. Application pour la prévision à moyen et à long-terme de la charge électrique 61

En plus de la nécessité des prévisions précises, il est également important d'estimer un intervalle dans lequel les futures observations de la charge électrique peuvent être situées. Dans la Figure 3.16, une simple combinaison des prévisions issues des modèles SARIMA et ANFIS pour 2015, est tracée et accompagnée ainsi avec des intervalles prévus. Dans la Figure 3.17, la combinaison des prévisions entre ARIMA et ANFIS est également tracée avec les intervalles prévus pour la prévision des pics annuels sur la période entre 2015 et 2020.

Tableau 3.4 – SARIMA-ANFIS : Résultats pour la prévision des pics mensuels pour l'année 2015

Mois	Prévision	Mois	Prévision	Mois	Prévision
01-2015	9247	05-2015	8624	09-2015	10311
02-2015	9275	06-2015	10154	10-2015	9068
03-2015	8995	07-2015	11048	11-2015	9052
04-2015	8419	08-2015	11423	12-2015	9747

Tableau 3.5 – ARIMA-ANFIS : Résultats pour la prévision des pics annuels pour la période entre 2015-2020

Année	Prévision	Année	Prévision	Année	Prévision
2015	11627	2017	13098	2019	14626
2016	12430	2018	13811	2020	15431

Les charges prévues pour 2015 et pour la période 2015-2020 sont également présentées sur le Tableau 3.4 et le Tableau 3.5, respectivement. La consommation de pointe durant l'été (Août 2015) pour SARIMA-ANFIS est supposée être 11423 MW. Cette pointe est 4,539% plus élevée que le pic d'Août 2014 (10927 MW). Nous estimons aussi que la croissance moyenne de la charge de pointe annuelle sur la période entre 2015 et 2020 sera égale à 5,826%; et la charge pour 2020 est censée être multipliée par environ deux si nous prenons la demande d'électricité de l'année 2010 comme référence.

Les résultats obtenus ci-dessus en été présentées à une conférence internationale en 2015 [77]. Donc, à cette époque-là nous ne savions rien sur les performances de nos résultats pour l'année 2015 et pour la période entre 2015 et 2020. Cependant, durant la rédaction de la thèse nous avons l'occasion de mettre à jour la base de données de la charge électrique. A cet effet, nous avons collecté les données pour les premiers cinq mois de l'année 2015. Notre curiosité nous amène donc à vérifier les résultats de prévision des pics mensuels pour ces cinq mois. Nous présentons dans la Figure 3.18 l'erreur absolue en pourcentage pour la période entre Janvier 2015 et Mai 2015. Nous pouvons constater que ces résultats sont en consistence avec les résultats obtenus pour l'année 2014, puisque la valeur maximale de l'erreur absolue en pourcentage est de l'ordre de 4%.

3.4. Application pour la prévision à moyen et à long-terme de la charge électrique

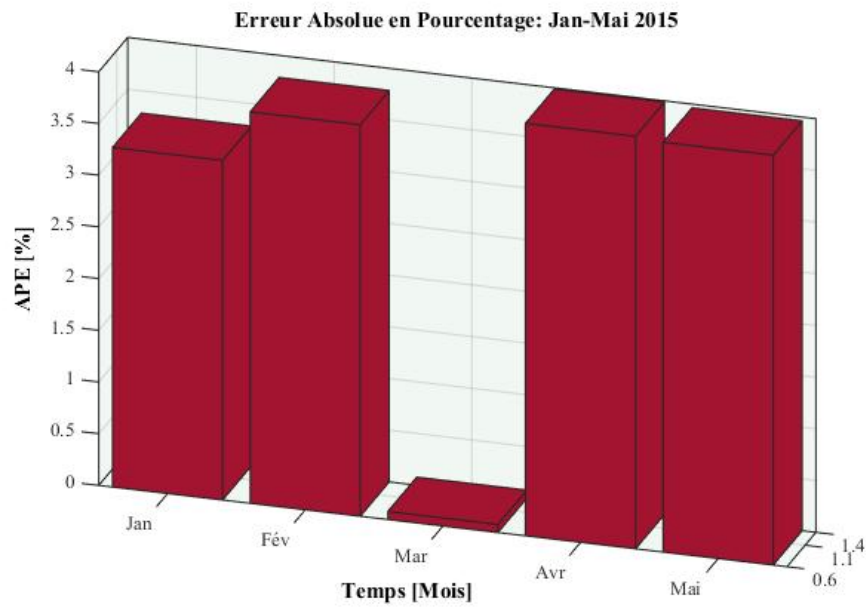


FIGURE 3.18 – SARIMA-ANFIS : Erreur absolue en pourcentage pour la période Jan-Mai 2015

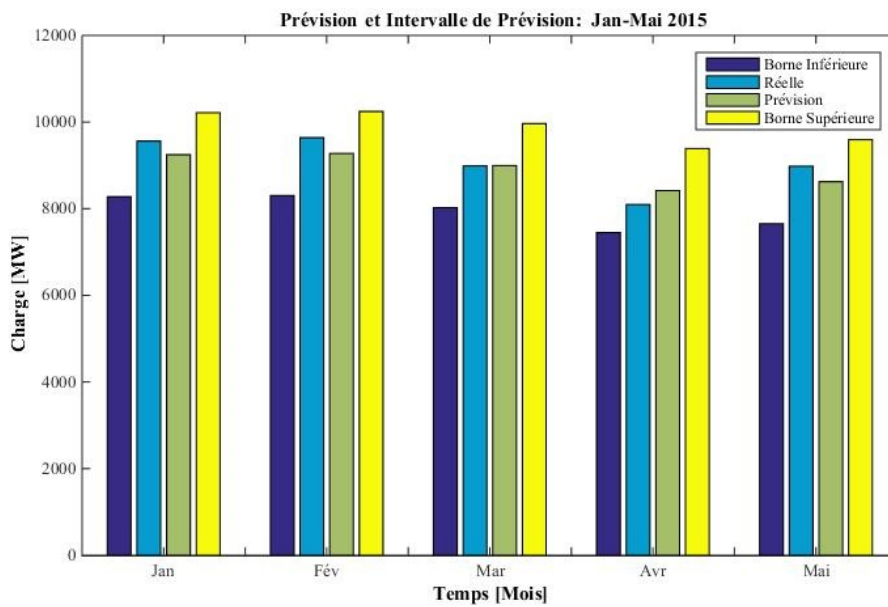


FIGURE 3.19 – SARIMA-ANFIS : Prévision et intervalle de prévision pour la période Jan-Mai 2015

La Figure 3.19 et le Tableau 3.6 montrent également les résultats de prévision pour la période entre Janvier 2015 et Mai 2015 conjointement avec l'intervalle de prévision pour la même période. Nous pouvons constater pour une autre fois l'efficacité de notre démarche, puisque aucun résultat de prévision n'est situé en dehors de l'intervalle prévu.

Tableau 3.6 – SARIMA-ANFIS : Prévision et intervalle de prévision pour la période Jan-Mai 2015

Mois	Borne inférieure	Pointe Réelle	Pointe prévue	Borne supérieure
Janvier 2015	8277	9561	9247	10217
Février 2015	8305	9640	9275	10245
Mars 2015	8025	8988	8995	9965
Avril 2015	7449	8096	8419	9389
Mai 2015	7654	8980	8624	9594

3.5 Conclusion

Ce chapitre a utilisé les observations de la demande d'électricité en Algérie pour évaluer des prédictions allant de quelques mois à plusieurs années à venir. Les résultats de prévision ont montré que le système d'inférence neuro-flou adaptatif et les modèles basés sur le processus autorégressif à moyenne mobile intégré ont été bien performés pour les deux cas de prévision à moyen et à long-terme de la demande d'électricité, étant donné que la valeur d'erreur absolue moyenne en pourcentage n'a pas excédée 5% dans les deux cas. Cependant, ce n'est pas très clair laquelle des deux méthodes est la plus performante : le système d'inférence neuro-flou adaptatif a été le meilleur modèle pour la prévision à long-terme des pics de charge électrique, alors que la meilleure performance a été obtenue par le processus saisonnier autorégressif à moyenne mobile intégré pour le cas de prévision à moyen-terme. Des méthodes plus simples telles que le lissage exponentiel de Holt ont été également bien concurrencées dans la situation où les données historiques contiennent une tendance.

Il est conseillé donc d'utiliser plusieurs méthodes de prévisions afin d'améliorer la précision de prévision et tirer la meilleure performance dans le maximum de cas possibles. On peut utiliser par exemple une méthode statistique classique telle que la méthode SARIMA lorsque les données comprennent un terme de saisonnalité. En outre, lorsque des facteurs exogènes sont disponibles, on peut utiliser des techniques de l'intelligence artificielle pour bien modéliser la relation existante entre ces facteurs exogènes et la charge électrique. Cette idée sera l'élément clé dans la méthodologie présentée dans le chapitre suivant pour prévoir les pics quotidiens de la consommation d'électricité.

Méthodologie hybride adaptative à deux étapes dédiée à la prévision de la pointe journalière

4.1 Introduction

Ce chapitre décrit la prévision des pics quotidiens de la charge électrique en utilisant une méthodologie hybride adaptative à deux étapes. Vu que la série chronologique de la consommation d'électricité est principalement influencée par des effets saisonniers, la méthode de lissage exponentiel à double saisonnalité de Holt-Winters-Taylor est utilisée en premier lieu pour prévoir la demande de pointe du jour suivant. Dans la deuxième étape, un modèle de prévision secondaire est appliqué en tenant compte des avantages de classification par la méthode de c-moyennes floues, algorithme des k-plus proches voisins, décomposition en paquets d'ondelettes, et le système d'inférence neuro-flou adaptatif ; pour améliorer encore la précision de prévision. L'ensemble de l'architecture du modèle proposé sera présentée et les résultats seront comparés avec des modèles basés sur les réseaux de neurones et le système d'inférence neuro-flou adaptatif en utilisant des données recueillies à partir du réseau électrique algérien.

4.2 Série chronologique de la demande de pointe journalière d'électricité en Algérie

Dans cette étude, nous considérons les données de charge de pointe qui se composent de trente-cinq mois d'observations pour la demande d'électricité en Algérie du 01 Janvier 2012 au 30 Novembre 2014. Ces données ont été recueillies à partir du site web de l'opérateur du système électrique algérien(OSE) [64]. La partie supérieure de la Figure 4.1 montre que les données utilisées sont constituées de deux séries temporelles.

La première série enregistre la demande de pointe du matin, tandis que la seconde série se compose des observations de la charge de pointe du soir. On peut observer que la charge quotidienne de pointe se produit généralement au cours de la période de soirée. Par conséquent, nous avons limité notre intérêt dans la présente étude à la

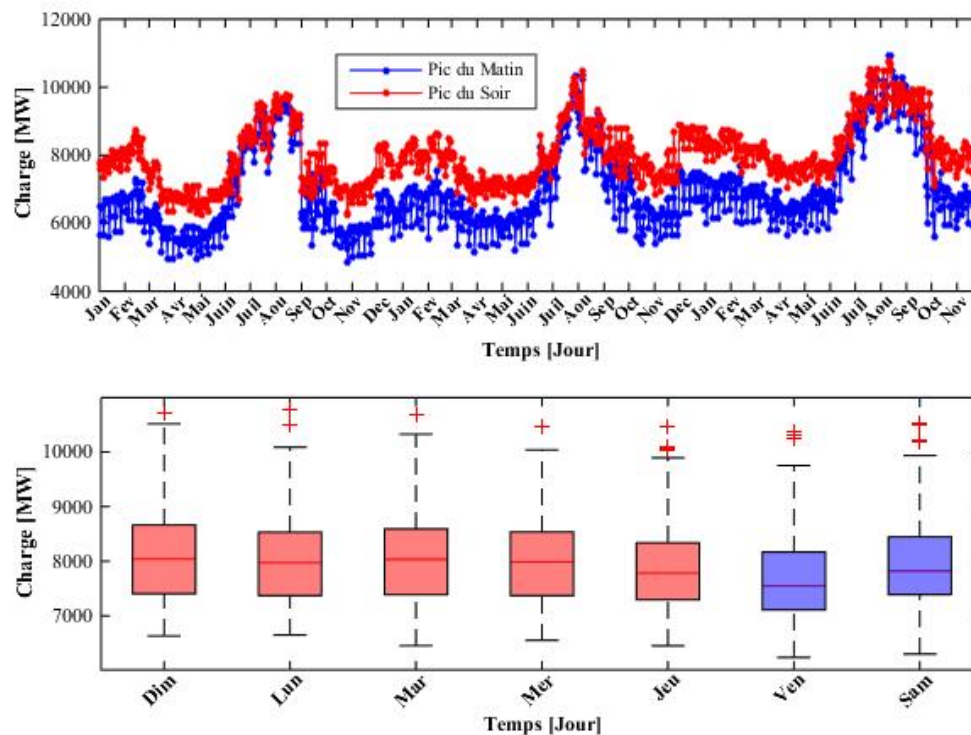


FIGURE 4.1 – La demande de pointe d'électricité du 01 Janvier 2012 au 30 Novembre 2014

prévision des pics du soir. La partie inférieure de la Figure 4.1 est une présentation en boîtes à moustaches pour les données de la demande de pointe du soir. On peut percevoir un cycle saisonnier hebdomadaire, où les jours ouvrables montrent des tendances similaires de la demande, mais les weekends (vendredis) sont quelque peu différentes et la charge sur lesquels est relativement faible par rapport aux jours ouvrables.

En considérant les deux séries citées ci-dessus, nous pouvons percevoir que les données contiennent un cycle d'une saisonnalité journalière composé par une durée de deux périodes (pointe du matin et pointe du soir), et un cycle hebdomadaire d'une durée de quatorze périodes (jours ouvrables et le weekend). Par conséquent, notre choix a été tombé sur l'utilisation du modèle de lissage exponentiel à double saisonnalité de Holt-Winters-Taylor (HWT), dont le but est d'essayer d'inclure le cycle saisonnier intra-journalier et le cycle intra-hebdomadaire dans une seule capsule. Ceci est l'approche prise par J.W. Taylor dans le cadre de leur modélisation de la consommation demi-horaire d'électricité en Angleterre et en Pays de Galles pour assurer des prévisions jusqu'à une journée à l'avance [78]. Cette méthode assure de très bons résultats dans le cas de jours normaux, mais il peut faire face à des dif-

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁶⁶

ficultés lorsqu'elle est soumise à des conditions journalières spéciales comme le cas de l'évolution rapide de la demande d'électricité pendant les événements spéciaux. L'architecture complète de la méthodologie proposée pour la prévision de la charge de pointe dans des conditions quotidiennes normales et spéciales sera présentée dans la section suivante.

Pour évaluer la performance de la stratégie proposée, les données de charge de pointe pendant quatre mois différents : Janvier 2014, Avril 2014, Juillet 2014, et Novembre 2014 ; et pour les jours fériés de 2014 ont été utilisées comme un exemple illustratif.

4.3 Méthodologie proposée pour la prévision de la pointe journalière

La méthode proposée pour la prévision des pics quotidiens de charge peut être vue comme une méthodologie hybride et adaptative qui consiste de deux phases principales [79]. Par le biais des records de la demande de pointe d'électricité dans les deux périodes du matin et du soir, HWT est appliquée en tant que premier volet au jour J_{L-1} pour assurer une prévision primaire pour le jour J_L . À son côté, la deuxième partie de notre modèle se compose d'un système de prévision secondaire, nommé FKW-ANFIS, et composé de cinq éléments de base. Tout d'abord, les observations de température sont utilisées comme entrées à la méthode de regroupement par c-moyennes floues (FCM) pour classifier les données de pics de charge en deux classes (charge d'une période chaude ou froide). Ensuite, les charges de pointe du matin, les données de pointe du soir, et la prévision primaire du HWT ; sont employés comme entrées pour l'algorithme des k-plus proches voisins (k-NN) pour distinguer si le jour J_L est un jour ouvrable ou un week-end et sélectionner alors les échantillons les plus proches qui ont une caractéristique similaire au jour J_L . La troisième partie du modèle secondaire décompose les données clustérisées de la charge de pointe du matin et du soir ; via la méthode de décomposition en paquets d'ondelettes (WPD), en parties d'approximation associées aux basses fréquences et parties détaillées associées à des fréquences élevées. Lors de la quatrième partie, le système d'inférence neuro-flou adaptatif est introduit pour prédire les futures valeurs de chaque espace de paquets d'ondelettes. Enfin, les valeurs de pics du soir prévues de tous les espaces sont additionnées pour produire le résultat final de la prévision de la demande d'électricité.

4.3.1 Méthode de Holt-Winters-Taylor

La technique de Holt-Winters-Taylor est une extension du lissage exponentiel de Holt-Winters standard, qui a été introduite pour tenir compte les cycles saisonniers journalier et hebdomadaire de la série de charge électrique. La formulation pour la méthode HWT est donnée dans les expressions suivantes [80] :

$$S_t = \alpha(y_t/D_{t-s1}W_{t-s2}) + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (4.1)$$

4.3. Méthodologie proposée pour la prévision de la pointe journalière 67

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)(T_{t-1}) \quad (4.2)$$

$$D_t = \delta(y_t/S_t W_{t-s_2}) + (1 - \delta)D_{t-s_1} \quad (4.3)$$

$$W_t = \omega(y_t/S_t D_{t-s_1}) + (1 - \omega)W_{t-s_2} \quad (4.4)$$

$$\hat{y}_t(k) = (S_t + kT_t)D_{t-s_1+k}W_{t-s_2+k} + \phi^k(y_t - ((S_{t-1} + T_{t-1})D_{t-s_1}W_{t-s_2})) \quad (4.5)$$

S_t et T_t , correspondent au niveau et la pente à l'instant t ; D_t et W_t sont respectivement le cycle journalier s_1 ($s_1 = 2$) et hebdomadaire s_2 ($s_2 = 14$), α , γ , δ et ω sont les paramètres de lissage; y_t est la valeur réelle de la série chronologique pour la période t ; et $\hat{y}_t(k)$ est la prévision pour k -étape à l'avance. Le terme impliquant le paramètre ϕ dans l'expression Eq. (4.5) est un simple ajustement pour l'autocorrélation d'ordre première.

Dans ce chapitre, les paramètres de la méthode de HWT sont estimés en minimisant l'erreur pour la prévision de la pointe du soir en trois étapes principales (Figure 4.2). Tout d'abord, nous avons généré 2000 vecteurs de paramètres à partir d'un générateur de nombre aléatoire entre 0 et 1. Pour chacun des vecteurs, nous avons ensuite évalué l'erreur absolue en pourcentage pour la prévision de la demande de pointe du 24 Juin 2013. Les 200 vecteurs produisant les erreurs les plus bas sont évalués à leur tour, au cours de la période entre 24 Juin 2013 et 30 Juin 2013. Parmi les 200 vecteurs, les 10 vecteurs qui ont les plus faibles valeurs d'erreur absolue moyenne en pourcentage sont évalués pour une dernière fois pour prévoir les pics de charge de la période entre 01 Juin 2013 et 30 Juin 2013. Le vecteur produisant la valeur la plus basse de MAPE a été choisi comme le vecteur de paramètre final. Les valeurs optimisées des paramètres de la méthode de HWT étaient : $\alpha = 0.189$, $\gamma = 0.049$, $\delta = 0.304$, $\omega = 0.112$, et $\phi = 0.660$.

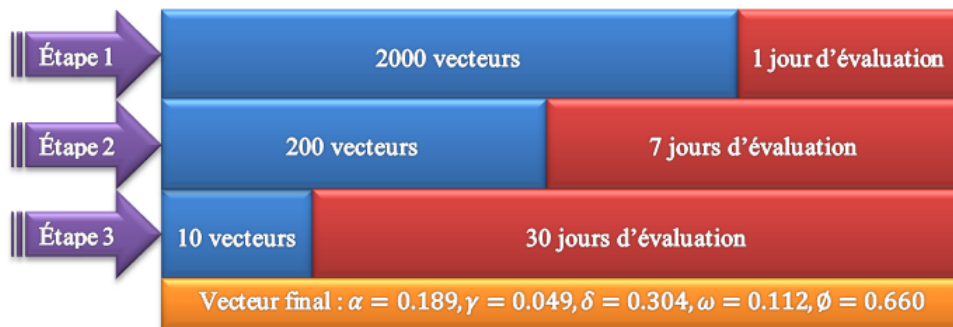


FIGURE 4.2 – Démarche suivie pour optimiser les paramètres de la méthode de HWT

4.3.2 Classification par la méthode des c-moyennes floues

Les conditions météorologiques ont des répercussions majeures sur la consommation d'électricité, surtout pour les deux catégories des consommateurs domes-

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁶⁸

tiques et agricoles. Le facteur météorologique le plus important pour la prévision à court-terme de la charge électrique est la température. Les changements affectent considérablement l'exigence de charge pour des fins de chauffage en hiver et des buts de climatisation en été. En Algérie, la forte augmentation de la demande sur le réseau national, sous l'impulsion de la hausse des températures en été, est une conséquence directe du changement des habitudes de consommation, notamment du fait de l'utilisation de plus en plus répandue de la climatisation. En hiver, la consommation d'électricité atteint ses plus hauts niveaux le soir, quand tous les foyers sont rassemblés chez eux et allument simultanément, éclairage, chauffages électriques et autres appareils ménagers (téléviseurs, PC, lave-linges,... etc).

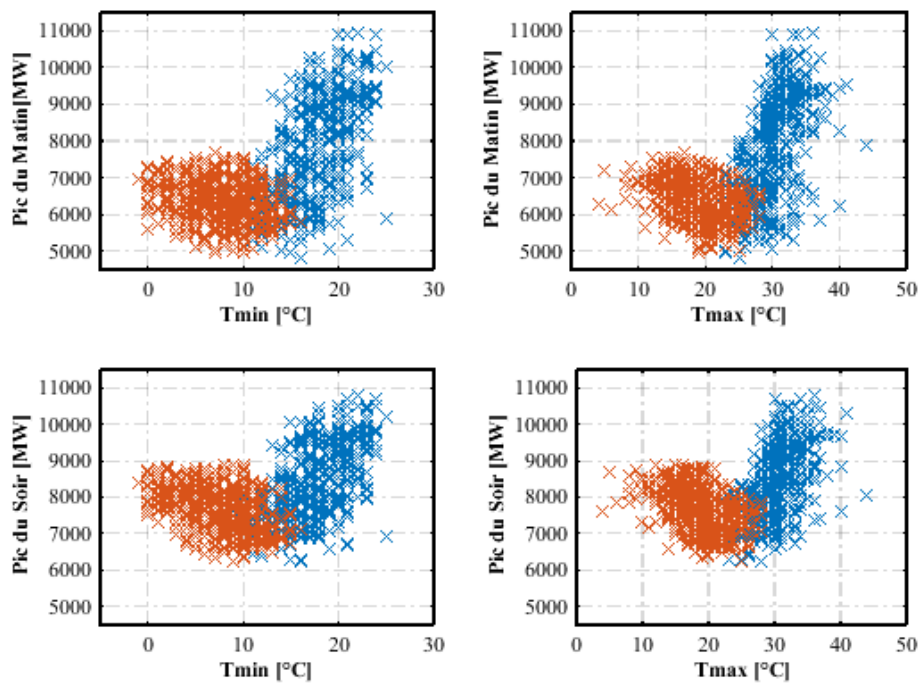


FIGURE 4.3 – Regroupement c-moyennes floues : résultats pour le 30 Novembre 2014

La Figure 4.3 montre les pointes de la demande d'électricité en Algérie pour la période du matin (ou du soir), en fonction des valeurs de température maximale (ou minimale) issues d'une station météorologique située à Alger [81]. La figure montre que la relation existante entre la température et la consommation d'électricité est très complexe et hautement non linéaire. Cependant, on peut voir que pendant l'été, l'augmentation de la température se traduira par une augmentation de la charge. Tandis qu'en hiver, la diminution de la température aura pour résultat une augmentation de la charge électrique. En fait, nous n'examinons pas dans la méthodologie proposée de modéliser la relation existante entre la température et la charge, mais nous essayons seulement de distinguer si le climat pour un jour donné

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁶⁹

est chaud ou froid. Pour ce faire, FCM est appliqué pour permettre aux données de température d'être classées en deux groupes liés à des périodes chauds ou froids de l'année.

1. Soit $X = (x_1, x_2, \dots, x_n)$ la collection des individus à classer
2. Fixer les paramètres
 - m : coefficient flou
 - c : le nombre de classes
 - ε : critère d'arrêt
3. Initialiser le vecteur V par c centres aléatoirement choisis
4. Calculer la matrice U de taille $c \times n$ par les équations :
$$u_{ij} = \frac{\mathbf{1}}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$
5. Calculer le nouveau centre de chaque classe à l'aide de l'équation :
$$V_i = \frac{\sum_{i=1}^n (u_{ij})^m x_i}{\sum_{i=1}^n (u_{ij})^m}$$
6. Mettre à jour la matrice U et incrémenter le compteur t
7. Calculer la distance entre les nouveaux et les anciens centres par :
$$h = \|V^{t-1} - V^t\|$$
8. Répéter les étapes de 3 à 6 tant que $h > \varepsilon$

FIGURE 4.4 – Algorithme de classification par c-moyennes floues

La méthode de regroupement en c-moyennes floues est une technique de classification de données qui permet à un point d'appartenir à un ou plusieurs groupes. La méthode a été développée par Dunn en 1973 [82] et améliorée par Bezdek en 1981 [83], et elle est fréquemment utilisée dans la reconnaissance des formes. L'algorithme de FCM tente de partitionner une collection finie de n éléments $X = x_1, \dots, x_n$ dans une collection de c classes floues par rapport à un critère donné. L'algorithme calcule à chaque fois les centres des classes et génère la matrice d'appartenance U des individus à ces classes. Étant donné un ensemble fini de données de température, l'algorithme renvoie une liste de deux centroïdes $V = \{c_1, c_2\}$ et une matrice de composition $U = u_{ij} \in [0, 1]$, $i = 1, \dots, n$; $j = 1, 2$ où chaque élément u_{ij} indique le degré auquel l'élément x_i appartient à la classe j . FCM vise à minimiser une fonction objective :

4.3. Méthodologie proposée pour la prévision de la pointe journalière 70

$$J_m = \sum_i^n \sum_j^c u_{ij}^m \|x_i - c_j\|^2, \quad (4.6)$$

Où m est un nombre réel supérieur à 1. $D_{ij} = \|x_i - c_j\|$ est une métrique définissant la distance entre l'élément x_i et le centroïde c_j . $u_{i,j}$ est le degré d'appartenance de x_i à la classe j :

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (4.7)$$

Nous donnons dans la Figure 4.4 un aperçu de l'algorithme qui peut être utilisé pour la classification par la méthode de c-moyennes floues. Dans notre cas, l'algorithme est utilisé pour permettre aux données de température d'être classées en deux groupes. Cependant, une fois que la classification des données de température est effectuée, il est facile à ce moment-là de classifier les données de charge en tenant compte les charges de pointe partageant les même index avec les valeurs de températures dans la base de données originale. Dans un but de clarification, nous voulons dire que les deux couleurs différentes apparues dans la Figure 4.3 sont liées aux deux groupes effectués avec le regroupement de FCM.

4.3.3 Algorithme des k-plus proches voisins

L'algorithme des k -plus proches voisins est une méthode très simple et performante pour les tâches de classement et de régression basée sur la ressemblance des échantillons d'un ensemble. La technique a été décrite dans les années 1950 et plus tard formalisée en 1960, lorsque des moyens de calcul puissants sont devenus disponibles. Le principe de l'algorithme est particulièrement très simple et directe. Il ne nécessite pas d'apprentissage mais simplement le stockage des données. Son principe est le suivant : une nouvelle donnée est comparée à toutes les données stockées, ensuite, on choisit pour la nouvelle donnée les k plus proches voisins au sens d'une distance choisie.

Il existe donc trois éléments clés pour cette approche : un ensemble d'enregistrements mémorisés ; une distance métrique pour calculer les distances entre les enregistrements ; et la valeur k , le nombre de voisins les plus proches à extraire. Pour un objet x non marqué, l'algorithme calcule la distance d (par exemple la distance Chebyshev) entre x et tous les objets y de l'ensemble d'entraînement afin de déterminer la liste des k -plus proches voisins sur l'ensemble de n éléments d'entraînement :

$$d_i = \max_i |x - y_i|, i = 1, \dots, n \quad (4.8)$$

Une fois la distance entre l'objet x et tous les échantillons de l'ensemble d'entraînement est calculée, les distances sont alors triées pour établir la liste des valeurs les plus proches en basant sur les k distances minimales. La Figure 4.5 résume l'algorithme utilisé pour extraire les k plus proches voisin de k -NN.

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁷¹

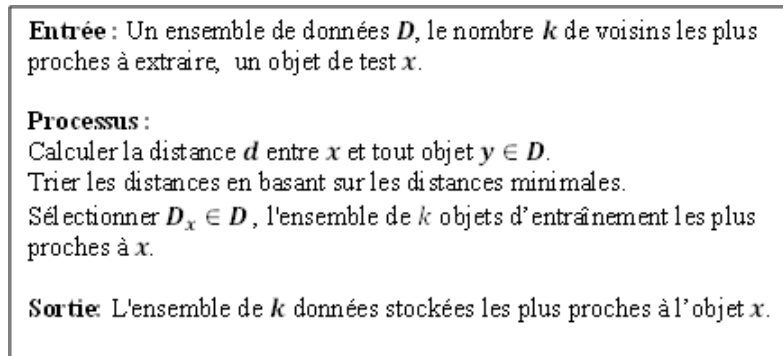


FIGURE 4.5 – Algorithme des k -plus proches voisins

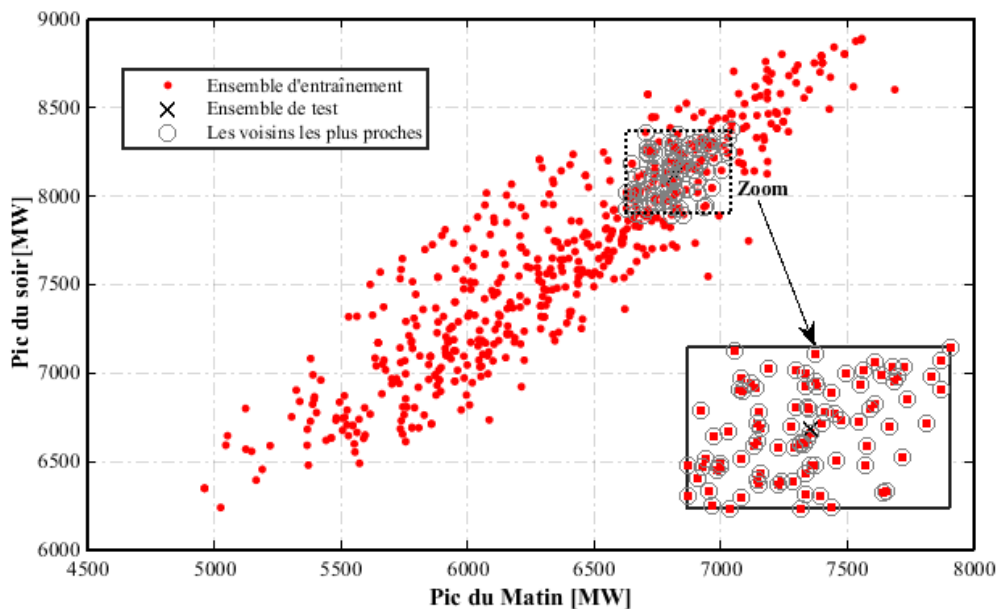


FIGURE 4.6 – Algorithme des k -plus proches voisins : résultats pour le 30 Novembre 2014

Pour la méthode k -NN proposée dans le présent chapitre, chaque individu est défini par les deux valeurs de charge de pointe du matin et du soir correspondant à un jour donné J_i . Ainsi, l'algorithme des k -NN tente de trouver des paires qui sont semblables à la valeur réelle de la demande de pointe du matin et la prévision primaire de la charge de pointe du soir pour le jour J_L . Par conséquent, l'algorithme permet l'identification des profils de charge, et par la manière de distinguer généralement si le jour J_L est un jour ouvrable ou un week-end. Un exemple des résultats de l'algorithme des k -plus proches voisins est présenté sur la Figure 4.6. Le point

4.3. Méthodologie proposée pour la prévision de la pointe journalière 72

d'être classé, montré avec "X" ; est lié aux données de la demande de pointe pour le 30 Novembre 2014. Cette figure montre que les classificateurs de données ne permettent pas seulement de réduire la dimension du problème, mais aussi d'améliorer la qualité des données à prendre en considération par la décomposition en paquets d'ondelettes et le système d'inférence neuro-flou adaptatif.

4.3.4 Décomposition en paquets d'ondelettes

La transformée en ondelettes (WT) est un outil puissant et flexible qui a été largement utilisé dans le domaine de la prévision de la charge électrique pour décomposer la série de charge à un ensemble de séries constitutives que l'on peut prédire avec plus de précision en comparaison avec le signal original [84], [85], [86]. La transformée en ondelettes est capable de révéler les points de rupture, les discontinuités, et les aspects d'autosimilarité de données [84]. WT peut être divisée en deux catégories : transformée en ondelettes continue (CWT), et transformée en ondelettes discrète (DWT).

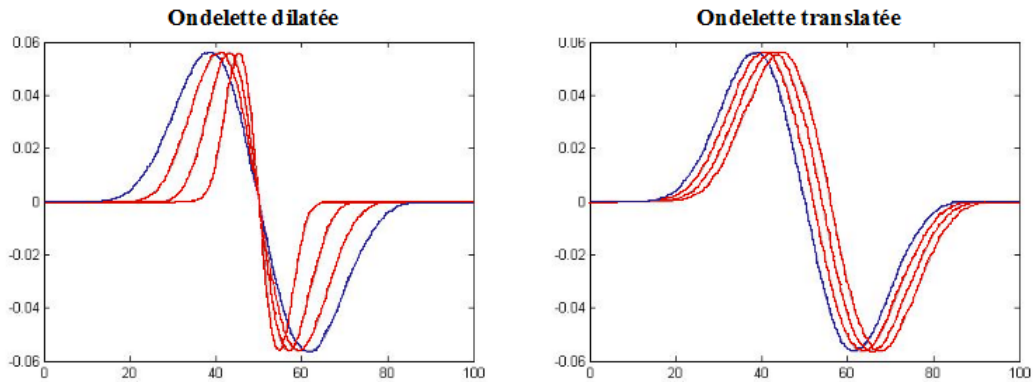


FIGURE 4.7 – Dilatation et translation d'une ondelette mère

Dans la transformée en ondelettes continue, chaque ondelette est créée par des opérations de translation et de dilatation d'une ondelette mère (Figure 4.7). L'ondelette mère est une fonction oscillante avec énergie finie et moyenne nulle. Il existe de nombreuses ondelettes mères possibles. Cependant, les familles d'ondelettes habituellement utilisables sont présentées dans la Figure 4.8.

La transformée en ondelettes continue d'un signal de temps $f(t)$ est défini comme suit :

$$CWT(a, b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt \quad (4.9)$$

Où ψ est l'ondelette mère, $a \in \mathbb{R}$ est un paramètre de dilatation, $b \in \mathbb{R}$ est un paramètre de translation, et :

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁷³

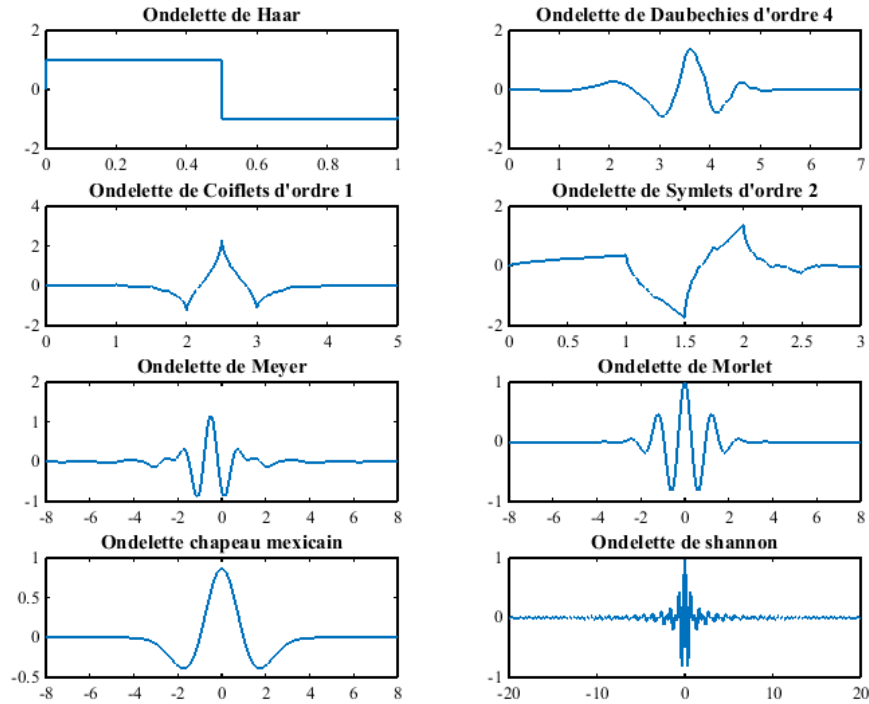


FIGURE 4.8 – Ondelettes mères couramment utilisées

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (4.10)$$

Cependant, le CWT est difficile à mettre en pratique et l'information ramassée peut se chevaucher et aboutir à une redondance. Il est donc évident qu'une discrétisation de la transformation devrait être envisagée si l'on souhaite obtenir une transformation non redondante [87]. Dans l'analyse multirésolution basé sur la DWT, le signal S est passé à travers des filtres passe-haut et passe-bas en utilisant seulement certaines échelles et positions. Le filtre passe-haut donne la résolution détaillée du signal (détail), tandis que le filtre passe-bas donne l'information approximative du signal (approximation). Une approximation (A) maintient la tendance générale d'un signal donné, alors qu'un détail (D) représente les composantes hautes fréquences de celui-ci.

En définissant $a = a_0^c, b = db_0 a_0^c$; avec a et $b \in \mathbb{Z}$, la DWT du signal échantillonné $f(t)$ de longueur T est donnée par :

$$DWT(c, d) = \sum f(t) \psi_{c,d}(t) \quad (4.11)$$

4.3. Méthodologie proposée pour la prévision de la pointe journalière74

Où :

$$\psi_{c,d}(t) = \frac{1}{\sqrt{a_0^c}} \psi\left(\frac{t - db_0 a_0^c}{a_0^c}\right) \quad (4.12)$$

L'analyse en paquets d'ondelettes est une extension de l'analyse multirésolution par ondelettes qui offre une meilleure résolution de fréquence et une analyse du signal plus riche. Le processus de WPD par filtres est similaire à celui de DWT. La différence entre la décomposition en paquet d'ondelettes et la décomposition en ondelettes est que seules les composantes basses fréquences sont divisées en deux parties de façon continue dans l'analyse multirésolution, alors que les coefficients du détail et d'approximation sont décomposés à la fois dans WPD. Cet attribut rend la WPD très appropriée dans notre cas d'étude pour la prédiction des pics de consommation d'électricité, étant donné que des informations pertinentes peuvent être situées dans les composantes haute fréquence des données de charge électrique.

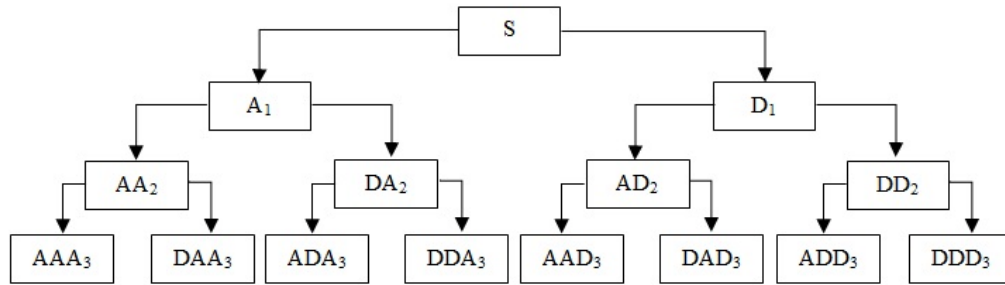


FIGURE 4.9 – Arbre en paquets d'ondelettes de trois niveaux

Pour une meilleure compréhension du processus de décomposition en paquets d'ondelettes, un arbre de WPD d'ordre trois est représenté sur la Figure 4.9. Dans la décomposition en paquets d'ondelettes, le signal original S est tout d'abord décomposé en détail (D_1) et approximation (A_1). Dans le deuxième niveau, non seulement la partie d'approximation est décomposée à nouveau en détail (DA_2) et approximation (AA_2), mais le vecteur de coefficients de détail (D_1) est également décomposé en deux parties (DD_2 , AD_2) en utilisant la même démarche dans le fractionnement du vecteur d'approximation. Ce processus est répété d'une manière similaire dans le troisième niveau de l'arbre de paquets d'ondelettes. Cependant, il n'y a pas une manière unique pour reconstituer le signal initial S à partir de ses coefficients d'ondelettes. Pour le modèle proposé dans le présent chapitre, les nœuds suivants dans l'arbre de paquets d'ondelettes à trois niveaux sont sélectionnés : $\{AAA_3, DAA_3, ADA_3, DDA_3, AD_2, DD_2\}$. L'ondelette mère utilisée est la Daubechies d'ordre 2 (en abrégé db2). Un exemple des résultats de la décomposition en paquets d'ondelettes est représenté sur la Figure 4.10, où nous avons exposé les sous-séries obtenues à partir du processus de prévision du 30 Novembre 2014.

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁷⁵

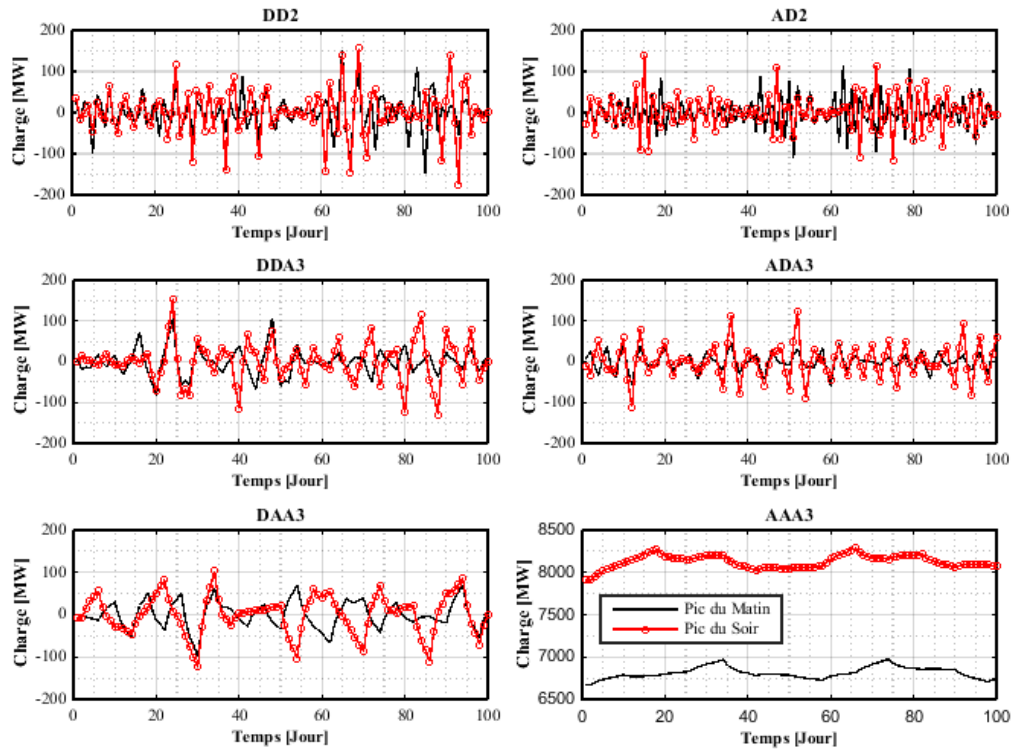


FIGURE 4.10 – Résultats de WPD pour la prévision de la charge de pointe du 30 Novembre 2014

4.3.5 Système d'inférence neuro-flou adaptatif

A cette partie de la méthode de prévision proposée, l'ANFIS est introduit pour prédire les futures tendances de chaque espace d'ondelettes. Cependant, comme il y a six nœuds de l'arbre en paquets d'ondelettes, nous atteignons éventuellement six modèles ANFIS séparés. Pour chaque modèle, l'entrée est associée à la partie d'approximation ou de détail du signal de la pointe de consommation pour la période du matin. Alors que la sortie est associée à la partie d'approximation ou de détail de la série de charge du soir. La structure proposée du système d'inférence neuro-flou adaptatif utilise la fonction d'appartenance généralisée en forme de cloche, deux règles floues, et deux cents époques pour identifier les paramètres du FIS et la topologie optimale du système d'inférence neuro-flou adaptatif.

4.3.6 Résultat final pour la prévision de la pointe journalière

Le résultat final de la prévision de la pointe de charge est produit au moyen d'additionner tous les composants prévus de chaque espace d'ondelettes.

4.3. Méthodologie proposée pour la prévision de la pointe journalière⁷⁷

Tableau 4.2 – L'erreur relative (MAPE) de toutes les méthodes proposées

Méthode	Prévisions primaires						Prévisions secondaires					
	HWT	P-BPNN	P-NARX	P-ANFIS	S-BPNN	S-NARX	S-ANFIS	FKW-BPNN	FKW-NARX	FKW-ANFIS		
Janvier	1.872	2.021	2.535	1.985	1.889	2.187	1.719	2.000	1.553	1.717		
Avril	1.465	1.693	1.998	1.985	1.273	1.934	1.301	1.382	1.353	1.434		
Juillet	2.848	3.143	2.828	2.845	2.360	2.630	2.571	2.157	2.908	2.238		
Novembre	1.329	2.034	2.486	1.742	1.655	2.164	1.610	1.335	1.704	1.176		
tout	1.879	2.223	2.462	2.048	1.794	2.229	1.800	1.719	1.879	1.641		

4.4 Evaluation de la méthodologie proposée pour la prévision des pics quotidiens de charge

Dans cette section, nous évaluons la précision des prévisions de notre méthodologie constituée de deux étapes (HWT et FKW-ANFIS). Nous comparons également leur performance contre les prévisions primaires du réseau de neurones avec rétro-propagation de l'erreur (P-BPNN) ; réseau neuronal non linéaire autorégressif avec entrée exogène (P-NARX) ; et le système d'inférence neuro-flou adaptatif (P-ANFIS), et aussi contre les prévisions secondaires de S-BPNN ; S-NARX ; S-ANFIS ; FKW-BPNN ; et FKW-NARX. Nous avons choisi de mettre en œuvre le réseau neuronal autorégressif non linéaire avec entrée exogène dans notre étude parce que nous avons été séduits par la structure de ce réseau : la sortie dépend non seulement des valeurs précédentes des intrants, mais aussi sur les valeurs précédentes du signal de sortie (Figure 4.11).

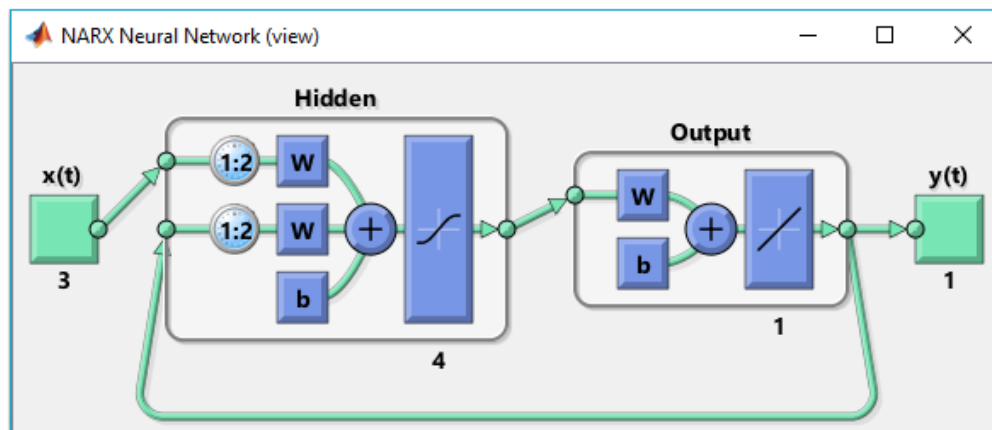


FIGURE 4.11 – Architecture du réseau neuronal non linéaire autorégressif avec entrée exogène

Nous avons également opté la mise en œuvre du système d'inférence neuro-floue adaptative seul et le perceptron multicouche avec rétro-propagation de l'erreur, parce que ces deux modèles ont été intensivement utilisés pour la prévision à court-terme de la charge électrique [88],[89].

Les données de la pointe du soir pour le jour précédent (par exemple le jour J_{L-1}) et les données du même type de jour pour les deux semaines avant la journée à prévoir (par exemple jours J_{L-7} et J_{L-14}) sont adoptées comme intrants pour les modèles d'intelligence computationnelle primaires considérés dans le présent chapitre. La température minimale, la température maximale, et la charge du matin pour le jour J_L sont utilisés comme intrants pour S-BPNN, S-NARX, et S-ANFIS. L'architecture de FKW-BPNN et FKW-NARX est très similaire à celle de FKW-ANFIS, où seule la pointe du matin de la journée pour laquelle nous tenons à prédire la charge ; est

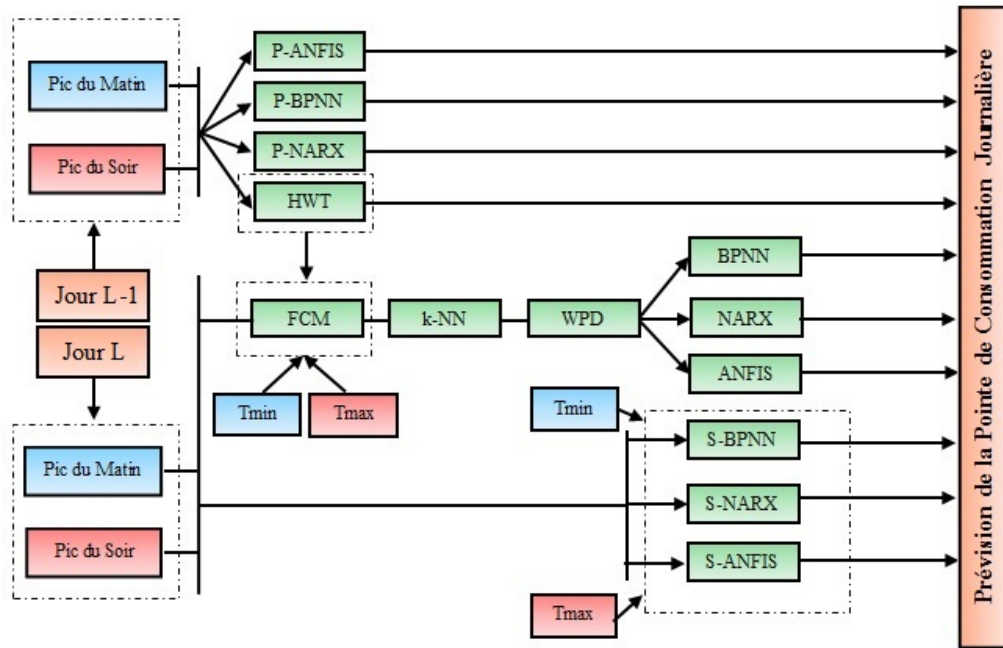


FIGURE 4.12 – Structure générale des méthodes proposées

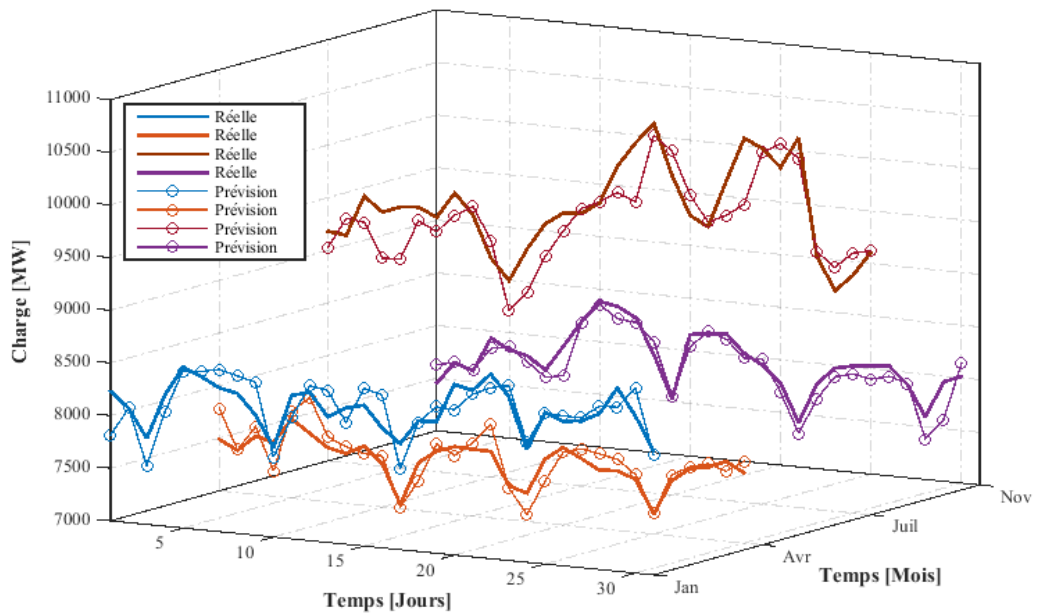


FIGURE 4.13 – Résultats de FKW-ANFIS pour quatre mois différents : Janvier, Avril, Juillet et Novembre 2014

utilisée comme entrée à la phase finale pour les réseaux de neurones. Les éléments et les paramètres principaux de ces approches sont indiqués sur le Tableau 4.1 et schématisés sur la Figure 4.12. Les blancs dans le Tableau 4.1 signifient que l'entrée ou le paramètre concerné n'est pas inclus par un modèle de prévision spécifique, alors que "*" désigne l'inclusion de l'entrée/paramètre par le modèle de prévision.

La Figure 4.13 montre les résultats de FKW-ANFIS pour la prévision de la demande de pointe du soir le long de quatre mois différents : janvier 2014, avril 2014, juillet 2014 et novembre 2014. Cette figure montre que le modèle FKW-ANFIS proposé a réussi à prédire la demande quotidienne de pointe, étant donné que l'allure des courbes prédites de la charge de pointe suite dans la majorité des cas la forme des courbes réelles de la demande de pointe. Cependant, les résultats de prévision détaillés ; comme il est indiqué dans le Tableau 4.2 ; démontrent clairement l'efficacité de la stratégie de prévision proposée et montrent que FKW-ANFIS a surpassé toutes les autres modèles. Les résultats de MAPE illustrent que la méthode de HWT est la plus précise parmi les modèles de prévision primaires, tandis que la meilleure technique parmi les modèles de prévision secondaires est FKW-ANFIS. Les résultats montrent également l'existence d'une amélioration notable de la précision des prévisions lors de l'utilisation du modèle d'ANFIS secondaire plutôt qu'en utilisant la technique à double saisonnière de HWT, et montrent aussi l'existence d'une autre amélioration plus perceptibles qui a été obtenue en utilisant le modèle FKW-ANFIS. D'après ces résultats, nous pouvons remarquer que FKW-ANFIS a une précision 08.838% et 12.678% plus élevée par rapport à S-ANFIS et HWT, respectivement. L'amélioration la plus importante a été enregistrée en été, avec une valeur de MAPE inférieure respectivement de 12.958% et 21.418%. Ceci est un point très important, puisque nous savons que la consommation d'électricité atteint ces plus hauts niveaux en été à la suite d'une énorme utilisation des appareils de climatisation. Dans le reste de cette section, nous nous intéressons à d'autres comparaisons entre les deux composantes de notre méthodologie : HWT et FKW-ANFIS.

Nous présentons dans la Figure 4.14 les résultats de MAPE pour les quatre mois évoqué ci-dessus, et consolidés en fonction du type de jour. Cette figure montre que HWT présente les valeurs de MAPE minimales pour les vendredis et les samedis, et FKW-ANFIS présente les valeurs de MAPE minimales pour les jours ouvrables. Pour les lundis, la valeur MAPE obtenue par FKW-ANFIS est un peu autour de la moitié de la valeur obtenue par HWT.

Nous présentons dans le Tableau 4.3 les résultats des prévisions de HWT et FKW-ANFIS pour les jours fériés de l'année 2014. Le tableau montre que la valeur MAPE de HWT et FKW-ANFIS est respectivement 3.283% et 2.796%, ce qui signifie que le modèle FKW-ANFIS à une précision 14.834% plus élevée par rapport à HWT. Par conséquent, le modèle FKW-ANFIS proposé dans le présent chapitre est un moyen efficace pour améliorer la précision de prévision, et il peut servir, ensemble avec la méthode de HWT, comme un outil très approprié pour la prévision de la pointe journalière de la demande d'électricité durant à la fois les conditions normales et spéciales de la charge électrique.

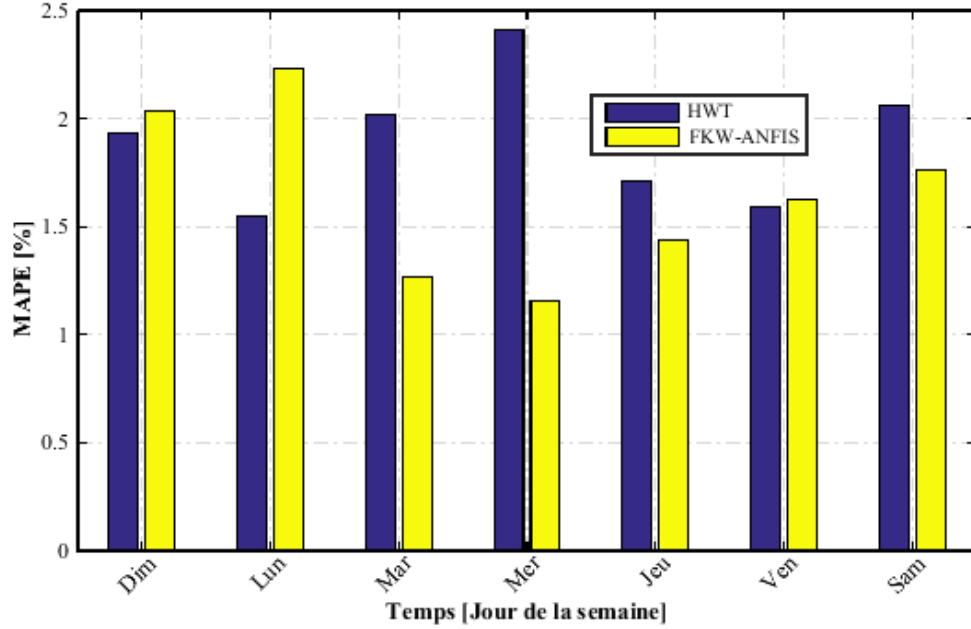


FIGURE 4.14 – Résultats en MAPE de HWT et FKW-ANFIS pour chaque jour de la semaine

Tableau 4.3 – Résultats obtenus lorsque HWT et FKW-ANFIS sont utilisés pour prévoir la charge de pointe des jours fériés en 2014

Résultats	pic réel [MW]	HWT				FKW-ANFIS			
		Pic prévu [MW]	Erreur [MW]	APE [%]	MAPE [%]	Pic prévu [MW]	Erreur [MW]	APE [%]	MAPE [%]
01/01/2014	8237	8709	472	5.730	7813	-424	5.147		
14/01/2014	8294	8127	-167	2.013	8156	-138	1.664		
19/03/2014	7795	7668	-127	1.629	7741	-054	0.693		
01/05/2014	7400	7471	071	0.959	7475	075	1.013		
29/07/2014	9093	9570	477	5.246	9316	223	2.452	2.796	
30/07/2014	9258	9198	-060	0.648	9470	212	2.290		
05/10/2014	7292	7205	-087	1.193	6910	-382	5.239		
06/10/2014	8031	7061	-970	12.078	7505	-526	6.550		
03/11/2014	7602	7598	-004	0.053	7611	009	0.118		

4.5 Conclusion

Ce chapitre a abordé le problème de la prévision des pics quotidiens de la charge à l'aide d'une structure hybride et adaptative à deux étages. Cette structure est composée principalement de deux modèles de base : un modèle primaire de HWT, et un modèle secondaire de FKW-ANFIS. Les résultats et les performances de prévision obtenues ont montré que : (1) la valeur moyenne de MAPE du modèle de prévision primaire de HWT et du modèle secondaire FKW-ANFIS dans les jours normaux est 1.879% et 1.641%, respectivement, et (2) la valeur de MAPE moyenne du modèle de HWT et de FKW-ANFIS dans les jours fériés est 3.283% et 2.796%, respectivement. En outre, le modèle final de FKW-ANFIS a montré de meilleurs résultats de prévision en comparaison avec d'autres modèles fondés sur les réseaux de neurones artificiels et le système d'inférence neuro-flou adaptatif seul. Par conséquent, la méthodologie en deux étapes proposée dans le présent chapitre est un outil très efficace pour améliorer la précision de prévision de la demande de pointe d'électricité durant à la fois, les conditions normales et spéciales de la charge électrique.

Prévision en ligne basée sur une méthodologie efficace de combinaison de prédictions

5.1 Introduction

Ce chapitre présente une nouvelle méthodologie combinée pour assurer les prévisions en ligne, et à très court-terme, de la demande d'électricité. Ces prévisions de quelques minutes à l'avance sont importantes pour la planification en temps réel de la production d'électricité, ainsi que pour le contrôle de fréquence et les fonctions de dispatching économiques. Cependant, les grandes erreurs de prévision peuvent grandement affecter les niveaux de production, la sûreté et la qualité du fonctionnement du réseau électrique. Un compromis performance-temps de calcul faible est nécessaire donc pour assurer le fonctionnement économique et fiable du réseau électrique. A cet effet, nous proposons dans cette étude une méthodologie combinée, dont l'idée de base consiste à utiliser un système de prévision de charge qui a la capacité d'obtenir une bonne précision de prévision, assurer un temps de calcul faible, et éviter les grandes erreurs de prédiction. Les données de charge en temps-réel du système électrique français et l'ensemble de données australiennes pour l'Etat de la Nouvelle-Galles du Sud sont utilisées comme un exemple illustratif pour évaluer la performance de la méthode proposée. Les résultats de prévision de cette démarche seront également comparés avec les résultats issus d'un certain nombre d'algorithmes de référence et méthodes de combinaison de prévisions.

5.2 Analyse et traitement des séries chronologiques de la demande d'électricité

Pour évaluer la performance de notre approche, nous utilisons les données de charge électrique provenant de deux pays différents (l'Australie et la France), avec des intervalles de temps aussi différents (5 minutes et 15 minutes). La base de données pour la France est recueillie à partir du site web de l'opérateur du réseau de transport d'électricité [90]. Elle couvre la période du 07 Avril 2013 au 31 Janvier

2015 avec un intervalle d'enregistrement de 15 minutes. La base de données australienne est pour l'Etat de la Nouvelle Galles du Sud, fournie par l'opérateur du marché de l'énergie en Australie (AEMO) [91]. Les données sont pour deux ans, 2006 et 2007, et sont mesurées toutes les 5 minutes. Toutefois, étant donné que nous avons choisi d'utiliser les informations de la demande d'électricité comme unique condition d'entrée pour développer un algorithme approprié pour la prévision en ligne de la charge électrique, il est donc de première importance d'examiner les principales caractéristiques de la série de charge et de comprendre le comportement de la demande d'électricité.

Nous présentons dans la Figure 5.1, la consommation d'électricité en France pour la période du 01 Janvier 2014 au 31 Décembre 2014. Il ressort de cette figure que le comportement de la charge électrique est principalement influencé par des profils de saisonnalité quotidienne où la consommation d'électricité commence généralement avec des valeurs faibles au début de matinée et augmente de façon significative jusqu'à atteindre une charge de pointe journalière. On peut percevoir également l'existence d'une saisonnalité annuelle qui se caractérise par une demande de pointe annuelle en hiver et une creuse en été.

Nous présentons dans la Figure 5.2 la charge de pointe quotidienne en France pendant quatre semaines différentes. Il ressort de cette figure que le comportement de la charge électrique montre un cycle saisonnier hebdomadaire comprenant les jours ouvrables avec des profils de demande similaires et le week-end, qui se caractérise par une diminution de la consommation.

Des méthodes classiques univariées telles que les modèles de séries temporelles de Box-Jenkins et les méthodes de lissage exponentiel sont souvent considérées comme suffisantes pour la modélisation des caractéristiques saisonnières des séries de charge électrique. Cependant, il semble que l'inclusion du cycle annuel dans une démarche de prévision ne conduira pas à une amélioration de la performance pour des prédictions aussi courtes qu'une heure à l'avance. À cet égard, notre choix est tombé sur la considération du processus saisonnier autorégressif à moyenne mobile intégré et le modèle à double saisonnalité de Holt-Winters-Taylor en tant que techniques efficaces pour modéliser les cycles journaliers et hebdomadaires figurant dans les données de la demande d'électricité.

Comme conséquence d'ignorer les jours spéciaux au cours du processus de modélisation, l'inconvénient des deux approches séquentielles citées précédemment se trouve sur son incapacité d'obtenir des résultats satisfaisants lorsqu'ils sont soumis à des conditions anormales de charge tel que le cas de la diminution rapide de la demande pendant les vacances. Pour obvier à cet inconvénient, nous proposons une méthodologie parallèle pour générer les prévisions à très court-terme de la demande pour à la fois les conditions normales et spéciales de la charge électrique. Dans cette approche, la série temporelle de la consommation d'électricité est séparée en points autonomes.

Un ensemble de points indépendants signifie que la charge à chaque quart d'heure (ou cinq minutes pour les données australiennes) de la journée est indépendante de la charge reliée à un autre quart d'heure (ou cinq minutes). Notre approche parallèle

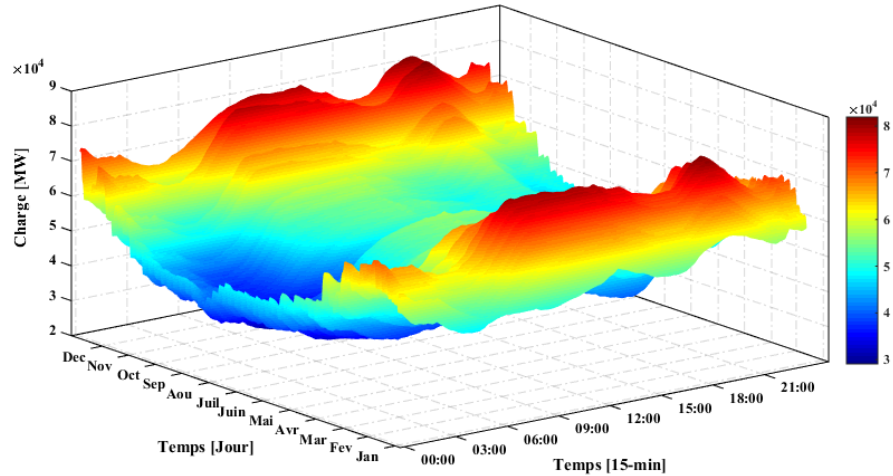


FIGURE 5.1 – Les données de la demande d'électricité du système électrique français : 01 Jan-31 Dec 2014

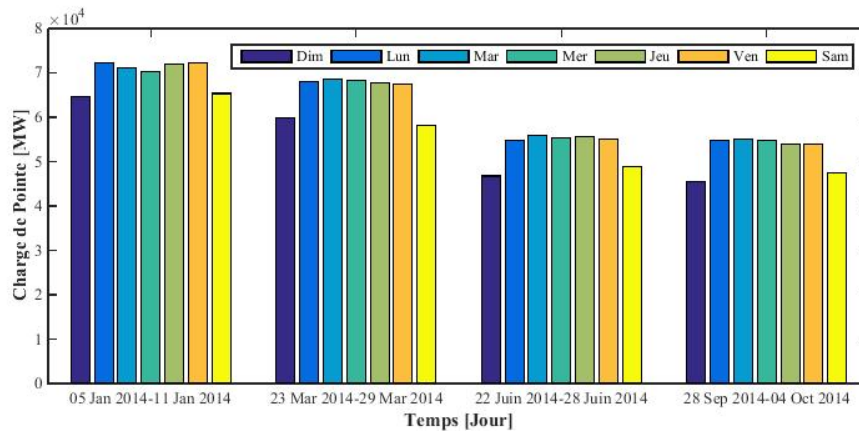


FIGURE 5.2 – La pointe journalière d'électricité en France durant quatre semaines différentes

considère deux cas principaux. Pour les jours normaux, les séries parallèles de charge sont classés par calendrier selon le type de jour : les samedis, les dimanches, et les jours ouvrables. Dans des travaux effectués au cours de la préparation de notre thèse de doctorat, cette méthode d'arrangement a montré des résultats satisfaisants pour la prévision à très court-terme de la charge durant des jours normaux [92], [93], [94]. Cependant, la particularité du présent travail est située dans la prise en compte des conditions anormales de la charge électrique. La prise en compte d'une telle charge dans les travaux existants dans la littérature a été réalisée en utilisant un processus de modélisation complexe qui nécessite une quantité importante de données de charge et une intervention de la part de l'utilisateur est pratiquement essentielle pour restructurer le modèle. Contrairement à ces travaux existants, notre algorithme de modélisation des jours spéciaux est très simple et ne nécessite que peu de données récentes pour effectuer les prévisions. Par conséquent, chaque série parallèle dans notre processus de modélisation des charges atypiques contienne soit l'information de la demande d'électricité des dimanches précédents plus la charge de la journée de prévision (si la charge à prévoir est liée à un jour ouvrable), ou les informations de la demande d'électricité des samedis ou dimanches (si le jour à prévoir est un week-end). Les méthodes de prévision destinées à être utilisées dans une implémentation parallèle sont le réseau de neurones avec rétro-propagation de l'erreur (BPNN), le système d'inférence neuro-flou adaptatif (ANFIS), et une version adaptative de lissage exponentiel de Holt récemment développée (AHES). Pour ces modèles parallèles, on peut obtenir les prévisions pour quatre étapes à l'avance (ou douze pour les données australiennes) en effectuant une seule étape de prévision à l'avance sur chaque série parallèle liée à la période d'évaluation (une heure). Pour les méthodes basées sur l'utilisation d'une implémentation séquentielle, nous pouvons tout simplement faire des prévisions pour un horizon temporel k à partir de l'origine de prévision t ; où k peut atteindre 5 (ou 13 pour les données australiennes), et t est un temps lié à la dernière observation de la série temporelle de la charge électrique. L'horizon de prévision k est un terme allant jusqu'à 5 (ou 13), car pour une mise en œuvre en ligne, l'information de la demande d'électricité d'un temps actuel donné (notée $t + 1$) n'est pas disponible au moment de la prévision.

5.3 Méthodologie proposée pour la prévision en ligne de la charge électrique

La structure de la méthodologie proposée pour la prévision à très court-terme de charge (abrégée HFCM), se compose de trois étapes principales [95]. À la première partie, plusieurs modèles sont utilisés pour effectuer des prévisions primaires de la charge électrique pour une heure à l'avance. Puisque la série chronologique de la demande d'électricité connaît de multiples variations, il est de première importance de prendre en compte plusieurs hypothèses et méthodes pour traiter ces variations et capter la tendance récente de la demande d'électricité. Pour cette raison, deux hypothèses ont été considérées dans notre méthode. La première est basée

Tableau 5.1 – Les principaux éléments et paramètres des différentes méthodes considérées

Méthodes Éléments	Méthodes classiques		Méthodes d'intelligence artificielle		Méthodes combinées			
	SARIMA	HWT	AHES	BPNN	ANFIS	SFCM	HFCM-WAM	HFCM-TM
Modèle mathématique	Eq.3.10	Eq.4.1-4.5	Eq.3.1-3.3	Non requis	Non requis	Nécessaire dans certaines parties	Nécessaire dans certaines parties	Nécessaire dans certaines parties
Détection des sure/sous-estimation						*	*	*
Moyenne simple						*	*	*
Moyenne pondérée							*	
moyenne tronqué								*
Approche séquentielle	*	*				*	*	*
Approche parallèle			*	*	*	*	*	*
Modélisation des jours spéciaux			*	*	*	*	*	*

sur l'utilisation d'une approche séquentielle et la seconde repose sur l'utilisation d'une approche parallèle. De plus, un compromis entre les méthodes classiques et de l'intelligence artificielle a été organisé. De ce fait, deux méthodes classiques (HWT et SARIMA) et deux modèles d'intelligence artificielle (BPNN et ANFIS) ont été choisis pour faire partie des modèles primaires de prévision. L'inclusion de ces modèles est basée sur leur précision raisonnable, la simplicité d'utilisation, et le temps d'exécution acceptable. Le lissage exponentiel de Holt-Winters et le processus ARIMA saisonnier sont, sans doute, les méthodes classiques les plus couramment utilisées dans la littérature de la prévision de charge. Cette popularité est attribuée à la capacité de ces méthodes à gérer la composante saisonnière de la série chronologique de la demande d'électricité. D'un autre côté, les réseaux de neurones artificiels et les systèmes neuro-flous comptent parmi les méthodes de l'intelligence artificielle qui ont une grande capacité à modéliser la non-linéarité et la relation complexe qui existe entre la demande de charge et les autres facteurs exogènes. Un cinquième modèle individuel, nommé lissage exponentiel adaptative de Holt, est également inclus. Nous avons choisi d'utiliser cette cinquième méthode, car elle est une approche récente qui a montré des résultats prometteurs pour traiter les variations rapides de la demande d'électricité et assurer une très faible valeur d'erreur pour le cas de prévision à très court-terme de la charge électrique [92]. Une simple combinaison des prévisions issues des cinq modèles individuels est également incluse pour faire partie des modèles primaires de prévision. Pour éviter les grandes erreurs de prévision, un filtre de prétraitement de données est introduit à la deuxième partie de notre approche pour remplacer les sous- et sur-estimations par des valeurs normalisées. Ensuite, l'estimation finale de la charge est obtenue en considérant la moyenne tronquée des prévisions normalisées.

La structure générale de notre méthodologie proposée pour assurer la prévision en ligne de la demande de charge électrique est montrée dans la Figure 5.3, et les éléments et paramètres clés de toutes les méthodes considérées sont présentés dans le Tableau 5.1. Les cellules qui apparaissent vides dans le Tableau 5.1 signifient que l'élément ou le paramètre concerné n'est pas inclus pour un modèle de prévision spécifique, alors que «*» désigne l'inclusion de l'élément/paramètre par le modèle de prévision.

5.3.1 Modèles primaires de prévision

5.3.1.1 Processus saisonnier autorégressif à moyenne mobile intégrée

Comme nous l'avons vu dans le troisième chapitre, le processus saisonnier autorégressif à moyenne mobile intégré est une extension du processus autorégressif à moyenne mobile intégré qui a été mis en place pour modéliser des séries chronologiques avec des tendances, comportements cycliques et d'autres caractéristiques non stationnaires. Dans le présent cas, la composante de saisonnalité provient du cycle journalier de la série chronologique de la charge électrique. Les modèles $SARIMA(p, d, q) \times (P, D, Q)_s$ adoptés pour les données de la France et de

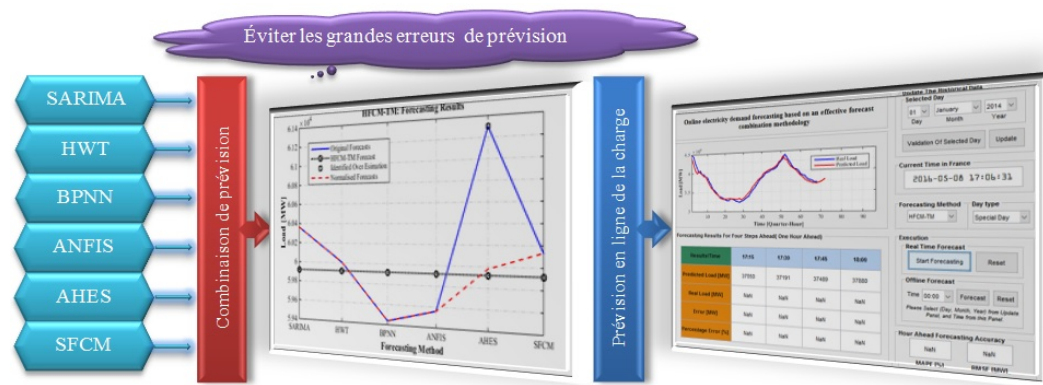


FIGURE 5.3 – Structure générale de la méthodologie proposée pour la prévision en ligne de charge

l'Australie peuvent être décrits comme SARIMA $(1, 1, 1) \times (2, 1, 1)_{96}$ et SARIMA $(3, 1, 3) \times (2, 1, 1)_{288}$, respectivement.

5.3.1.2 Méthode de Holt-Winters-Taylor

Comme nous l'avons vu dans le quatrième chapitre, la technique de HWT est une extension du lissage exponentiel de Holt-Winters standard, qui a été introduite pour permettre l'accumulation des cycles journaliers et hebdomadaires de la série de charge électrique. Dans le présent cas de prévision, le cycle journalier s_1 a une durée de 96 périodes pour les données de France, et une durée de 288 périodes pour les données de l'Australie. Pour le cycle hebdomadaire (s_2), ce cycle présente une durée de 672 périodes pour le cas des données de France et 2016 périodes pour les données de l'Australie.

Dans ce chapitre, les paramètres de la méthode de HWT sont estimés en tenant compte de l'erreur minimale de prévision pour trois étapes principales. Pour la série de charge française, nous avons généré en première étape 2000 vecteurs de paramètres à partir d'un générateur de nombres aléatoires entre 0 et 1. Pour chacun des vecteurs, nous évaluons alors l'erreur absolue en pourcentage pour la prévision de la pointe journalière du 02 Août 2014. Les 200 vecteurs produisant les APE les plus bas sont évalués; à leur tour, pour prédire la courbe de charge du 02 Août 2014 (96 valeurs de charge). Parmi les 200 vecteurs, les 10 vecteurs qui ont les plus faibles valeurs d'erreur absolue moyenne en pourcentage sont évalués à une dernière fois pour assurer une heure de prévision de charge le long de la période entre le 02 Août 2014 et le 08 Août 2014. Le seul produisant la valeur de MAPE la plus basse a été choisi comme vecteur de paramètre final. Les valeurs optimisées des paramètres de la méthode de HWT sont : $\alpha = 0.987$; $\gamma = 0.057$; $\delta = 0.006$; $\omega = 0.272$; et $\phi = 0.008$. Pour la série chronologique de la consommation d'électricité en Australie, les paramètres sont estimés par une manière semblable et les valeurs optimisées sont

les suivantes : $\alpha = 0.996$; $\gamma = 0.191$; $\delta = 0.273$; $\omega = 0.545$; et $\phi = 0.206$.

5.3.1.3 Réseau de neurones à rétro-propagation de l'erreur

En raison de leur capacité inhérente pour résoudre toute relation complexe et non linéaire entre les entrées et les sorties à travers un processus d'apprentissage, l'utilisation des modèles de réseaux neuronaux est de nos jours une méthode alternative pour résoudre le problème de prévision à très court-terme de la charge électrique [96], [97]. À cet égard, un réseau de neurones avec rétro-propagation de l'erreur est utilisé dans ce travail. La configuration du réseau de neurones proposé est très semblable à celle utilisée dans [94]. Notre modèle de réseau neuronal avec rétro-propagation de l'erreur consiste d'une couche cachée de quatre neurones sigmoïdes suivi par une couche de sortie du neurone linéaire. Les données de charge d'une série parallèle précédente sont adoptées comme un vecteur d'entrée, et les informations de charge pour le quart d'heure (ou les cinq minutes) à prévoir sont considérés comme la sortie du réseau. L'algorithme de Levenberg-Marquardt est conçu pour entraîner le réseau et la fonction de critère de performance est l'erreur quadratique moyenne. Le nombre maximum d'itérations pour entraîner le réseau est limitée à 20 époques.

5.3.1.4 Système d'inférence neuro-flou adaptatif

La structure du système d'inférence neuro-flou proposée utilise la fonction d'appartenance généralisée en forme de cloche, deux règles floues, et vingt itérations pour identifier les paramètres du système d'inférence floue et la topologie optimale de l'ANFIS. Les paires d'entrée/sortie sont similaires à celles du modèle du réseau de neurones à rétro-propagation de l'erreur.

5.3.1.5 Lissage exponentiel adaptatif de Holt

AHES est une technique adaptative qui a été introduite dans [92] pour assurer la prévision de charge pour une seule étape à venir. La méthode est adaptative à la manière dont les paramètres de lissage sont changés à chaque fois qu'une nouvelle observation est enregistrée. Il est constaté que l'architecture adaptative de cette approche présente l'avantage de percevoir la variation qui peut apparaître dans la consommation d'électricité, et d'ailleurs la capacité de fournir une faible valeur d'erreur lors de la prévision de charge d'un temps spécifié.

Afin de prévoir la charge dans un temps déterminé m ; nous créons dans une première étape deux séries de charge parallèle : une première série pour un moment antérieur $m - 1$, et une seconde série pour l'instant m . Dans la deuxième étape, nous cherchons à trouver une paire de paramètres de lissage donnant la plus faible valeur d'erreur lors du passage de l'avant-dernière à la dernière valeur dans la première série de charge. Étant donné que les techniques de lissage exponentiel dépendent principalement des dernières observations, et en particulier la capacité de la méthode de Holt pour détecter la tendance dans les séries chronologiques non saisonnière, il semble que le fait de chercher un modèle optimal assurant l'erreur minimale au

moment $m-1$ peut aider de manière significative à fournir une petite valeur d'erreur pour la prédiction de la charge du temps m spécifié.

5.3.1.6 Méthode de combinaison de prévisions basée sur la moyenne simple

Dans cette méthode, une simple moyenne des prévisions en provenance des modèles individuels est réalisée avec un objectif d'améliorer la précision de prévision. Cette méthode est basée sur l'attribution d'une moyenne arithmétique sur les prévisions individuelles pour un temps m et génère souvent des prévisions fiables pour les jours normaux. Sa représentation mathématique est donnée ci-dessous :

$$\hat{Y}_m^c = \frac{1}{n} \sum_{i=1}^n \hat{Y}_{m,i} \tag{5.1}$$

Où \hat{Y}_m^c est la prévision combinée ; $\hat{Y}_{m,i}$ est la prévision individuelle obtenue à partir du modèle i ; n est égal à 5, le nombre total des modèles individuels proposés.

5.3.2 Une procédure d'identification de valeurs aberrantes pour éviter les grandes erreurs de prévision

Parfois, même les méthodes de combinaison les plus avancées ne parvient pas à améliorer la précision des prévisions, surtout pour les charges anormales où la prédiction précise est toujours considérée comme une tâche difficile. Pour de tels cas, nous pouvons trouver fréquemment que la prévision de l'un des modèles individuels est nettement en dehors de la plage de la majorité des méthodes. Pour éviter les grandes erreurs de prévision, le filtre de Hampel est introduit à la deuxième partie de notre approche pour remplacer les surestimations et les sous-estimations par des valeurs normalisées, ce qui peut aider à améliorer la précision lors de la combinaison des prévisions normalisées.

Le filtre Hampel est un filtre à fenêtre variable qui recherche des valeurs aberrantes ou des données locales anormales dans une séquence temporelle et les remplace par des valeurs alternatives plus raisonnables [98]. Pour une séquence de données (Y_1, Y_2, \dots, Y_n) , le filtre forme un ensemble d'échantillons à partir desquels il calcule la médiane et la médiane des écarts absolus à la médiane (MAD) et effectue le test des valeurs aberrantes. La médiane d'une séquence de données est obtenue en rang ordonnant en premier les valeurs du plus petit au plus grand, à savoir $Y_1 \leq Y_2 \leq Y_3 \leq \dots Y_{n-1} \leq Y_n$ puis prendre soit la valeur du milieu (si n est impaire) ou la moyenne des deux valeurs intermédiaires (si n est paire) [99]. Le filtre Hampel définit les valeurs aberrantes ces points de données Y_j dont la différence absolue à la médiane de la série courte dans la fenêtre mobile est supérieure à un seuil prédéterminé :

$$|Y_j - Y_0| > T \times S_0 \tag{5.2}$$

5.4. Application de HFCM-TM pour la prévision à très court-terme de la charge électrique 92

Où Y_0 représente la valeur médiane des échantillons classés par ordre de la séquence de données de longueur n , T est un seuil scalaire, et S_0 est la médiane des écarts absolus à la médiane, tel que définie par :

$$S_0 = \text{median} \{|Y_j - Y_0|\} \quad (5.3)$$

Si le filtre Hampel déclare un point de donnée pour être une valeur aberrante, sa valeur est alors remplacée par la médiane de ses voisins. Pour le filtre Hampel utilisé dans notre modèle de prévision de charge, la séquence temporelle est définie par les prévisions en provenance des modèles individuels et en provenance de la méthode de combinaison basée sur la moyenne simple.

5.3.3 Estimation finale de la charge électrique

Une fois le filtre Hampel est effectué, la prévision final de la charge est obtenue en tenant compte la moyenne tronquée des prévisions normalisées (la moyenne des prévisions normalisées, en excluant la valeur la plus élevée et la plus basse). Aux fins de comparaison, nous proposons également une deuxième variante de HFCM. Dans cette méthode, la moyenne pondérée est utilisée pour combiner les prévisions normalisées. Mathématiquement, cette prévision combinée a la forme suivante :

$$F(f_1(Y), f_2(Y), \dots, f_n(Y)) = w_1 f_1(Y) + w_2 f_2(Y) + \dots + w_n f_n(Y) \quad (5.4)$$

Où chaque f_i représente un modèle de prévision primaire qui transforme la charge Y d'un quart d'heure (ou 5 minutes) à sa prévision \hat{Y} . Dans le modèle de prévision proposé, les erreurs inverses sont utilisées, de sorte qu'un modèle avec moins d'erreur reçoit plus de poids [100]. À savoir :

$$w_i = \frac{e_i^{-1}}{\sum_{i=1}^n e_i^{-1}}, \quad i = 1, 2, \dots, n \quad (5.5)$$

Où e_i désigne l'erreur absolue en pourcentage obtenue par l' $i^{\text{ème}}$ modèle lors de la prévision de charge des 15 minutes (ou 5 minutes) qui précèdent l'heure de prévision. Par cette méthode, le modèle de prévision qui a un pourcentage d'erreur absolue inférieure reçoit un poids plus élevé dans la composition de la prévision combinée. Pour le reste de ce chapitre, la méthode basée sur l'algorithme de la moyenne arithmétique pondérée sera définie par HFCM-WAM, alors que la méthodologie proposée basée sur la moyenne tronquée sera notée par HFCM-TM.

5.4 Application de HFCM-TM pour la prévision à très court-terme de la charge électrique

Dans cette section, nous décrivons une évaluation de la méthodologie de combinaison de prévisions basée sur le filtre Hampel pour assurer une heure de prédiction

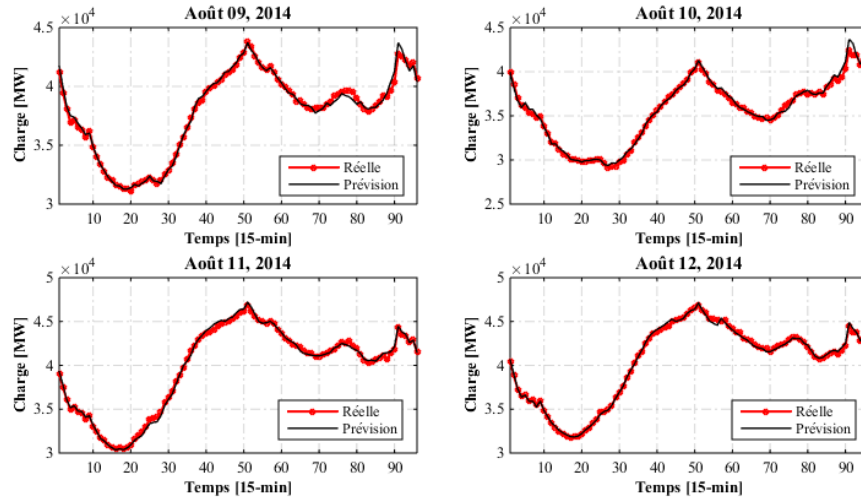


FIGURE 5.4 – Résultats de HFCM-TM pour les données de charge français : du 09 Août au 12 Août 2014

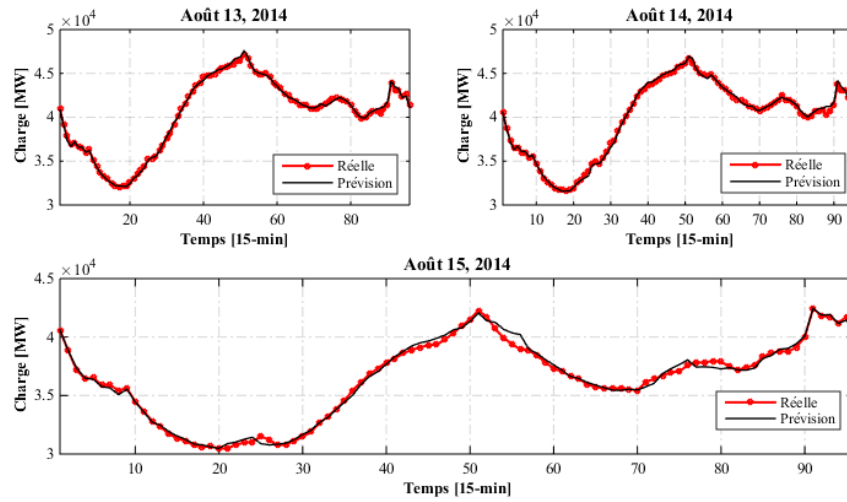


FIGURE 5.5 – Résultats de HFCM-TM pour les données de charge français : du 13 Août 2014-15 Août 2014

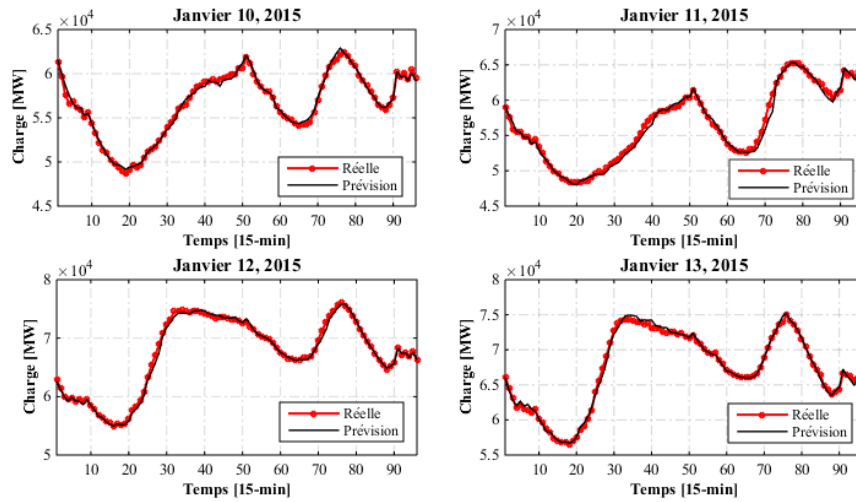


FIGURE 5.6 – Résultats de HFCM-TM pour les données de charge français : du 10 Janvier 2015 au 13 Janvier 2015

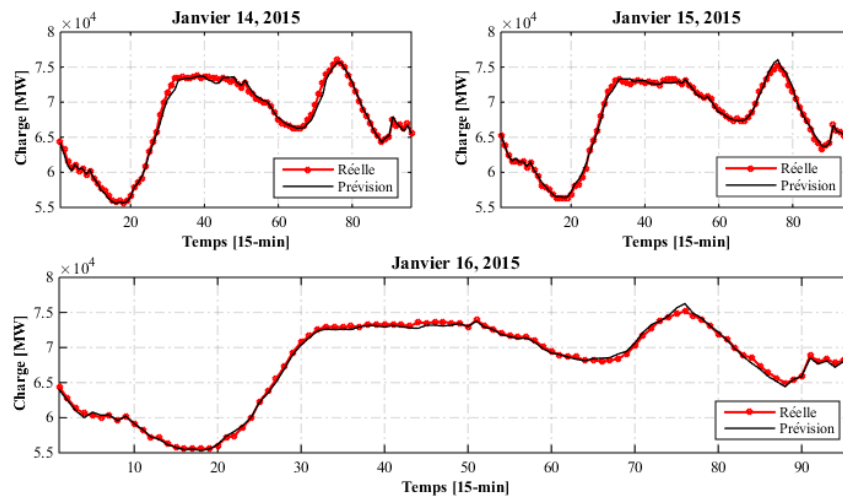


FIGURE 5.7 – Résultats de HFCM-TM pour les données de charge français : du 14 Janvier 2015 au 16 Janvier 2015

Tableau 5.2 – L’erreur relative (MAPE) pour la prévision de la consommation d’électricité en France pendant deux semaines

Temps	SARIMA	HWT	BPNN	ANFIS	AHES	SFCM	HFCM-WAM	HFCM-TM
09/08/2014	1.006	0.566	0.979	0.843	0.721	0.583	0.556	0.570
10/08/2014	1.337	0.758	1.080	0.906	0.739	0.661	0.641	0.642
11/08/2014	0.666	0.569	1.420	0.840	0.583	0.582	0.565	0.552
12/08/2014	0.533	0.488	0.685	0.740	0.393	0.406	0.400	0.404
13/08/2014	0.430	0.634	0.703	0.788	0.479	0.404	0.399	0.376
14/08/2014	0.437	1.511	0.877	0.777	0.334	0.499	0.416	0.421
15/08/2014	1.088	1.710	0.982	0.737	0.743	0.649	0.740	0.642
10/01/2015	1.027	0.483	1.095	0.513	0.703	0.472	0.460	0.438
11/01/2015	1.204	0.777	1.058	0.649	1.040	0.690	0.663	0.648
12/01/2015	1.004	0.608	0.524	0.446	0.685	0.439	0.437	0.430
13/01/2015	0.806	0.758	0.519	0.490	0.617	0.470	0.431	0.430
14/01/2015	0.684	0.728	0.560	0.500	0.589	0.493	0.486	0.486
15/01/2015	0.574	0.728	0.454	0.521	0.641	0.444	0.425	0.415
16/01/2015	0.452	0.829	0.520	0.530	0.445	0.447	0.393	0.407
Jours normaux	0.781	0.726	0.806	0.657	0.613	0.507	0.483	0.478
Tous les jours	0.803	0.796	0.818	0.663	0.622	0.517	0.501	0.490

de la demande d’électricité pour à la fois les conditions normales et spéciales de la charge électrique. Nous allons également comparer les résultats avec un certain nombre d’algorithmes de référence et à d’autres méthodes de combinaison de prévisions. En outre, une discussion sur le temps de calcul de notre outil fourni pour réaliser la prédiction en ligne de la demande d’électricité sera présentée.

5.4.1 Évaluation des prévisions de la consommation française d’électricité

Les Figures 5.4-5.5 et les Figures 5.6-5.7 montrent les résultats de HFCM-TM pour une heure de prédiction de la demande d’électricité en France le long de la semaine du 09 au 15 Août 2014 et la semaine du 10 au 16 Janvier 2015, respectivement. Il est évident à partir de ces figures que notre méthode de combinaison de prévisions basée sur le filtre Hampel a réussi à suivre les variations de la charge électrique, puisque la forme de courbes de charge prévues est très similaire à celle des courbes de charge réelles.

En outre, les résultats de prévision détaillés, comme indiqué sur le Tableau 5.2, démontrent clairement l’efficacité de la stratégie de prévision proposée et montrent également que HFCM-TM surpasse toutes les autres méthodes. Il est visible depuis la performance relative que la valeur moyenne de MAPE pour HFCM-TM dans les jours normaux est égale à 0.478%, ce qui est 5.720% et 22.023% plus inférieur que la méthode de combinaison de prévisions basée sur la moyenne simple et la meilleure approche individuelle (AHES), respectivement. Les résultats montrent également que la méthode basée sur une moyenne pondérée (HFCM-WAM) a été légèrement surpassée par HFCM-TM. Parmi les méthodes de prévision individuelles, AHES

a montré une très bonne précision de prévision. La performance obtenue de cette méthode est principalement attribuée à son caractère adaptatif où les paramètres de lissage, comme le montre la Figure 5.8, sont mis à jour chaque fois qu'une nouvelle observation est enregistrée.

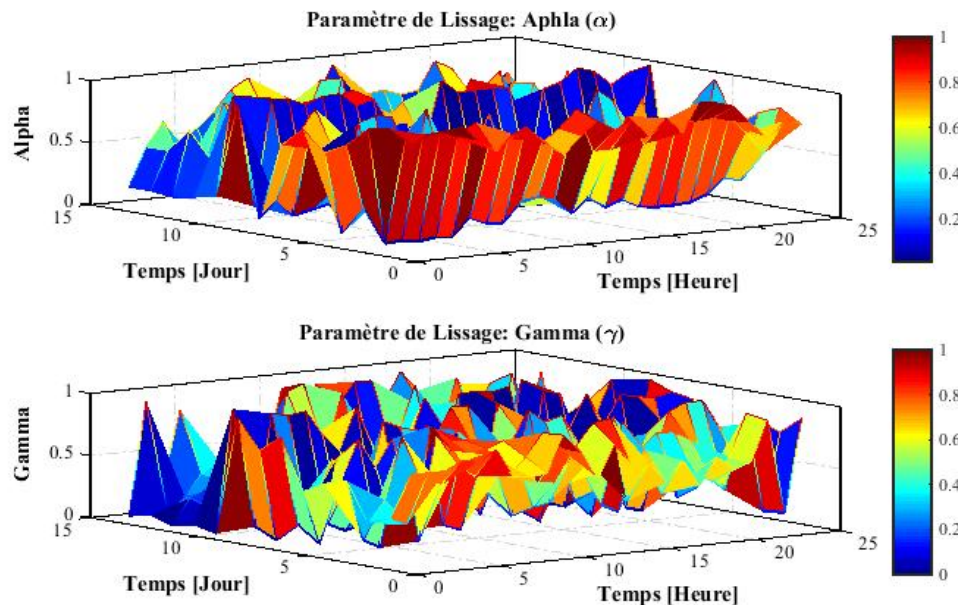


FIGURE 5.8 – Les paramètres de lissage exponentiel de AHES le long d'une semaine en été et une semaine en hiver

Une performance satisfaisante d'une valeur de MAPE égale à 0.743% a été également obtenue par AHES pour le jour férié du 15 Août 2014. Pour cette journée particulière, une nouvelle amélioration de la précision de prévision a été obtenue en utilisant HFCM-TM. Cependant, la prise en compte d'un seul jour férié n'est pas suffisante pour juger la performance de notre modèle de prévision dans de telles journées spéciales. Pour cette raison, une évaluation approfondie d'AHES, SFCM, HFCM-WAM et HFCM-TM, est effectuée hors ligne au cours des jours fériés en France pour l'année 2014 et les résultats des MAPE sont énumérés dans le Tableau 5.3.

On peut donc constater que la MAPE moyenne de HFCM-TM est inférieure à tous les MAPE des autres méthodes. Par rapport au cas des jours normaux où HFCM-TM a montré une précision seulement 5.720% de plus sur SFCM, nous pouvons remarquer dans le cas des jours anormaux qu'une amélioration très significative est obtenue par HFCM-TM, avec une valeur moyenne de MAPE inférieure par 15.887%. En comparant les résultats obtenus par HFCM-TM avec les résultats obtenus par HFCM-WAM, nous pouvons observer clairement que cette dernière méthode a une performance plus faible. Nous pensons que la raison principale de cette situation est

Tableau 5.3 – L’erreur relative (MAPE) pour les jours fériés de 2014 en France

Temps	AHES	SFCM	HFCM-WAM	HFCM-TM
01/01/2014	1.433	1.517	1.382	1.400
21/04/2014	0.603	0.992	0.722	0.747
01/05/2014	1.128	1.111	0.901	0.837
08/05/2014	1.115	0.900	0.789	0.783
29/05/2014	0.946	1.190	0.917	0.898
09/06/2014	1.218	0.883	0.797	0.730
14/07/2014	0.649	0.803	0.628	0.636
15/08/2014	0.743	0.644	0.740	0.642
01/11/2014	1.132	0.812	0.835	0.799
11/11/2014	0.572	0.991	0.670	0.639
25/12/2014	1.563	1.586	1.547	1.525
Tous les jours	0.996	1.026	0.890	0.863

attribuée à l’algorithme pondéré de HFCM-WAM. Dans cette méthode de combinaison, le modèle de prévision ayant la plus faible valeur d’erreur de prévision dans une période d’entraînement reçoit un poids plus élevé dans la composition de la prévision combinée. Cependant, un modèle avec une meilleure performance dans une période d’entraînement n’implique pas un modèle avec une meilleure performance dans la période de validation. Par conséquent, si une sur- ou sous-estimation de la charge est obtenue par ce modèle, nous obtenons alors une prévision erronée si la méthode de combinaison pondérée n’est pas en mesure de détecter et de remplacer cette sur- ou sous-estimation. Pour éviter cette limitation, il sera plus approprié d’attribuer des poids égaux à chaque modèle individuel ; ou mieux encore, d’exclure la valeur la plus élevée et la plus basse de l’ensemble des prévisions normalisées et de n’utiliser que la moyenne arithmétique des quatre valeurs restantes. Par conséquent, la combinaison de prévision présentée dans ce dernier cas est utilisée dans notre modèle HFCM-TM.

Afin de mieux clarifier l’idée clé de notre modèle HFCM-TM, nous présentons sur la Figure 5.9 et dans le Tableau 5.4 les résultats détaillés issus du processus de prévision pour le 26 Décembre 2014 à 00 :00. Nous pouvons citer trois points : (1) une surestimation de la charge a été produite par la meilleure méthode individuelle, (2) HFCM-TM est dépassée par les modèles de BPNN et d’ANFIS, (3) bien que le modèle d’ANFIS a reçu le deuxième poids le plus faible dans la composition de la prévision combinée de HFCM-WAM, les résultats obtenus indiquent que cette approche individuelle a réalisé la meilleure performance de prévision. Cela est cohérent avec ce que nous avons signalé au sujet de la non-existence d’une méthode qui donne les meilleurs résultats pour tous les temps. Cependant, l’avantage de HFCM-TM est situé sur sa capacité à réduire l’incertitude et de fournir des prévisions qui sont très proches des valeurs réelles de charge. Cela peut être clairement observé par les Figures 5.10-5.11, où nous avons comparé la distribution des APE pour AHES, SFCM, HFCM-WAM et HFCM-TM pour les conditions de charge normales et anormales.

Tableau 5.4 – Résultats de prévision pour le 26 Décembre 2014, à 00 :00

Résultats	Charge Réelle	Charge Prédite	APE
Méthode	[MW]	[MW]	[%]
SARIMA		60363	1.285
HWT		60005	0.685
BPNN		59439	0.265
ANFIS	59597	59560	0.062
AHES		61300	2.857
SFCM		60133	0.895
HFCM-WAM		60027	0.721
HFCM-TM		59919	0.540

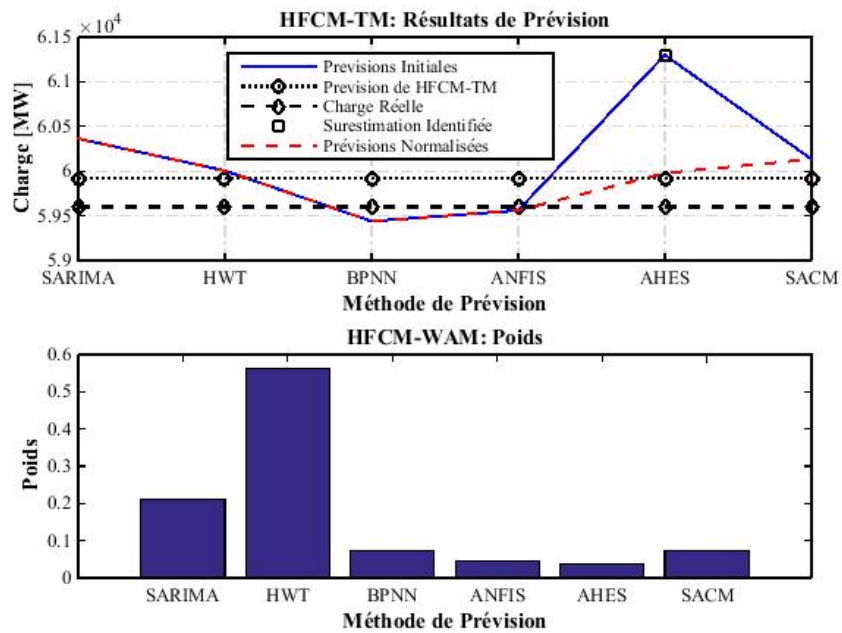


FIGURE 5.9 – Résultats de HFCM-TM et HFCM-WAM pour le processus de prévision du 26 Décembre 2014 à 00 :00

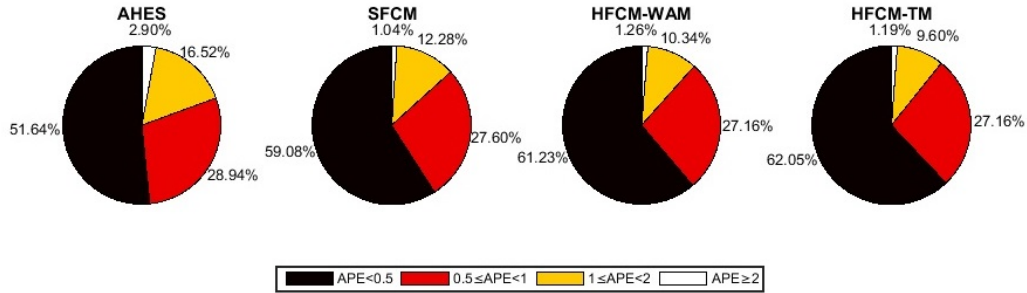


FIGURE 5.10 – Répartition de APE pour AHES, SFCM, HFCM-WAM, et HFCM-TM pour les charges normales

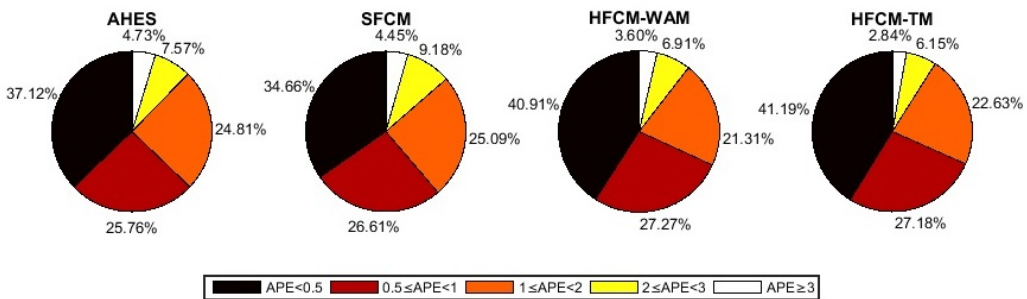


FIGURE 5.11 – Répartition de APE pour AHES, SFCM, HFCM-WAM, et HFCM-TM pour les charges anormales

Bien que la moyenne simple utilisée dans SFCM compte parmi les approches les plus couramment utilisées pour produire les prévisions combinées, nous pouvons percevoir à partir des deux dernières figures que cette méthode de combinaison n'a pas permis à une amélioration de performance dans le cas des conditions anormales de charge. En revanche, nous pouvons observer que HFCM-WAM et HFCM-TM donnent pratiquement les meilleurs résultats et montrent d'excellentes performances de prévision, car une erreur absolue en pourcentage en dessous de 1% a été produite pour plus de 85% et 65% de la période d'évaluation de journées normales et spéciales, respectivement. Cependant, les résultats des Figures 5.10-5.11 montrent que HFCM-TM est plus efficace et permet de mieux éviter les grandes erreurs de prévision.

5.4.2 Évaluation des prévisions de la consommation australienne d'électricité

Dans cette partie, nous procédons à une nouvelle évaluation de notre méthode grâce à une simulation des données australiennes de l'Etat de la Nouvelle-Galles du Sud. La raison principale de cette application est de comparer notre modèle avec

celui proposé dans [101]. Le modèle proposé dans cette référence (AWNN) combine un algorithme d'ondelette avancé pour la décomposition de charge, l'information mutuelle pour la sélection des caractéristiques, et les réseaux de neurones comme algorithme de prédiction. AWNN a été testé sur l'ensemble de données de charge australienne pour l'année 2007 et a réalisé une MAPE égale à 0,938% dans le cas où l'horizon de prévision est de soixante minutes. Pour une comparaison équitable, nous avons utilisé la même période d'évaluation (2007) pour construire une heure de prévision de la charge en utilisant l'algorithme HFCM-TM développé dans le présent chapitre.

La Figure 5.12 présente la charge prévue avec HFCM-TM en fonction d'une charge réelle, et trace aussi les valeurs de charge réelles et prévues pour une journée normale de travail (31 décembre 2007). Cette figure indique la forte corrélation entre les valeurs de charge prévues et réelles.

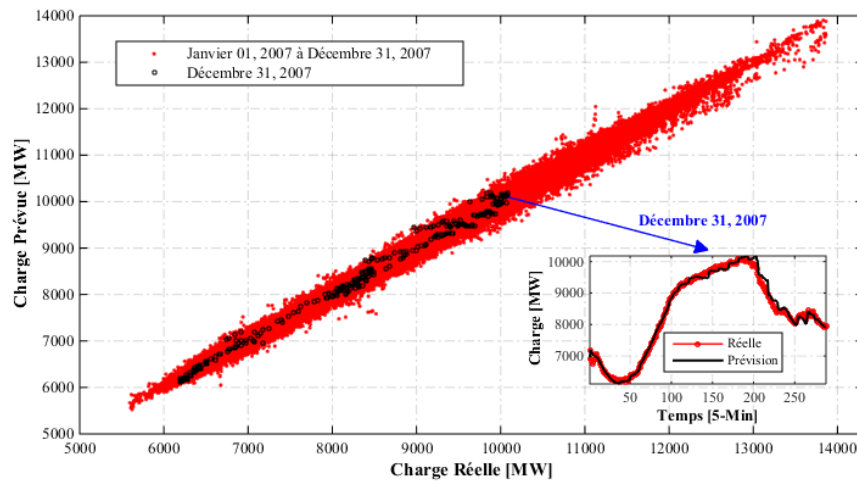


FIGURE 5.12 – Charge prédite par HFCM-TM tracée en fonction de la charge réelle de la base de données australienne

Le Tableau 5.5 compare les performances de prédiction de HFCM-TM avec celles d'AWNN. Les résultats de ce tableau montrent que HFCM-TM atteint des erreurs globales de prédiction comme suit : MAPE = 1.048% pour les jours fériés et MAPE = 0.860% pour tous les jours de 2007. En outre, HFCM-TM a montré de meilleurs résultats de prévision et l'amélioration globale de la MAPE en utilisant HFCM-TM par rapport à AWNN est égale à 8.316%.

5.4.3 Présentation de l'outil proposé pour la prévision en ligne de charge

En plus d'un modèle performant qui assure une très grande précision, un outil de prévision convivial prend également un rôle important dans la mise en œuvre réelle

5.4. Application de HFCM-TM pour la prévision à très court-terme de la charge électrique 101

Tableau 5.5 – L’erreur relative (MAPE) pour les données de charge de la Nouvelle-Galles du Sud

Type de jour Méthode de prévision	Jours fériés de l’année 2007	Tous les jours de l’année 2007
AWNN	-	0.938
HFCM-TM	1.048	0.860

d’une technique de prévision de la charge électrique. En tant que prolongement naturel du travail dans [92], AHES est la méthode qui a été mis en œuvre pour la prévision en ligne de la demande d’électricité en France au cours de la semaine d’été et la semaine d’hiver de la période d’évaluation. Nous avons opté pour l’utilisation des données de charge françaises pour la simple raison que nous avons trouvé le site web du réseau de transport d’électricité de France très utile et précis pour fournir les données de charge électrique en temps réel. Cependant, le but de parvenir à des erreurs de prévision ci-dessous 01% dans les deux conditions de charge normales et anormales nous a conduit, à travers une étude approfondie, au développement de la méthode de combinaison de prévisions basée sur le filtre Hampel (HFCM-TM). Pour l’exécution en ligne de cette méthode, de nombreuses considérations, algorithmes et fonctions ont été prises en compte pour le développement d’un logiciel approprié. Comme le montre la Figure 5.13, notre méthode est actuellement accompagnée d’un outil de prévision et fonctionne très bien pour la prédiction en ligne de la demande d’électricité du système électrique français. Cet outil de prévision peut effectuer des prévisions pour une heure à l’avance à partir de n’importe quel moment sélectionné par l’utilisateur. Cependant, pour notre étude actuelle, les mêmes points sont utilisés pour chaque jour de la période d’évaluation. Par conséquent, les premières étapes qui constituent les prévisions d’une heure à l’avance sont 00 :15, 01 :15, 02 :15 ... 23 :15. Ainsi, de cette manière, nous obtenons seulement une seule valeur prévue pour chaque moment particulier de la journée.

Cependant, trois tâches principales peuvent être réalisées en utilisant notre outil de prévision. Évidemment, la principale tâche de notre outil est de permettre la prédiction en ligne de la demande d’électricité. Tout ce que l’utilisateur est tenu de faire pour réaliser cette fonction est de sélectionner la méthode de prévision et le type de jour dans lequel il aimerait effectuer les prédictions et le logiciel complète alors la tâche en passant par quatre étapes principales : (1) l’obtention de l’heure actuelle en France, (2) la connexion au site web de RTE et le téléchargement des données de charge disponibles pour la journée de prévision, (3) la mise à jour de la base de données en additionnant les données téléchargées au fichier principal des données historiques de charge, (4) l’exécution des algorithmes nécessaires et l’exposition des résultats de prévision obtenu. En plus de la mise en œuvre en ligne, l’outil proposé permet également d’effectuer les prédictions hors ligne. Ceci est important si l’utilisateur souhaite voir les performances de prévision précédentes en prenant en compte différents critères pour la mesure d’erreur. La troisième tâche qui peut être

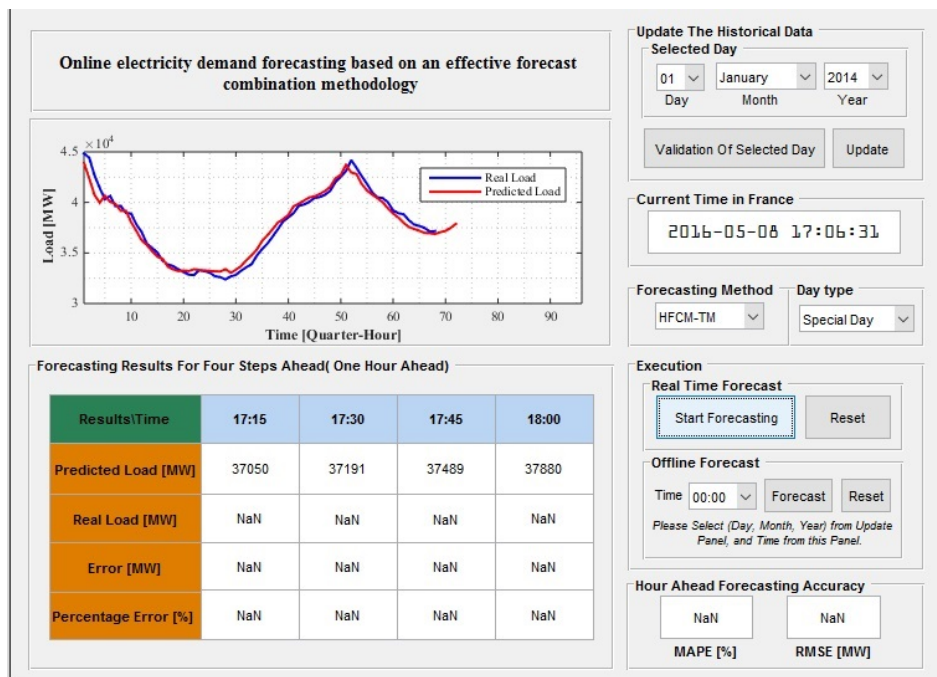


FIGURE 5.13 – Outil développé pour la prédiction en ligne de la demande d’électricité

réalisée par notre logiciel est la capacité de mettre à jour les données historiques de charge dans le cas où l'utilisateur a négligé l'utilisation de l'outil pendant plusieurs jours. Dans ce cas, une intervention de l'utilisateur est nécessaire pour sélectionner un jour souhaité à partir du panneau «Update The Historical Data» afin d'inclure les données manquées dans le fichier de données historiques principal.

Nous présentons dans le Tableau 5.6 le temps de calcul de notre logiciel sur un ordinateur portable de 4 Go de RAM, processeur i3 380M, un disque dur de 5200 tr/min ; et qui fonctionne sous l'environnement de Windows 10. Le tableau montre le temps de calcul de chaque méthode considérée pour quatre cas différents : le premier cas (09 août 2014) est lié à une exécution hors ligne pour assurer une heure de prévision de charge au cours d'un week-end, le deuxième cas (25 décembre 2014) est lié à une mise en œuvre hors ligne pour effectuer une heure de prévision de charge au cours d'un jour férié, le troisième cas (27 octobre 2016) est lié à une exécution en ligne pour assurer une heure de prévision pendant un jour ouvrable, et le quatrième cas (01 novembre 2016) est lié à une exécution en ligne pour effectuer une heure de prévision de la consommation d'électricité pendant un jour férié. On peut clairement percevoir à partir de ces résultats que les modèles individuels nécessitent moins de temps de calcul par rapport aux méthodes combinées. Cependant, la différence de temps de calcul est très acceptable si l'on tient compte que les méthodes combinées nécessitent l'utilisation simultanée de nombreux algorithmes et méthodes pour effectuer les prédictions. De plus, le temps d'exécution de notre méthode de HFCM-TM est inférieur à quinze secondes. Ce temps inclut le temps de mettre à jour les données historiques de charge et d'écrire les résultats de prévision sur un fichier XLS. Si

Tableau 5.6 – Le temps de calcul de toutes les méthodes considérées dans quatre cas de prévision différents

Méthode	SARIMA	HWT	BPNN	ANFIS	AHES	SFCM	HFCM-WAM	HFCM-TM
Résultats								
Période d'évaluation	Sam, 09 Aou 2014, 12 :45-13 :30							
Temps de calcul [s]	4.38	5.16	6.08	4.29	4.71	7.96	11.85	8.39
Période d'évaluation	Jeu, 25 Dec 2014, 23 :15-00 :00							
Temps de calcul [s]	6.57	7.39	9.20	5.28	7.66	11.24	15.83	11.61
Période d'évaluation	Jeu, 27 Oct 2016, 21 :15-22 :00							
Temps de calcul [s]	6.77	7.28	7.84	5.97	8.51	11.39	16.61	12.35
Période d'évaluation	Mar, 01 Nov 2016, 21 :30-22 :15							
Temps de calcul [s]	6.27	6.93	8.35	5.84	8.25	10.84	13.24	11.06

nous comparons le temps d'exécution du HFCM-TM pour le jour férié de 2014 avec celui du jour férié de 2016, on peut clairement percevoir que les deux cas nécessitent pratiquement le même temps d'exécution d'environ onze secondes. Ceci peut être facilement justifié si l'on tient compte que nous avons limité les données à utiliser par chaque méthode individuelle à une quantité prédéfinie d'échantillons. Par exemple, même si les données historiques de charge sur le fichier principale sont constituées de vingt années d'observations de la demande d'électricité, la méthode HWT devrait utiliser uniquement la dernière année d'observations comme une quantité suffisante de données pour traiter la saisonnalité de la série chronologique et assurer les prédictions de charge. Par conséquent, avec un temps d'exécution inférieur à quinze secondes, notre méthode combinée peut être facilement implémentée à n'importe quelle compagnie d'électricité pour effectuer les prédictions en ligne de la demande d'électricité.

5.5 Conclusion

Dans ce chapitre, une nouvelle technique a été développée pour réaliser la prédiction en ligne de la demande d'électricité pour un horizon d'une heure d'avance. Étant donné que le comportement de la charge électrique est complexe et la consommation est principalement affectée par plusieurs facteurs, il n'y a pas de méthode de prévision qui donne le meilleur résultat tout le temps. Il est donc nécessaire d'intégrer différents modèles afin de mieux réduire l'incertitude et améliorer la précision des prévisions. À cet égard, la méthode proposée est classée dans la catégorie des méthodes de prévision combinées. Pour faire face au comportement saisonnier et au caractère complexe et non linéaire de la série chronologique de la charge électrique, deux méthodes statistiques et deux approches de l'intelligence computationnelle, ont été prises en considération. Une méthode de lissage exponentiel adaptative récemment développée et une simple combinaison des prévisions des cinq modèles individuels ont été également utilisées comme modèles candidats. Pour améliorer la performance lors de l'association des prévisions des modèles primaires, un filtre de prétraitement de données et la moyenne tronquée ont été employés comme des éléments clés dans cette fonction d'éviter d'importantes erreurs de prévision. Les résultats obtenus après une évaluation sur les données de charge électrique provenant de deux pays différents (la France et l'Australie) ont montré des performances satisfaisantes dans les deux conditions normales et spéciales de la charge électrique, avec une erreur absolue moyenne en pourcentage de l'ordre de 1% dans les deux cas. En outre, les résultats reflètent de meilleures performances de l'approche développée que d'autres méthodes considérées dans cette étude. Par exemple, les résultats provenant des jours fériés en France ont montré que les améliorations de précision en termes de d'erreur absolue moyenne en pourcentage par rapport à, une méthode de combinaison basée sur une moyenne simple, la meilleure méthode individuelle, et une combinaison pondérée sont 15.887%, 13.353%, et 3.034%, respectivement. En outre, une amélioration significative de 8.316% a été obtenue en utilisant le sys-

tème de prévision proposé en comparaison avec un algorithme de référence issu de la littérature. De plus, notre méthode est également accompagnée d'une interface graphique conviviale et le temps d'exécution de ce logiciel sur un ordinateur portable de 4 Go de RAM, processeur i3 380M, et un disque dur de 5200-tpm, est d'environ quinze secondes. Par conséquent, la méthodologie proposée peut être facilement implémenté à tout réseau électrique pour effectuer les prédictions en ligne de la charge électrique.

Conclusions et perspectives

6.1 Conclusions

Ce travail a porté sur la prévision de la charge électrique pour quatre horizons temporels. Deux grandes catégories de techniques de prévision ont été mises en évidence. Il s'agit des techniques statistiques et celles de l'intelligence artificielle.

Les techniques statistiques se caractérisent par des modélisations simples. Elles aboutissent, généralement, à des prévisions raisonnables basées sur des données historiques ayant des termes de tendance, de saisonnalité et d'autres composantes constituantes.

Les techniques univariées (e.g. ARIMA) montrent parfois une précision limitée en ignorant les effets des facteurs exogènes. Par contre, les modèles classiques multivariées telles que la régression, tiennent compte de ces derniers mais leur caractère linéaire peut faire face à une grande difficulté à déterminer empiriquement une relation correcte entre la demande d'électricité et les autres entrées explicatives.

Les techniques d'intelligence artificielle ont permis une modélisation plus perfectionnée des relations non linéaires et complexes entre les données de charge électrique et les différents facteurs influents la consommation d'électricité. Les études de la prévision de la charge électrique se sont, alors, penchées à leur application.

Cependant, ce n'est pas une tâche facile d'assurer les prédictions précises de la charge, car le comportement de la demande d'électricité est complexe et la série chronologique de la charge est non-stationnaire dans la moyenne et la variance, contenant des termes de tendance, de multiples cycles de fluctuations saisonnières, ainsi qu'un comportement de variation aléatoire. Donc, le problème est encore considéré comme difficile pour les chercheurs en raison de sa grande complexité. L'estimation de la charge future avec des données historiques est restée une difficulté jusqu'à présent, en particulier pour la prévision dans des conditions atypiques telles que le cas de la variation rapide de la consommation d'électricité pendant les vacances, les jours avec des conditions météorologiques extrêmes et les autres jours anormaux. L'objectif de notre thèse été donc de contribuer au développement des techniques et méthodologies permettant de prévoir efficacement la consommation d'électricité dans les différents situations de charge. Nous avons effectué principalement trois études : (1) une étude pour à la fois la prévision à moyen et à long-terme de la charge électrique, (2) une étude pour la prévision à court-terme de la charge électrique, (3) et une dernière étude pour la prévision à très court-terme de la demande d'électricité.

Dans un premier lieu, nous avons utilisé les observations de la demande d'électricité en Algérie pour évaluer des méthodes de prévision à moyen et à long-terme des pics de charge électrique. Les résultats de prévision ont montré que le système d'inférence neuro-flou adaptatif et les modèles basés sur le processus autorégressif à moyenne mobile intégré ont été bien performés pour les deux cas de prévision à moyen et à long-terme de la demande d'électricité. Cependant on ne peut pas dire définitivement laquelle des deux méthodes est la plus performante : le système d'inférence neuro-flou adaptatif a été la meilleure approche pour la prévision à long-terme des pics de charge électrique, alors que la meilleure performance a été obtenue par le processus saisonnier autorégressif à moyenne mobile intégré dans le cas de prévision à moyen terme. Des méthodes plus simples telles que le lissage exponentiel de Holt ont été également bien concurrencées dans la situation où les données historiques contiennent une tendance. Donc, il n'y a pas de méthode de prévision qui donne le meilleur résultat pour tous les temps : certaines approches peuvent performer très bien dans certaines circonstances, et ne refléteraient pas les variations de la charge dans d'autres. Afin d'améliorer la précision de prévision et tirer la meilleure performance dans le maximum de cas possible, nous avons pensé sur l'utilisation simultanée de plusieurs méthodes de prévisions. Cette idée a été également l'élément clé dans la méthodologie présentée dans l'étude pour la prévision à court-terme de la consommation d'électricité.

La deuxième étude est consacrée à la prévision des pics quotidiens de la charge électrique en Algérie en utilisant une méthodologie hybride adaptative à deux étapes. Parce que la série chronologique considérée est principalement influencée par des effets saisonniers journaliers et hebdomadaires, la méthode de lissage exponentiel à double saisonnalité de Holt-Winters-Taylor a été utilisée en premier lieu pour prévoir la demande de pointe du jour suivant. Cette méthode peut concurrencer bien dans les jours normaux, mais elle peut faire face à des difficultés lorsqu'elle est soumise à des conditions journalières spéciales comme le cas de l'évolution rapide de la demande de charge pendant les vacances. A cet effet, un modèle secondaire a été proposé dans le but d'améliorer la performance et d'assurer les prévisions satisfaisantes pour à la fois les conditions quotidiennes normales et spéciales de la charge électrique. Notre vision pour cette méthodologie a été donc d'utiliser un système adaptatif à deux étages, où nous donnons dans une première étape une prévision primaire et nous essayons ensuite d'améliorer la précision lorsque des variables telles que les valeurs réelles de température pour le jour à prévoir seront disponibles. Cette méthodologie à deux étages a montré de meilleurs résultats de prévision en comparaison avec la méthode à double saisonnalité de Holt-Winters-Taylor, des modèles de réseaux de neurones, et des approches fondées sur le système d'inférence neuro-flou adaptatif. Par conséquent, cette méthodologie est un outil efficace pour améliorer la précision des prévisions de la pointe de consommation pour à la fois les conditions normales et spéciales de la charge électrique.

Pour le cas de prévision à très court-terme il n'y aura pas de temps pour tenter une amélioration de performance pour un horizon d'exécution aussi court que quelques minutes à l'avance. Donc, il fallait chercher la meilleure performance dès un

premier temps. Par conséquent, notre idée pour ce cas de prévision est d'utiliser une combinaison de prévisions et de tirer la meilleure performance à partir d'un ensemble de méthodes primaires. En plus, la méthodologie proposée doit, pas seulement assurer des prédictions satisfaisantes pour des conditions normales et anormales de la charge électrique, mais exige également l'inclusion de beaucoup d'autres paramètres, telle que la capacité d'adaptation en ligne, la simplicité d'utilisation et le temps de calcul faible. La structure de notre méthodologie proposée se compose de trois étapes principales. À la première partie, plusieurs modèles ont été utilisés pour prévoir la consommation d'électricité une heure d'avance. Nous avons choisi d'utiliser deux méthodes classiques, deux méthodes d'intelligence artificielle, une méthode récemment développée basée sur le lissage exponentiel adaptatif, et une simple combinaison des prévisions issus des cinq modèles individuels. Pour éviter les grandes erreurs de prévision, un filtre de détection de valeurs aberrantes est introduit à la deuxième partie de notre approche en tant qu'un élément clé pour remplacer les sous- et sure-estimations par des valeurs normalisés. Ensuite, l'estimation finale de la charge est obtenue en considérant la moyenne tronquée des prévisions normalisées. Les données de charge en temps-réel du système électrique français et l'ensemble de données australiennes pour l'Etat de la Nouvelle-Galles du Sud sont utilisées comme un exemple illustratif pour évaluer la performance de la méthode proposée. Les résultats obtenus ont reflété de meilleures performances de l'approche développée que d'autres méthodes considérées dans cette étude. En outre, une amélioration significative a été obtenue en utilisant le système de prévision proposé en comparaison avec un algorithme de référence de la littérature, nommé AWNN, et testé sur l'ensemble de données de charge australienne pour l'année 2007. En plus d'un modèle performant qui assure une très grande précision, un outil de prévision convivial prend également un rôle important dans la mise en œuvre réelle d'une technique de prévision de la charge électrique. A cet effet, notre méthode a été accompagnée également d'un outil de prévision qui a très bien fonctionné pour la prédiction en ligne de la demande d'électricité du système français durant à la fois les conditions normales et spéciales de la charge électrique. Cependant, avec un temps d'exécution d'environ quinze secondes, notre méthodologie combinée peut être facilement implémenté à tout réseau électrique pour effectuer les prédictions en ligne de la charge électrique.

6.2 Perspectives

Globalement, notre thèse a apportée trois contributions principales :

- nous avons développé un modèle pour la prévision à court-terme des pics journaliers de la charge électrique où la performance de prévision obtenue a été meilleur que celle d'un modèle de référence, abrégé HWT, et développé par J.W. Taylor en 2003.
- nous avons développé un modèle pour la prévision à très court-terme de la charge électrique où la performance de prévision obtenue a été meilleur que celle d'un modèle de référence, abrégé AWNN et développé par M. Rana et

al en 2016.

- le modèle de prévision à très court-terme a été accompagné également avec un outil de prédiction convenable, dont le but est de faciliter la tâche de prédiction à l'utilisateur et d'assurer des prédictions satisfaisantes et adaptatives quelle que soit la situation exposée.

Cependant, les perspectives de ce travail s'orientent sur les optiques suivantes :

- la réalisation d'une étude plus approfondie pour la prédiction à moyen et à long-terme de la charge électrique. En fait, on été limité dans notre étude présentée par une quantité insuffisante de données de charge. Pour une meilleure performance de prévision, la considération des paramètres reliés au développement économique, accroissement démographique et innovation technologique serait d'un apport bien souhaitable pour des périodes de prédiction aussi longues que plusieurs années à venir.
- pour l'étude présentée dans le cadre de prévision à court-terme des pics de charge électrique, la performance finale de prévision est liée directement à la prévision primaire provenant de la méthode de Holt-Winters-Taylor. Cette méthode de lissage exponentiel présente un clair avantage pour modéliser les cycles de saisonnalité contenus dans la série chronologique de la consommation d'électricité. Cependant, il sera plus intéressant d'utiliser un modèle primaire de prévision qui tient en compte la modélisation des conditions typiques et atypiques de la charge électrique dans une seule capsule.
- il est fortement souhaité, dans le cadre de prévision à très court-terme de la charge électrique, d'utiliser des données de température pour modéliser les jours avec des conditions climatiques extrêmes.
- pour faciliter la tâche de prédiction, il est conseillé de développer un logiciel plus complet qui peut assurer, dans un seul volet, les prédictions des différents horizons temporels, et pour différentes régions et pays.
- au cours de la décennie la plus récente, la modernisation du réseau a apporté de nombreux changements à l'industrie des services électriques. Des changements tels que l'augmentation de la concurrence sur le marché, la pénétration croissante des énergies renouvelables et la communication bidirectionnelle entre le fournisseur et les utilisateurs finaux ont suscité beaucoup plus d'incertitudes pour les pratiques de prévision de charge des services publics. La prévision unique qui donne une prévision déterministe de la charge future ne fournit aucune information sur ces incertitudes. Pour diminuer l'incertitude, les prévisions probabilistes de la charge peuvent considérablement contribuer afin d'assurer une meilleure planification et exploitation des systèmes énergétiques.

Bibliographie

- [1] RTE France, *Production d'électricité par filière*, lien = "www.rte-france.com/fr/eco2mix/eco2mix-mix-energetique". (Cité en pages viii et 4.)
- [2] E.A. Feinberg, and D. Genethliou, *Load Forecasting*, In : Applied Mathematics for Restructured Electric Power Systems, Part of the Series Power Electronics and Power Systems, pp. 269-285, 2005. (Cité en page 5.)
- [3] E.B.U. Islam, *Comparison of Conventional and Modern Load Forecasting Techniques Based on Artificial Intelligence and Expert Systems*, In : International Journal of Computer Science, vol. 8, Issue 5, no. 3, 2011. (Cité en page 6.)
- [4] P.R. Sircar, and J.W. Murphy, *An Econometric Load Forecasting Model for Iowa Utilities : An Empirical Assessment*, In : Proceedings of Symposium on Computer Analysis of Electric Load Forecasting and Generating Capacity Expansion, National Regulatory Research Institute, Columbus, Ohio, 1983. (Cité en page 6.)
- [5] A. Goett, *Household appliance choice : Revision of REEPS behavioural models*, In : Electric Power Research Institute, Report EA-3409, Palo Alto, CA, USA, 1984. (Cité en page 6.)
- [6] J.E. McMahon et al, *The LBL residential energy and hourly demand models*, In : Strategic Planning in Energy and Natural Resources, 1987. (Cité en page 6.)
- [7] B. Lapillonne, *MEDEE 2 : A Model for Long-Term Energy Demand Evaluation*, In : IIASA Research Report, 1978. (Cité en page 6.)
- [8] R. Bartels, D.G. Fiebig, M. Garben, and R. Lumsdaine, *An end-use electricity load simulation model : Delmod*, In : Utilities Policy, vol. 2, Issue 1, pp. 71-82, 1992. (Cité en page 6.)
- [9] S.L. Corpening, N.D. Reppen and R.J. Ringlee, *Experience with Weather Sensitive Load Models for Short and Long-Term Forecasting*, In : IEEE Transactions on Power Apparatus and Systems, vol. PAS-92, no. 6, pp. 1966-1972, 1973. (Cité en page 6.)
- [10] T.K. Basu, T.K. Bhattacharya, and P. Purkayastha, *Medium range forecasting of hourly power system load by time series analysis using the Walsh transform*, In : International Journal of Electrical Power & Energy Systems, vol. 13, no. 4, pp. 193-200, 1991. (Cité en page 6.)
- [11] T.K. Bhattacharya, and T.K. Basu, *Medium range forecasting of power system load using modified Kalman filter and Walsh transform*, In : International Journal of Electrical Power & Energy Systems, vol. 15, no. 2, pp. 109-115, 1993. (Cité en page 6.)

- [12] S.M. Islam, Saleh M. Al-Alawi, and K.A. Ellithy, *Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network*, In : Electric Power Systems Research, vol. 34, no. 1, pp. 1-9, 1995. (Cité en page 7.)
- [13] S.K. Padmakumari, K.P. Mohandas, and D. Thiruvengadam, *Application of fuzzy system theory in land use based long term distribution load forecasting*, In : International Conference on Energy Management and Power Delivery, Proceedings of EMPD'98, vol. 1, pp. 311-316, 1998. (Cité en page 7.)
- [14] K. Padmakumari, K.P. Mohandas, and S. Thiruvengadam, *Long term distribution demand forecasting using neuro fuzzy computations*, In : International Journal of Electrical Power & Energy Systems, vol. 21, no. 5, pp. 315-322, 1999. (Cité en page 7.)
- [15] M.S. Kandil, S.M. El-Debeiky, and N.E. Hasanien, *Long-term load forecasting for fast developing utility using a knowledge-based expert system*, In : IEEE Transactions on Power Systems, vol. 17, no. 2, pp. 491-496, 2002. (Cité en page 7.)
- [16] W.R. Christiaanse, *Short-Term Load Forecasting Using General Exponential Smoothing*, In : IEEE Transactions on Power Apparatus and Systems, vol. PAS-90, no. 2, pp. 900-911, 1971. (Cité en page 7.)
- [17] M.T. Hagan, and S.M. Behr, *The Time Series Approach to Short Term Load Forecasting*, In : IEEE Transactions on Power Systems, vol. 2, no. 3, pp. 785-791, 1987. (Cité en page 7.)
- [18] A.D. Papalexopoulos, and T.C. Hesterberg, *A regression-based approach to short-term system load forecasting*, In : IEEE Transactions on Power Systems, vol. 5, no. 4, 1990. (Cité en page 7.)
- [19] D.C. Park, M.A. El-Sharkawi, R.J. Marks II, L.E. Atlas, and M.J. Damborg, *Electric load forecasting using an artificial neural network*, In : IEEE Transactions on Power Systems, vol. 6, no. 2, pp. 442-449, 1991. (Cité en page 7.)
- [20] K. Liu et al, *Comparison of very short-term load forecasting techniques*, In : IEEE Transactions on Power Systems, vol. 11, no. 2, pp. 877-882, 1996. (Cité en page 7.)
- [21] R.H. Liang, and C.C. Cheng, *Short-term load forecasting by a neuro-fuzzy based approach*, In : International Journal of Electrical Power & Energy Systems, vol. 24, no. 2, pp. 103-111, 2002. (Cité en page 7.)
- [22] A.J.R. Reis, and A.P.A. da Silva, *Feature extraction via multiresolution analysis for short-term load forecasting*, In : IEEE Transactions on Power Systems, vol. 20, no. 1, pp. 189-198, 2005. (Cité en page 7.)
- [23] S. Fan, and L. Chen, *Short-term load forecasting based on an adaptive hybrid method*, In : IEEE Transactions on Power Systems, vol. 21, no. 1, pp. 392-401, 2006. (Cité en page 7.)

- [24] M. El-Telbany, *Short-term forecasting of Jordanian electricity demand using particle swarm optimization*, In : Electric Power Systems Research, vol. 78, pp. 425-433, 2008. (Cit  en page 7.)
- [25] W.C. Hong, *Hybrid evolutionary algorithms in a SVR-based electric load forecasting model*, In : International Journal of Electrical Power & Energy Systems, vol. 31, no. 7-8, pp. 409-417, 2009. (Cit  en page 7.)
- [26] J.J. Liang, H. Song, B. Qu, W. Liu, and A.K. Qin, *Neural Network Based on Dynamic Multi-swarm Particle Swarm Optimizer for Ultra-Short-Term Load Forecasting*, In : Advances in Swarm Intelligence, the series Lecture Notes in Computer Science, vol. 8795, pp. 384-391, 2014. (Cit  en page 7.)
- [27] C.E. Borges, Y.K. Penya, and I. Fernandez, *Evaluating Combined Load Forecasting in Large Power Systems and Smart Grids*, In : IEEE Transactions on Industrial Informatics, vol. 9, no. 3, pp. 1570-1577, 2013. (Cit  en page 7.)
- [28] R. Clemen, *Combining forecasts : A review and annotated bibliography*, In : International Journal of Forecasting, vol. 5, pp. 559-583, 1989. (Cit  en page 7.)
- [29] J.W. Taylor, *An Evaluation of Methods for Very Short-Term Load Forecasting Using Minute-by-Minute British Data*, In : International Journal of Forecasting, vol. 24, pp. 645-658, 2008. (Cit  en page 8.)
- [30] J. Wang, J. Wang, Y. Li, S. Zhu, and J. Zhao, *Techniques of applying wavelet de-noising into a combined model for short-term load forecasting*, In : International Journal of Electrical Power & Energy Systems, vol. 62, pp. 816-824, 2014. (Cit  en page 8.)
- [31] Z. Liu, W. Sun, and J. Zeng, *A new short-term load forecasting method of power system based on EEMD and SS-PSO*, In : Neural Computing and Applications, vol. 24, no. 3, pp. 973-983, 2014. (Cit  en page 8.)
- [32] L. Xiao, J. Wang, R. Hou, and J. Wu, "A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting", Energy, vol. 82, pp. 524-549, 2015. (Cit  en page 8.)
- [33] J. Nowotarski, B. Liu, R. Weron, and T. Hong, "Improving short term load forecast accuracy via combining sister forecasts", Energy, vol. 98, pp. 40-49, 2016. (Cit  en page 8.)
- [34] L. Xiao, W. Shao, T. Liang, and C. Wang, *A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting*, In : Applied Energy, vol. 167, pp. 135-153, 2016. (Cit  en page 8.)
- [35] M. Ghofrani, M. Ghayekhloo, A. Arabali, and A. Ghayekhloo, *A hybrid short-term load forecasting with a new input selection framework*, In : Energy, vol. 81, pp. 777-786, 2015. (Cit  en page 9.)
- [36] S. Arora, and J.W. Taylor, *Short-Term Forecasting of Anomalous Load Using Rule-Based Triple Seasonal Methods*, In : IEEE Transactions on Power Systems, vol. 28, no. 3, pp. 3235-3242, 2013. (Cit  en page 9.)

- [37] K.H. Kim, H.S. Youn, and Y.C. Kang, *Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method*, In : IEEE Transactions on Power Systems, vol. 15, no. 2, pp. 559-565, 2000. (Cité en page 9.)
- [38] M. Ghayekhloo, M. B. Menhaj, and M. Ghofrani, *A hybrid short-term load forecasting with a new data preprocessing framework*, In : Electric Power Systems Research, vol. 119, pp. 138-148, 2015. (Cité en page 9.)
- [39] M. Moazzami, A. Khodabakhshian, and R. Hooshmand, *A new hybrid day-ahead peak load forecasting method for Iran's National Grid*, In : Applied Energy, vol. 101, pp. 489-501, 2013. (Cité en page 11.)
- [40] L.C. Moreira de Andrade, and I. Nunes da Silva, *Very Short-Term Load Forecasting Based on ARIMA Model and Intelligent Systems*, In : 15th International Conference on Intelligent System Applications to Power Systems (ISAP'09), Curitiba, pp. 1-6, 2009. (Cité en page 11.)
- [41] P. Murto, *Neural network models for short-term load forecasting*, Master's Thesis, Helsinki University of technology, Department of Engineering, 1998. (Cité en page 12.)
- [42] RTE France, *Consommation française d'électricité : Caractéristiques et méthode de prévision*, Rapport, 2014. (Cité en pages 12 et 18.)
- [43] P.J. Santos, A.G. Martins, A.J. Pires, J.F. Martins, and R.V. Mendes, *Short Term load forecast using trend information and process reconstruction*, In : International Journal of Energy Research, 2006. (Cité en page 12.)
- [44] D. Chikobvu , and C. Sigauke, *Regression-SARIMA modelling of daily peak electricity demand in South Africa*, In : Journal of Energy in Southern Africa, vol. 23, no. 3, pp. 23-30, 2012. (Cité en page 19.)
- [45] A. Laouafi, *Prédiction de la courbe de charge par des techniques intelligentes*, Mémoire présenté pour obtenir le diplôme de Master en électrotechnique industrielle, Université 20 Août 1955 - Skikda, 2012. (Cité en pages viii, 18, 19, 20, 21 et 31.)
- [46] H.C. Jiann, *Seasonal Adjustment of Time Series*, In : Statistics Singapore Newsletter, 2005. (Cité en page 22.)
- [47] F. Avram, *Séries temporelles : régression, et modélisation ARIMA(p,d,q)*, Notes de Cours, Université de Pau et des pays de l'Adour, 2012. Lien = "<http://web.univ-pau.fr/~avram/sertemp/ser.pdf>". (Cité en pages viii et 23.)
- [48] J. Durbin, and P. B. Kenny, *Seasonal Adjustment When the Seasonal Component Behaves Neither Purely Multiplicatively nor Purely Additively*, In : Seasonal Analysis of Economic Time Series, pp. 173-200, 1979. (Cité en page 24.)
- [49] E. Ostertagová, and O. Ostertag, *The simple exponential smoothing model, modeling of mechanical and mechatronic systems*, In : The 4th International conference, Faculty of Mechanical engineering, Technical university of Košice, 2011. (Cité en page 25.)

- [50] G.U. Yule, *On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers*, In : Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences, vol. 226, pp. 636-646, 1927. (Cit  en page 25.)
- [51] E. Slutsky, *The summation of random causes as the source of cyclic processes*, In : Econometrica, vol. 5, pp. 105-146, 1927. (Cit  en page 26.)
- [52] G.E.P. Box, and G.M. Jenkins, *Time Series Analysis : Forecasting and Control*, In : San Francisco : Holden-Day, 1970. (Cit  en page 26.)
- [53] Y. Chakhchoukh, P. Panciatici, and L. Mili, *Electric Load Forecasting Based on Statistical Robust Methods*, In : IEEE Transactions on Power Systems, vol. 26, no. 3, pp. 982-991, Aug. 2011. (Cit  en page 27.)
- [54] C. Guan, P.B. Luh, L.D. Michel, Y. Wang, and P.B. Friedland, *Very Short-Term Load Forecasting : Wavelet Neural Networks With Data Pre-Filtering*, In : IEEE Transactions on Power Systems, vol. 28, no. 1, 2013. (Cit  en pages viii et 28.)
- [55] B. Farouk, *M thodes Heuristiques pour la Pr diction des S ries Temporelles*, Th se pr sent e en vue de l'obtention du dipl me de Doctorat en Informatique, Universit  Badji Mothtar-Annaba, 2012. (Cit  en page 30.)
- [56] P. Bretschneider, T. Rauschenbach, and J. Wernstedt, *Forecast using an adaptive fuzzy classification algorithm for load*, In : Proceedings of the 5th European Control Conference, Karlsruhe, Germany, 1999. (Cit  en page 30.)
- [57] Y.Y. Hsu, and C.C. Yang, *Design of artificial neural networks for short-term load forecasting Part I : Self-organizing feature maps for day type identification*, In : IEE Proceedings C (Generation, Transmission and Distribution), vol. 138, no. 5, pp. 407-413, 1991. (Cit  en page 30.)
- [58] E. Pelikan, P. Matejka, M. Slama, and K. Vinkler, *Interactive forecasting of the electric load using Kohonen self-organizing feature maps*, In : International Neural Network Society Annual Meeting, Mahwah, NJ, USA. NJ : Lawrence Erlbaum Assoc, pp. 443-446, 1996. (Cit  en page 30.)
- [59] D. Fay, J.V. Ringwood, M. Condon, and M. Kelly, *24-h electrical load data-a sequential or partitioned time series ?*, In : Neurocomputing, vol. 55, no. 3-4, pp. 469-498, 2003. (Cit  en pages viii et 30.)
- [60] M. Sifuzzaman, M.R. Islam, and M.Z. Ali, *Application of Wavelet Transform and its Advantages Compared to Fourier Transform*, In : Journal of Physical Sciences, vol. 13, pp. 121-134, 2009. (Cit  en page 31.)
- [61] S.G. Mallat, *A theory of multiresolution signal decomposition : the wavelet representation*, In : IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, 1989. (Cit  en page 32.)
- [62] N. Smail, *Application de la transform e en ondelettes pour la recherche et l'analyse de textes*, Th se pour obtenir le grade de Docteur de l'Universit  Paris-Est, 2009. (Cit  en page 32.)

- [63] *The Mathworks : Leading developer of mathematical computing software*, Website="<http://www.mathworks.com/help/finance/examples/forecasting-corporate-default-rates.html>". (Cit  en pages viii et 37.)
- [64] OSE : Op rateur du Syst me Electrique, *Filiale du groupe SONELGAZ*, Site Web = "<http://www.ose.dz/>". (Cit  en pages 39 et 64.)
- [65] E. S. Gardner, *Exponential smoothing : The state of the art*, In : Journal of Forecasting, vol. 4, pp. 1-28, 1985. (Cit  en page 41.)
- [66] R.G. Brown, *Statistical forecasting for inventory control*, McGraw Hill, 1959. (Cit  en page 41.)
- [67] C.C. Holt, *Forecasting Trends and Seasonal by Exponentially Weighted Averages*, ONR Memorandum, no. 52, Carnegie Institute of Technology, Pittsburgh, USA (Published in International Journal of Forecasting, vol. 20, no. 1, pp. 5-13, 1957. (Cit  en page 41.)
- [68] P.R. Winters, *Forecasting sales by exponentially weighted moving averages*, In : Management Science, vol. 6, pp. 324-342, 1960. (Cit  en page 41.)
- [69] M. Mordjaoui, B. Boudjema, and M. Bouabaz, *Electric load consumption prediction using statistical approaches*, In : International Conference on Applied Analysis and Algebra (ICAAA), Istanbul, Turkey, 2012. (Cit  en page 41.)
- [70] K. LE, *Gestion optimale des consommations d' nergie dans les b timents*, Th se en cotutelle internationale pour obtenir le grade de Docteur de l'institut polytechnique de Grenoble et de l'Universit  de Danangle, 2008. (Cit  en page 43.)
- [71] F. Laouafi, *L'optimisation du plan de tension et la r partition optimale de la puissance r active par la m thode de newton, les r seaux de neurones et les algorithmes g n tiques*, M moire pr sent  en vue de l'obtention du dipl me de Magister en  lectrotechnique, Universit  20 Ao t 1955-Skikda, 2005. (Cit  en pages viii et 45.)
- [72] M. Parizeau, *Le perceptron multicouche et son algorithme de r tropropagation des erreurs*, Notes de cources, D partement de g nie  lectrique et de g nie informatique, Universit  Laval, Lien = "<http://wcours.gel.ulaval.ca/2010/h/IFT3901/default/5notes/RetroPerceptron.pdf>". (Cit  en page 46.)
- [73] L.A. Zadeh, *Fuzzy sets*, In : Information and Control, vol. 8, no. 3, pp. 338-353, 1965. (Cit  en page 48.)
- [74] M. Mordjaoui, *Mod lisation des effets  lectromagn tiques "Apport de la logique floue et neuro-floue"*, Th se pr sent e pour l'obtention du dipl me de Doctorat en Sciences, Universit  de Batna, 2008. (Cit  en page 51.)
- [75] J.S.R. Jang, *ANFIS : adaptive network based fuzzy inference system*, In : IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665-685, 1993. (Cit  en page 51.)
- [76] O.E.V. Dragomir, *Contribution au pronostic de d faillances par r seau neuro-flou : ma trise de l'erreur de pr diction*, Th se pr sent e pour l'obtention du

- grade de Doctorat en Automatique. Université de Franche-Comte, 2008. (Cité en page 55.)
- [77] A. Laouafi, M. Mordjaoui, and F. Laouafi, *An evaluation of conventional and computational intelligence methods for medium and long-term load forecasting in Algeria*, In : 3rd International Conference on Control, Engineering & Information Technology (CEIT), Tlemcen, Algeria, pp. 1-6, 2015. (Cité en page 61.)
- [78] J.W. Taylor, *Short-term electricity demand forecasting using double seasonal exponential smoothing*, In : Journal of the Operational Research Society, vol. 54, pp. 799-805, 2003. (Cité en page 65.)
- [79] A. Laouafi, M. Mordjaoui, F. Laouafi, and T.E. Boukelia, *Daily peak electricity demand forecasting based on an adaptive hybrid two-stage methodology*, In : Electrical Power and Energy Systems, vol.77, pp. 136-144, 2016. (Cité en page 66.)
- [80] A. Laouafi, M. Mordjaoui and D. Dib, *A Modified Holt-Winters Method for Daily Peak Load Forecasting in Algeria*, In : International Conference on Information Processing and Electrical Engineering (ICIPEE), Tebessa, Algeria, 2014. (Cité en page 66.)
- [81] Données de temperature, *Wilaya d'Alger*, Lien = "<http://www.wunderground.com/history/airport/DAAG> " (Cité en page 68.)
- [82] J.C. Dunn, *A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters*, In : Journal of Cybernetics, vol. 3, pp. 32-57, 1973. (Cité en page 69.)
- [83] J.C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, New York : Plenum Press, 1981. (Cité en page 69.)
- [84] M. Mordjaoui, B. Boudjema, and M. Bouabaz, *A hybrid wavelet transform and ANFIS model for short term electric load prediction*, In : 2nd International conference on advances in computational tools for engineering applications (ACTEA), Beirut, Liban, 2012. (Cité en page 72.)
- [85] C. Guan, P.B. Luh, L.D. Michel, Y. Wang, and P.B. Friedland, *Very short-term load forecasting : wavelet neural networks with data pre-filtering* , In : IEEE Transactions on Power Systems, vol. 28, no. 1, pp. 30-41, 2013. (Cité en page 72.)
- [86] G. Sudheer, and A. Suseelatha, *Short term load forecasting using wavelet transform combined with Holt-Winters and weighted nearest neighbor models*, In : International Journal of Electrical Power & Energy Systems, vol. 64, pp. 340-346, 2015. (Cité en page 72.)
- [87] J.P.S. Catalao, H.M.I. Pousinho, and V.M.F. Mendes, *Hybrid intelligent approach for short-term wind power forecasting in Portugal*, In : IET Renewable Power Generation, vol. 5, no. 3, pp. 251-257, 2011. (Cité en page 73.)

- [88] K. Y. Lee, Y. T. Cha, and J. H. Park, *Short-term load forecasting using an artificial neural network*, In : Transactions on Power Systems, vol. 7, no. 1, pp. 124-132, 1992. (Cit  en page 78.)
- [89] M. Mordjaoui, B. Boudjema, M. Bouabaz, and R. Daira, *Short term electric load forecasting using Neuro-fuzzy modeling for nonlinear system identification*, In : 3rd Conference on Nonlinear Science and Complexity, Ankara, Turkey, 28-31 July, 2010. (Cit  en page 78.)
- [90] RTE : Transmission System Operator of France, Website = "www.rte-france.com" (Cit  en page 83.)
- [91] AEMO : Australian Energy Market Operator, Website = "www.aemo.com.au" (Cit  en page 84.)
- [92] A. Laouafi, M. Mordjaoui, and D. Dib, *Very Short-Term Electricity Demand Forecasting using Adaptive Exponential Smoothing methods*, In : 15th international conference on Sciences and Techniques of Automatic control & computer engineering - STA'2014, pp. 553-557, Hammamet, Tunisia, 2014. (Cit  en pages 86, 88, 90 et 101.)
- [93] A. Laouafi, M. Mordjaoui, and D. Dib, *One-hour ahead electric load forecasting using neuro-fuzzy system in a parallel approach*, In : Computational intelligence applications in modeling and control, Studies in computational intelligence, vol. 575, pp. 95-121, 2015. (Cit  en page 86.)
- [94] A. Laouafi, M. Mordjaoui, and D. Dib, *One-hour ahead electric load and wind-solar power generation forecasting using artificial neural network*, In : 6th International renewable energy congress (IREC), Sousse, Tunisia, pp. 1-6, 2015. (Cit  en pages 86 et 90.)
- [95] A. Laouafi, M. Mordjaoui, S. Haddad, T.E. Boukelia, and A. Ganouche, *On-line electricity demand forecasting based on an effective forecast combination methodology*, In : Electric Power Systems Research, vol. 148, pp. 35-47, 2017. (Cit  en page 86.)
- [96] W. Charytoniuk, and M. S. Chen, *Very short-term load forecasting using artificial neural networks*, In : IEEE Transactions on Power Systems, vol. 15, no. 1, pp. 263-268, 2000. (Cit  en page 90.)
- [97] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, *One-hour-ahead load forecasting using neural network*, In : IEEE Transactions on Power Systems, vol. 17, no. 1, pp. 113-118, 2002. (Cit  en page 90.)
- [98] F. R. Hampel, *The influence curve and its role in robust estimation*, In : Journal of the American Statistical Association, vol. 69, no. 346, pp. 383-393, 1974. (Cit  en page 91.)
- [99] R.K. Pearson, *Outliers in Process Modeling and Identification*, In : IEEE Transactions on Control Systems Technology, vol. 10, no. 1, 2002. (Cit  en page 91.)
- [100] J.X. Che, and J.Z. Wang, *Short-term load forecasting using a kernel-based support vector regression combination model*, In : Applied Energy, vol. 132, pp. 602-609, 2014. (Cit  en page 92.)

- [101] M. Rana, and I. Koprinska, *Forecasting electricity load with advanced wavelet neural networks*, In : Neurocomputing, vol. 182, pp. 118-132, 2016. (Cité en page 100.)

Liste des publications

- [1] A. Laouafi, M. Mordjaoui, and D. Dib, *A Modified Holt-Winters Method for Daily Peak Load Forecasting in Algeria*, In: International Conference on Information Processing and Electrical Engineering (ICIPEE), Tebessa, Algeria, 2014.
- [2] A. Laouafi, M. Mordjaoui, and D. Dib, *Very Short-Term Electricity Demand Forecasting using Adaptive Exponential Smoothing methods*, In: 15th international conference on Sciences and Techniques of Automatic control & computer engineering - STA'2014, pp. 553-557, Hammamet, Tunisia, 2014.
- [3] A. Laouafi, M. Mordjaoui, and D. Dib, *One-hour ahead electric load forecasting using neuro-fuzzy system in a parallel approach*, In: Computational intelligence applications in modeling and control, Studies in computational intelligence, vol. 575, pp. 95-121, 2015.
- [4] A. Laouafi, M. Mordjaoui, and D. Dib, *One-hour ahead electric load and wind-solar power generation forecasting using artificial neural network*, In: 6th International renewable energy congress (IREC), Sousse, Tunisia, pp. 1-6, 2015.
- [5] A. Laouafi, M. Mordjaoui, and F. Laouafi, *An evaluation of conventional and computational intelligence methods for medium and long-term load forecasting in Algeria*, In: 3rd International Conference on Control, Engineering & Information Technology (CEIT), Tlemcen, Algeria, pp. 1-6, 2015.
- [6] A. Laouafi, M. Mordjaoui, F. Laouafi, and T.E. Boukelia, *Daily peak electricity demand forecasting based on an adaptive hybrid two-stage methodology*, In: Electrical Power and Energy Systems, vol.77, pp. 136-144, 2016.
- [7] T.E. Boukelia, M.S. Mecibah, and A. Laouafi, *Performance simulation of parabolic trough solar collector using two fluids (thermic oil and molten salt)*, In: Journal of Fundamental and Applied Sciences, vol. 8, no. 2, 2016.
- [8] M. Mordjaoui, A. Medoued, A. Laouafi, and S. Haddad, *Electric Load Forecasting By Using Dynamic Neural Network*, In: 4th European Conference on Renewable Energy Systems, Istanbul, Turkey, 2016.

LISTE DES PUBLICATIONS

- [9] A. Laouafi, M. Mordjaoui, S. Haddad, T.E. Boukelia, and A. Ganouche, *Online electricity demand forecasting based on an effective forecast combination methodology*, In: Electric Power Systems Research, vol. 148, pp. 35-47, 2017.
- [10] A. Laouafi, M. Mordjaoui, A. Medoued, T.E. Boukelia, and A. Ganouche, *Wind Power Forecasting Approach using Neuro-Fuzzy System Combined with Wavelet Packet Decomposition, Data Preprocessing, and Forecast Combination Framework*, In: Wind Engineering, 2017.

