

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des Sciences de l'Ingénieur
Département d'Informatique

Année Universitaire : 2011/2012

THESE

Présentée en vue de l'obtention du diplôme de *DOCTORAT en Informatique*

Méthodes Heuristiques pour la Prédiction des Séries Temporelles

Option
Informatique

Par
BENABBAS Farouk

Directeur de Thèse

KHADIR Mohamed Tarek

Prof. Univ. Badji Mokhtar, Annaba

Devant le Jury

Président : **BOUFAIDA** Mahmoud

Prof. Univ. Mentouri, Constantine

Examineurs :

SERIDI Hamid

Prof. Univ. 8 Mai 1945, Guelma

FARAH Nadir

Prof. Univ. Badji Mokhtar, Annaba

MOUSSAOUI Abdelouahab

MC. Univ. Ferhat Abbas, Sétif

Année Universitaire : 2011/2012

Remerciements

J'adresse ma plus sincère reconnaissance à monsieur Khadir Med Tarek, Professeur en informatique de l'université Badji Mokhtar Annaba pour sa disponibilité, ses encouragements, ses conseils, sa confiance et ses qualités de chercheur qui m'ont permis de mener à bon terme cette thèse.

Je tiens également à adresser mes plus vifs remerciements aux membres du jury qui m'ont fait l'honneur d'examiner l'ensemble de ces recherches.

Je remercie Monsieur BOUFAIDA professeur à l'université de Constantine, qui m'a fait l'honneur de présider le jury de soutenance de thèse.

Je remercie Monsieur Mr SERIDI Hamid, professeur à l'université de Guelma et Mr MOUSSAOUI, Maître de conférences à l'université de Sétif qui ont accepté d'évaluer ce travail et de participer au jury de soutenance.

Je remercie également Monsieur FARAH Nadir, professeur à l'université d'Annaba et directeur du laboratoire LABGED (Laboratoire de Gestion Electronique des documents) d'avoir accepté de participer à mon jury de thèse, pour ses encouragements et ses aides multiples.

Je tiens ces quelques lignes aux personnes sans qui cette aventure n'aurait vraisemblablement jamais commencé : Monsieur D.C. GILLES, professeur émérite au département informatique de l'université Glasgow, Ecosse, Grande Bretagne qui m'avait encadré pour mon mémoire de magistère et à monsieur SHARMA, docteur en informatique, enseignant au département d'informatique de l'université d'Annaba pour leur soutien pour aller loin dans mes études.

J'adresse ma gratitude à monsieur José RAGOT, professeur à l'INPL (Institut National Polytechnique de Lorraine) à l'université de Nancy pour m'avoir suivi lors de mes stages effectués au CRAN (Centre de Recherche en Automatique de Nancy) dans le cadre de l'accord Programme Franco-algérien, Tassili. Mes remerciements vont également à monsieur Gilles MOUROT, professeur à l'INPL pour ses remarques pertinentes et les moyens fournis lors de mes séjours. Je remercie vivement Mr HARKAT, Maître de conférences à l'université d'Annaba et Professeur DOGHMANE les principaux initiateurs du projet Tassili pour leur disponibilité et leur aide précieuse pour le bon déroulement des séjours au sein du CRAN.

Il m'est impossible d'omettre de saluer chaleureusement les doctorants du laboratoire du CRAN de Nancy : Dalil, Rodolfo, Yvon, Anca, Farah et Hichem et tous les autres collègues que je n'ai pas cité, pour l'ambiance conviviale qu'ils ont su entretenir et les joyeux moments passés à discuter ensemble.

J'ai vécu ces années de thèse comme une aventure passionnante. Ce temps dédié à la recherche m'a beaucoup apporté tant sur le plan scientifique que sur le plan humain. Je tiens à remercier, pour ces raisons, ma famille pour leur soutien et leurs encouragements. Trop nombreux sont ceux que je n'ai pu nommer, qu'ils trouvent ici l'expression de ma gratitude.

الملخص

إن تنبؤ الشحنة الكهربائية في الوقت الحاضر هو في غاية الأهمية سواء لأسباب اقتصادية أو بيئية فالكهرباء كمورد لا يمكن تخزينه يجب أن تستجيب لإنتاج يعادل الطلب الحالي لتخفيف توازن بين الإنتاج والاستهلاك، هذا ما يحدد من الخسائر ويضمن إمدادات مستمرة للمستهلكين، فأدوات التنبؤ يمكن لها أن توفر معرفة كمية الطاقة اللازمة للإنتاج، هذه الأدوات تركز على مناهج معروفة في مجال تحليل السلاسل الزمنية للتنبؤ، إن تطور تكنولوجيا المعلومات وإدراج العديد من المناهج الذكية والشبكات العصبية برهنت على قدرتها العالية فيما يخص تقريب الوظيفة والتصنيف.

ومع ذلك لا يمكن معالجة مشكلة التنبؤ بالشحنة الكهربائية إلا بعد دراسة وفهم دقيق لبيانات هذه الأخيرة، بالإضافة إلى ذلك فإن منحى الشحنة (الحمل) يختلف كثيرا من بلد لآخر، ومن ناحية أخرى في البلد نفسه لأنه يعتمد على العامل السوسيو اقتصادي، ثقافي وديني، فالتعرف على أنواع الأيام الكهربائية تصبح إذن ضرورة لتحديد مدى تأثير هذه العوامل على منحى الشحنة، وكذا الفهم الأحسن وتجزئة البيانات.

في الواقع، معرفة أنواع الأيام وعلاقتها مع احتياجات المستهلكين والذي يترجم بتأثير مباشر على منحى الشحنة (يومي، أسبوعي أو شهري) هو الهدف الأسمى، وعليه، وبعد دراسة مكثفة لأبحاث خصت مناهج التصنيف البسيطة والهجينة، مجموعة من النماذج الذكاء الاصطناعي وكأداة أساسية شبكات الذكاء الاصطناعي وخصوصا خرائط التنظيم الذاتي لكهونن (والمنطق الغامض تأخذ بعين الاعتبار، لقد تم الحصول على نتائج حاسمة ومبتكرة (K-Moyenne) وكذا التهجين مع (Kohonen) بالنسبة للشحنة الجزائرية الوطنية، وكدى الشحنة الجوية من خلال التركيز على مناطق الشمال والجنوب.

على أساس هذا التصنيف، أخذنا مثال لمنهج تنبؤ لإثبات مساهمة التصنيف العصبي، والتميز من خلال مناهج صعبة وغامضة. على أساس هذا التصنيف، أخذنا مثال لنظام خبير غامض لتنبؤ الشحنة الكهربائية لكل نوع من الأيام لمنهج متعدد النماذج.

في هذه الأطروحة، قمنا بإستراتيجية للتعرف لنوع الأيام وهذا باستعمال التصنيف الاتوماتيكي تعين الطريقة المستعملة لإيجاد المجموعات، على أساس هذا التصنيف، تم إعداد هذا النظام خبير الغامض لتنبؤ الشحنة الكهربائية.

Résumé

La prévision de la charge électrique est de nos jours d'une importance primordiale tant pour des raisons économique qu'environnementales. L'électricité en tant que ressource non stockable, doit de ce fait répondre à une production équivalente à la demande en cours, afin de réaliser un équilibre entre la production et la consommation ; ce qui limiterait les pertes et assurerait un approvisionnement continu des consommateurs. Des outils de prédictions peuvent assurer une connaissance a priori sur la quantité d'énergie requise à la production. Ces outils se basent sur des approches reconnues dans le domaine de l'analyse des séries temporelles pour prédire les valeurs antérieures de cette dernière. Avec l'évolution de l'informatique, et l'introduction de plusieurs approches intelligentes, les réseaux de neurones ont marqué leur puissance en termes d'approximation de fonction et de classification.

Toutefois, le problème de prédiction de la charge électrique ne peut être traité qu'après une étude et une compréhension exhaustive des données de cette dernière. De plus, la courbe de charge est très différente d'un pays à l'autre voir d'une région à l'autre dans un même pays car dépendant de paramètres socio-économique culturel, religieux etc. L'identification des types de journées électrique devient alors nécessaire afin de refléter l'influence de ses paramètres sur la courbe de charge et permettre une meilleure segmentation des données.

En effet la connaissance des types de jours ainsi que leur relation avec les besoins des consommateurs qui se traduit par un effet direct sur la courbe de charge (journalière, hebdomadaire, ou mensuelle), est primordiale. Pour cela, et après l'examen d'une recherche exhaustive sur les approches de classifications simple et hybrides, une palette de paradigmes de l'IA, avec comme outil principal les RNA et plus particulièrement les cartes auto-organisatrices de Kohonen, ainsi qu'une hybridation avec les K-moyenne et la logique floue a été considérée. Des résultats probants et novateurs, pour la charge Algérienne nationale ainsi que la charge régionale en mettant l'accent sur des zones du nord et du sud, ont été trouvés.

Sur la base de cette classification a été élaboré un système expert flou pour la prédiction de la charge électrique pour chaque type de jours, préambule à une approche multi-modèles.

Dans cette thèse, nous nous intéressons à une stratégie pour l'identification de types de jours en utilisant une classification automatique dans le cadre de la classification non supervisée. Quelle approche utiliser et quels sont les indices de validités pour pouvoir déceler les différentes classes ? Sur la base de cette classification a été élaboré comme exemple explicatif, un système expert flou pour la prédiction de la charge électrique

Mots clés : Classification automatique, prédiction, Cartes auto-organisées de Kohonen, K-moyenne, C-moyenne flou , Système expert flou, Perceptron Multi Couche.

Abstract

The prediction of the electric charge is nowadays of most importance both for economic reasons and environmental reasons. Electricity as a resource not be stored, must therefore answer a production equivalent to the current request, in order to achieve a balance between production and consumption, which would limit losses and ensure a continuous supply of consumers. Prediction tools can provide a priori knowledge on the amount of energy required for production. These tools are based on accepted approaches in the field of time series analysis to predict past values of the latter. With the evolution of computing and the introduction of several approaches to intelligent, neural networks have expressed their power in terms of function approximation and classification.

However, the problem of predicting the electric charge can be processed only after a thorough study and understanding of that data. In addition, the load curve is very different from one country to another to see a region to another in the same country as dependent on socio-economic parameters cultural, religious etc.. Identifying types of days supply becomes necessary to reflect the influence of its parameters on the load curve, and to better understand and segment our data.

Indeed knowledge of the types of days and their relationship with consumers' needs resulting in a direct effect on the load curve (daily, weekly, or monthly), is paramount. To do this, and after reviewing extensive research on approaches to simple classifications and hybrids, a range of paradigms of Artificial Intelligence (AI), as the primary tool with the Artificial Neural Network (ANN) and especially Kohonen self-organizing maps (SOM), as well as a hybridization with the K-Means and fuzzy logic was considered. Conclusive results and innovative support for Algerian national and regional load by focusing on areas north and south, were found.

Based on this classification is an example on realization of a model of short-term prediction (24 hours) using two paradigms of the Fuzzy logic and Expert systems for short terms Algerian load forecast led to appreciable results.

Keywords : Clustering, forecasting, Self-Organizing Maps Kohonen, K-Means, Fuzzy C-Means, Fuzzy expert system.

Table des matières

Chapitre 1 Introduction générale

Chapitre 2 Séries temporelles et analyse des données de la charge électrique nationale Algérienne

2.1 Introduction	5
2.2 Concepts de base des séries chronologiques	5
2.2.1 Caractéristiques des séries chronologiques	5
2.2.2 Les objectifs principaux	5
2.2.3 Composantes d'une série chronologique	6
2.2.4 Description schématique de l'étude complète d'une série chronologique	6
2.2.5 Prédiction les méthodes de lissage exponentielle	9
2.2.6 Processus stochastique	10
2.2.6.1 Processus stationnaire	10
2.2.6.2 Bruit blanc	11
2.2.6.3 Fonction d'auto-corrélation partielle	11
2.2.6.4 Opérateur de retard B et de différenciation Δ	11
2.2.7 Modèles pour processus stochastiques	12
2.2.7.1 Modèle autorégressif : Processus AR(p)	12
2.2.7.2 Modèle moyenne mobile : Processus MA(q)	12
2.2.7.3 Modèle autorégressif – moyenne mobile : ARMA(p,q)	13
2.2.7.4 Processus ARIMA et SARIMA	13
2.2.8 Identification et estimation des paramètres	14
2.2.8.1 Identification du modèle : Méthode de Box & Jenkins	14
2.2.8.2 Détermination du type et de l'ordre du modèle	14
2.2.8.3 Estimation des paramètres et prédiction	15
2.2.8.4 Validation	15
2.3 Concepts de la base de charge électrique	16
2.3.1 Introduction	16
2.3.2 Notion de charge électrique	16
2.3.3 La consommation et ses variations	17
2.3.4 Méthodes de prédiction pour le lendemain	18
2.4 Analyse des données de la charge électrique algérienne	18
2.4.1 Données de la charge électrique algérienne	18
2.4.2 Prétraitement des données de la charge électrique nationale Algérienne	20
2.4.3 Caractéristiques de la charge électrique nationale Algérienne	21
2.5 Conclusion et contributions	24

Chapitre 3 Classification non supervisée : Etat de l'art

3.1 Introduction	26
3.2 Concepts de base de la classification automatique	26

3.2.1 Les données	26
3.2.2 Les groupes d'objets	26
3.2.3 Partition d'un groupe d'objets	29
3.2.4 Partitionnement dur ou stricte	29
3.2.5 Partitionnement flou	29
3.3 Algorithme classiques de classification	30
3.3.1 Algorithme de classification hiérarchique	30
3.3.2 Les Nuées dynamiques	32
3.3.2.1 L'algorithme des K-moyennes	32
3.3.2.2 L'algorithme des C-moyennes floues	32
3.3.3 Approche à bases de réseaux de neurones : les cartes auto-organisées de Kohonen	34
3.4 Approche de classification à deux niveaux	36
3.5 Evaluation et critère de validité	36
3.5.1 Indices propres aux cartes-organisées de Kohonen	37
3.5.1.1 Erreur de quantification	37
3.5.1.2 Taux d'erreurs topologiques	37
3.5.2 Indices propres à la classification dure	38
3.5.2.1 Erreur Quadratique Moyenne	38
3.5.2.2 Indice de Dunn	38
3.5.2.3 Indice de Davies-Bouldin	39
3.5.3 Indices propres à la classification floue	39
3.5.3.1 Coefficient de Partition	39
3.5.3.2 Entropie de classification	39
3.5.3.3 Indice de Partition	40
3.5.3.4 Indice de Xie et Beni	40
3.6 Résumé et Conclusion	40

Chapitre 4

Identification par type de jour de la charge électrique nationale Algérienne

4.1 Introduction	43
4.2 Approche de segmentation de la charge électrique	43
4.3 Identification par type de jour de la charge électrique nationale Algérienne avec les les cartes de Kohonen	44
4.4 Identification à deux niveaux par type de jour de la charge électrique nationale Algérienne : Approche à deux niveaux : SOM et K-moyenne	48
4.5 Conclusion et contributions	52

Chapitre 5

Identification par type de jour de la charge électrique régionale Algérienne

5.1 Introduction	54
5.2 Choix des paramètres et critères de qualités de SOM	54
5.3 Identification à deux niveaux par type de jour de la charge électrique régionale Algérienne.....	56
5.3.1 Approche à deux niveaux avec la classification dure	56
5.3.2 Approche à deux niveaux avec la classification floue	60

5.4 Conclusion et contributions	65
---------------------------------------	----

Chapitre 6

La prédiction à court terme de la charge électrique Algérienne

6.1 Introduction	68
6.2 Prévission de la charge électrique	68
6.2.1 Types de prévision de charge	68
6.2.2 Facteurs influençant la charge électrique	69
6.2.3 Les méthodes de prévision de charge	70
6.2.3.1 Méthode de régression	70
6.2.3.2 Système expert	70
6.2.3.3 Méthode basée sur les séries chroniques stochastiques	71
6.2.3.4 Les techniques non-linéaires de la charge électrique	73
6.3 Méthode basée sur les systèmes experts flous.....	73
6.3.1 Système expert flou	73
6.3.2 Moteurs d'inférences pour système flou	75
6.3.3 Génération automatique de la base des règles	77
6.4 Méthode basée sur les réseaux de neurones	78
6.4.1 Méthode des réseaux de neurones, pourquoi ?	78
6.4.2 Théorie des réseaux de neurones artificiels	79
6.4.2.1 Modèle des réseaux de neurones (RNA)	79
6.4.2.2 Apprentissage	82
6.4.2.3 Réseau rétro-propagation	83
6.5 Cas d'application : Approche de prédiction par système expert flou	85
6.5.1 Méthodologie de conception du programme	85
6.5.2 Préparation des données initiales	86
6.5.3 Choix de la base	86
6.5.4 Architecture du système projeté	86
6.5.5 Apprentissage	87
6.5.6 Validation	88
6.5.7 Test	88
6.5.8 Résultats de calcul	88
6.6 Conclusion	95

Chapitre 7

Conclusion générale et Perspectives

Bibliographie

Table des figures

Figure 2.1 Exemple de charge électrique	16
Figure 2.2 Carte géographique de l'Algérie avec les différentes régions (Oran, Alger, Hassi Messaoud et Ain Amenas) en (a) et Carte géographique de l'Algérie en Afrique du Nord (b)	18
Figure 2.3 Charge électrique Algérienne pour la période 2000-2004	20
Figure 2.4 Charge avec une courbe de tendance approximative	20
Figure 2.5 Charge électrique d'une journée de travail en été et en hiver	22
Figure 2.6 Charge électrique hebdomadaire	22
Figure 2.7 Charge électrique par région pour un jour particulier (Samedi 18 Mars 2000)	23
Figure 3.1 Quelques formes de classes : (a) En forme de boules . (b) Formes étirées (c) Compacte mais pas bien séparées	26
Figure 3.2 Exemple de classification automatique (trois classes détectées jaune, bleu et rouge)	27
Figure 3.3 Classification ascendante hiérarchique et descendante hiérarchique	31
Figure 3.4 Structure de la carte de Kohonen	34
Figure 3.5 Une première abstraction est obtenue en créant un ensemble de prototypes en utilisant les cartes auto-organisées. Le second niveau d'abstraction classe ensuite les prototypes obtenus	36
Figure 4.1 Nœuds déclenchés pour les jours de l'année 2000 à 2004	46
Figure 4.2 Identification des types de jours pour une semaine (Samedi à Vendredi)	47
Figure 4.3 Identification des types de jours saisonnière (Janvier à Décembre)	48
Figure 4.4 Figure 4.4 – Evolution de l'indice <i>DB</i> minimum 0.75574 atteint pour 6 classes (a), Segmentation de SOM avec k-means en 6 classes (b) et Nombre d'items par classe pour Algérie (c)	50
Figure 4.5 Distribution hebdomadaire des classes de l'Algérie	51
Figure 4.6 Distribution mensuelle des classes de l'Algérie	51
Figure 5.1 Evolution de l'indice <i>DB</i> minimum 0.8540 atteint pour 4 classes (a), segmentation de SOM avec k-means en 4 classes (b) et Nombre d'items par classe pour Alger (c)	58
Figure 5.2 Evolution de l'indice <i>DB</i> minimum 0.7929 atteint pour 3 classes (a), Segmentation de SOM avec k-means en 3 classes (b) et Nombre d'items par classe pour Hassi Messaoud (c)	58
Figure 5.3 Distribution hebdomadaire des classes de la ville d'Alger	59
Figure 5.4 Distribution mensuelle pour Alger	59
Figure 5.5 Distribution hebdomadaire des classes de la ville de Hassi Messaoud	60
Figure 5.6 Distribution mensuelle pour Hassi Messaoud	60
Figure 5.7 Résultats graphiques de la classification floue pour Alger	63
Figure 5.8 Résultats graphiques de la classification floue pour Hassi Messaoud	63
Figure 6.1 Diagramme d'un système expert basé sur la logique floue	75
Figure 6.2 Schématisation d'un neurone biologique	79
Figure 6.3 Modèle d'un neurone artificiel	80
Figure 6.4 Modèle d'une couche de <i>n</i> neurones (à gauche) et Représentation matricielle d'une couche à <i>S</i> neurones (à droite)	81
Figure 6.5 Types d'apprentissage en RNA	83
Figure 6.6 Architecture d'un réseau rétro-propagation	83
Figure 6.7 Détermination et application d'un réseau rétro-propagation	84

Figure 6.8 Architecture du système	87
Figure 6.9 Valeurs réelles et prédites en phase validation pour l'automne	89
Figure 6.10 Erreur en pourcentage, concernant la série de l'automne en phase de validation	89
Figure 6.11 Valeurs réelles et prédites en phase test pour l'automne	90
Figure 6.12 Erreur en pourcentage, concernant la série de l'automne en phase de test	90
Figure 6.13 Valeurs réelles et prédites en phase validation pour l'été	90
Figure 6.14 Erreur en pourcentage, concernant la série de l'été en phase de validation	91
Figure 6.15 Valeurs réelles et prédites en phase de test pour l'été	91
Figure 6.16 Erreur en pourcentage, concernant la série de l'été en phase de test	91
Figure 6.17 Valeurs réelles et prédites en phase validation pour l'hiver	92
Figure 6.18 Erreur en pourcentage, concernant la série de l'hiver en phase de validation	92
Figure 6.19 Valeurs réelles et prédites en phase de test pour l'hiver	92
Figure 6.20 Erreur en pourcentage, concernant la série de l'hiver en phase de test	93
Figure 6.21 Valeurs réelles et prédites en phase validation pour le printemps	93
Figure 6.22 Erreur en pourcentage, concernant la série du printemps en phase de validation.	93
Figure 6.23 Valeurs réelles et prédites en phase de test pour le printemps	94
Figure 6.24 Erreur en pourcentage, concernant la série du printemps en phase de test	94

Liste des tableaux

Tableau 2.1 Les données de la charge électrique (Période, intervalle de temps et nombre valeurs)	19
Tableau 2.2 Coefficients de la courbe de la tendance	21
Tableau 3.1 Mesure de dissimilarité par le calcul de la distance des centres des classes	28
Tableau 4.1 Nombre de classe et indice de Davies-Bouldin	49
Tableau 4.2 Nombre d'items par classes pour l'Algérie	49
Tableau 4.3 Distribution hebdomadaire par classes pour l'Algérie	50
Tableau 4.4 Distribution mensuelle par classes pour l'Algérie	50
Tableau 5.1 Paramètres de SOM	55
Tableau 5.2 Mesures de qualités de SOM	55
Tableau 5.3 Nombre de classes et indices de Davies-Bouldin	57
Tableau 5.4 Nombre d'items par classes pour Alger	57
Tableau 5.5 Nombre d'items par classe pour Hassi Messaoud	57
Tableau 5.6 Distribution hebdomadaire par classes pour Alger	58
Tableau 5.7 Distribution mensuelle par classe pour Alger	58
Tableau 5.8 Distribution hebdomadaire par classe pour Hassi Messaoud	59
Tableau 5.9 Distribution mensuelle par classe pour Hassi Messaoud	59
Tableau 5.10 Valeurs de c préférées	61
Tableau 5.11 Résultats de la classification floue pour les quatre classes d'Alger	62
Tableau 5.12 Résultats de la classification floue pour les quatre classes de Hassi Messaoud.	62
Tableau 5.13 Jour représentatif et degré d'appartenance par classe pour Alger	64
Tableau 5.14 Jour représentatif et degré d'appartenance par classe pour Hassi Messaoud ...	64
Tableau 5.15 Classe identifiée pour Alger	64
Tableau 5.16 Classe identifiée pour Hassi Messaoud	65
Tableau 5.17 Valeurs d'appartenances par classe et par jour pour Alger	65
Tableau 5.18 Valeurs d'appartenances par classe et par jour pour Hassi Messaoud	65
Tableau 6.1 Les différents modèles de Box-Jenkins et polynômes associés ($F = 1, D = 1$)	71
Tableau 6.2 Fonction de transfert	82
Tableau 6.3 Tableau des paramètres d'apprentissage choisis	88
Tableau 6.4 Tableau des résultats de la phase de validation	94
Tableau 6.5 Tableau des résultats de la phase de test	95

Références personnelles

Revue internationale avec comité de lecture

1. Benabbas F., Khadir M.T. (2011). Fuzzy c-means clustering and kohonen maps for the indentification of regional electricity load day types. International Journal of Hybrid Intelligent Systems, 8(2), 81-92.
2. Benlembarek K., Khadir M.T., Benabbas F. (2010). A Web Based System for Short-Term Forecasting of Algerian Electricity Load Using Artificial Neural Network, Journal of Automation and Systems Engineering, 4(2), 94-100.

Chapitre dans un livre

1. Khadir M. T., Khedairia, S., Benabbas F. (2010). Kohonen Maps Combined to K-means in a Two Level Strategy for Time Series Clustering Application to Meteorological and Electricity Load data. Self-Organizing Maps, In Tech, 63-80.
2. Khadir M. T., Benabbas F. (2011). Kohonen Maps Combined to Fuzzy C-means, a Two Level Clustering Approach. Application to Electricity Load Data". Self Organizing Maps Applications and Novel Algorithm Design, InTech, 541-558.

Conférences internationales avec comité de lecture :

1. Benabbas F., Khadir M.T., Fay D., Boughrira A. (2007). Regional Day Type Identification For Algerian Electricity Load Using Kohonen Map. In Proceedings of The International Conference of Neural Networks, EANN 2007, Thessaloniki, Greece, 29-31 aout.
2. Benabbas F., Khadir M.T., Fay D., Boughrira A. (2008). Cartes de Kohonen Combiné à L'algorithme des K-Means pour l'identification des Types de Jours de la Charge Electrique Algerienne. Proceedings de la Conférence Internationale francophone d'automatique, CIFA 2008, Bucarest, Roumanie, 3-5 septembre.
3. Benabbas F., Khadir M.T., Fay D., Boughrira A. (2008). Kohonen Map Combined to the K-Means Algorithm for the Identification of Day Types of Algerian Electricity Load. Proceedings of The International Conference of Computer Information Systems and Industrial Management Applications, CISIM 2008, Ostrava, Czech Republic, 26-28 juin.
4. Benabbas F., Khadir M.T., Fay D., Boughrira A. (2008). Kohonen Map Combined to the K-Means Algorithm for the Identification of Day Types of Regional Algerian Electricity Load. Proceedings of Second International Conference on Electrical Engineering Design and Technologies, ICEEDT 2008, Hammamet, 8-10 novembre, Tunisia.

1

Introduction

En commun avec de nombreux pays, le marché de l'électricité en Algérie est entrain d'être dérégulé depuis 2006. Cette dérégulation pousse les fournisseurs d'énergie à penser à produire des prévisions exactes à court terme. Dans de tels environnements concurrentiels, l'outil de prédiction devient un outil puissant pour atteindre de meilleures performances. La prédiction à court terme est aussi nécessaire par l'opérateur national pour déterminer les puissances optimales des stations et déterminer l'excès de production nécessaire pour garantir la qualité de l'offre.

Cependant, la conception et la validité d'un modèle donné est souvent sujette à priori à une bonne connaissance et compréhension du comportement réel de la charge électrique. Dans le cas d'une prédiction à court terme, il a été prouvé que le type de jours ou les habitudes quotidiennes des consommateurs pour différentes périodes de temps comme les jours de travail, weekends ou jours de vacances affectent sensiblement la courbe de la charge.

L'identification de ces informations appelé identification de types de jours doit être incluse lors des étapes de conception en segmentant les données et modélisant chaque type de jour séparément ou en incluant le type de jour comme entrée, ceci nous mène à la classification automatique des données et la création des classes - *clusters*.

La classification automatique des données – *clustering* est une étape importante dans l'identification des différentes classes de la charge électrique algérienne. Elle vise à découvrir la structure intrinsèque d'un ensemble d'objet en formant des regroupement ou classes qui partagent des caractéristiques similaires. Elle ne doit pas être confondue avec la classification supervisée ou classement – *classification* – qui consiste à déterminer les règles qui ont permis de séparer un ensemble d'individus en classes connues à priori.

Dans cette thèse, nous nous intéressons à une stratégie pour l'identification de types de jours en utilisant une classification automatique dans le cadre de la classification non supervisée. Quelle approches utiliser et quels sont les indices de validités pour pouvoir déceler les différentes classes ? Sur la base de cette classification a été élaboré, comme exemple applicatif, un système expert flou pour la prédiction de la charge électrique pour chaque type de jours, préambule à une approche multi-modèles.

Par rapport à ces différents objectifs, nous avons structuré cette thèse en six chapitres.

Dans le chapitre 2, la notion de séries temporelle et l'analyse des caractéristiques de la charge électrique algérienne sont présentés.

Le troisième chapitre est une analyse de l'état de l'art de la classification automatique. Nous présentons les notions de base, les différentes approches avec leurs avantages et inconvénients ainsi que l'évaluation de la classification à l'aide des indices de validités. Ce chapitre met en évidence l'apport de l'approche hybride pour une classification moins couteuse en temps en particulier pour une masse importante d'informations.

Le quatrième chapitre est consacré à deux approches proposées concernant l'identification de types de jour de la charge électrique algérienne. La première utilise les cartes auto-organisées SOM - *Self-Organizing Maps* - de Kohonen. La seconde approche hybride combine les cartes auto-organisées et l'algorithme des K-moyennes – *K-means*. La première phase crée les prototypes puis une segmentation de la carte est faite avec les K-moyennes. Le nombre optimal de classes est déduit en utilisant l'indice de validité de Davies-Bouldin.

Le cinquième chapitre concerne l'identification de types jours de la charge électrique régionale algérienne. Deux approches sont présentées, la première combine les cartes SOM avec l'algorithme des K-moyennes, la seconde les cartes SOM avec l'approche C-Moyenne floue – *Fuzzy C-Means* -.

La validation du regroupement pour la classification dure ou floue est effectuée à l'aide d'indices de validités adéquats. La visualisation des classes est réalisée à l'aide de l'ACP (Analyse en Composantes Principales). Une comparaison entre les deux approches est établie.

Le sixième chapitre traite de l'apport de l'identification de types de jours pour la prédiction de la charge électrique algérienne par une étude de cas. Cet exemple utilise un système expert basé sur la logique flou avec génération automatique de règle d'inférence de type Mamdani (Mamdani, 1977).

A la fin de la thèse, nous présenterons les conclusions et les perspectives de cette étude.

2

Séries temporelles et analyse des données de la charge électrique nationale Algérienne

Sommaire

2.1 Introduction	5
2.2 Concepts de base des séries chronologiques	5
2.2.1 Caractéristiques des séries chronologiques	5
2.2.2 Les objectifs principaux	5
2.2.3 Composantes d'une série chronologique	6
2.2.4 Description schématique de l'étude complète d'une série chronologique	6
2.2.5 Prévission pour les méthodes de lissage exponentiel	9
2.2.6 Processus stochastique	10
2.2.6.1 Processus stationnaire	10
2.2.6.2 Bruit blanc	11
2.2.6.3 Fonction d'auto-corrélation et d'auto-corrélation partielle	11
2.2.6.4 Opérateur de retard B et différenciation Δ	11
2.2.7 Modèles pour processus stochastiques	12
2.2.7.1 Modèle autorégressif : Processus AR(p)	12
2.2.7.2 Modèle moyenne mobile : Processus moyenne mobile MA(q)	12
2.2.7.3 Modèle autorégressif – moyenne mobile : ARMA(p,q)	13
2.2.7.4 Processus ARIMA et SARIMA	13
2.2.8 Identification et estimation des paramètres	14
2.2.8.1 Identification du modèle : Méthode de Box & Jenkins	14
2.2.8.2 Détermination du type et de l'ordre du modèle	14
2.2.8.3 Estimation des paramètres et prévision	15
2.2.8.4 Validation	15
2.3 Concepts de bases de la charge électrique	16
2.3.1 Introduction	16
2.3.2 Notion de charge électrique	16
2.3.3 La consommation, ses variations et les facteurs l'influençant	17
2.3.4 Méthode de prévisions pour le lendemain	18
2.4 Analyse des données de la charge électrique Algérienne	18
2.4.1 Données de la charge électrique Algérienne	18
2.4.2 Prétraitement des données de la charge électrique nationale Algérienne	21
2.4.3 Caractéristiques de la charge électrique nationale Algérienne	22
2.5 Conclusion et contributions	24

2.1. Introduction

Dans ce chapitre, nous allons décrire les concepts de base des séries chronologiques. Nous mettrons l'accent plus particulièrement sur la méthodologie utilisée dans le traitement des séries chronologiques, l'évaluation de la qualité des données et les transformations nécessaires avant toute utilisation. Nous décrirons par la suite les concepts de la charge électrique. Nous présenterons l'analyse proprement dite à la charge électrique algérienne qui est primordiale pour l'identification de types de jours nécessaire à la construction d'un modèle approprié. Cette analyse inclut au chapitre 2.4.1, la présentation des données fournies par l'opérateur économique National SONELGAZ puis au chapitre 2.4.2 les phases de prétraitement nécessaires avant toute utilisation. Les caractéristiques des courbes spécifiques à la charge électrique sont présentées au paragraphe 2.4.3 Nous terminerons ce chapitre par une conclusion.

2.2. Concepts de base des séries chronologiques

2.2.1. Caractéristiques des séries chronologiques

La théorie des séries chronologiques (ou temporelles) abordée ici est appliquée de nos jours dans des domaines aussi variés que l'économie, la médecine ou la démographie, pour n'en citer qu'une partie. On s'intéresse à l'évolution au cours du temps d'un phénomène, dans le but de décrire, expliquer puis prévoir ce phénomène dans le futur. On dispose ainsi d'observations à des dates différentes, c'est-à-dire une suite de valeurs numériques indicées par le temps. La température d'une région donnée, la consommation électrique sont des exemples de séries chronologiques ou temporelles.

Cette suite d'observations d'une famille de variables aléatoires réelles notées $(X_t)_{t \in \Theta}$ est appelée série chronologique (ou temporelle). Dans la suite de cette thèse, nous la noterons $(X_t)_{t \in \Theta}$ ou $(X_t, t \in \Theta)$

Où l'ensemble Θ est appelé espace des temps qui peut être :

- discret (cas du nombre de voyageurs d'un vol d'avion).
- continu (signal radio, résultat d'un électrocardiogramme).

Dans ce qui suit, nous considérerons uniquement des processus stochastiques à temps discret et unidimensionnels.

2.2.2. Les objectifs principaux

L'étude d'une série chronologique permet d'analyser, de décrire et d'expliquer un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décision (marketing). L'un des objectifs principaux de l'étude d'une série chronologique est la prévision qui consiste à prévoir les valeurs futures X_{T+h} ($h = 1, 2, 3, \dots$) de la série chronologique à partir de ses valeurs observées jusqu'au temps : X_1, X_2, \dots, X_T .

Parmi les autres objectifs avoués de l'étude des séries temporelles, figure le problème de l'estimation d'une tendance ou l'impact d'un événement sur une variable.

2.2.3. Composantes d'une série chronologique

On considère qu'une série chronologique est la résultante de différentes composantes fondamentales :

La **tendance** (ou trend) représente l'évolution à long terme de la série étudiée. Elle traduit le comportement « moyen » de la série.

La **composante saisonnière** (ou saisonnalité) correspond à un phénomène qui se répète à intervalles de temps réguliers (périodiques). En général, c'est un phénomène saisonnier d'où le terme de variations saisonnières.

La **composante résiduelle** (ou bruit ou résidu) correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire. On parle aussi d'aléas.

Des **phénomènes accidentels** (grèves, conditions météorologiques exceptionnelles, crash financier) peuvent notamment intervenir.

Dans ce qui suit, la composante correspondant aux phénomènes accidentels sera intégrée aux fluctuations irrégulières de la série et la composante tendance regroupera à la fois la tendance et le cycle.

2.2.4. Description schématique de l'étude complète d'une série chronologique

Comme nous venons de le voir, l'un des objectifs principaux de l'étude d'une série chronologique est la prévision des valeurs futures de cette série. Pour cela, on a besoin de connaître ou tout au moins de modéliser le mécanisme de production de la série chronologique.

Schématiquement, les principales étapes de traitement d'une série chronologique sont les suivantes :

- Correction des données
- Observation de la série
- Modélisation
- Analyse de la série à partir de ces composantes
- Prédiction (= prévision)

Avant de se lancer dans l'étude d'une série chronologique, il est souvent nécessaire de traiter, modifier les données brutes.

Par exemple,

- Evaluation de données manquantes, remplacement de données accidentelles.
- Standardisation afin de se ramener à des intervalles de longueur fixes. Par exemple, pour des données mensuelles, on se ramène au mois standard en calculant la moyenne journalière sur le mois (total des observations sur le mois divisé par le nombre de jours du mois) ;

- Transformation des données : pour des raisons diverses, on peut être parfois amenés à utiliser des données transformées. Par exemple en économie, on utilise la famille de transformations de Box-Cox (Box et Cox, 1964) : $Y_t = \frac{1}{\lambda} [(X_t)^\lambda - 1]$, $\lambda \in \mathbb{R}^*$

Une règle générale en Statistique descriptive consiste à commencer par analyser les données avant d'effectuer le moindre calcul. Ainsi, une fois la série corrigée et prétraitée, on trace son graphique c'est à dire la courbe de coordonnées (t, X_t) . L'observation de ce graphique est souvent une aide à la modélisation de la série chronologique et permet de se faire une idée des différentes composantes de la série chronologique.

De manière générale, la courbe peut indiquer un mouvement à moyen terme de croissance ou décroissance (linéaire, quadratique...) révélant la présence d'une composante déterministe dans la série appelée tendance (ou trend) qui exprime donc l'évolution générale à moyen ou long terme de la série, du phénomène étudié. Par exemple, si on admet le scénario d'un réchauffement de la planète, la courbe des températures moyennes indique un mouvement de croissance à moyen terme.

Un **modèle** est une image simplifiée de la réalité qui vise à traduire les mécanismes de fonctionnement du phénomène étudié et permet de mieux les comprendre. Un modèle peut être meilleur qu'un autre pour décrire la réalité et bien sur, plusieurs questions se posent alors : comment mesurer cette qualité ? Comment diagnostiquer un modèle ? Nous présentons dans cette section une petite liste qui sert à résumer et classer les différents modèles envisagés dans ce cours. On distingue principalement deux types de modèles : les modèles déterministes et stochastiques. Les **modèles déterministes** relèvent de la Statistique Descriptive. Ils ne font intervenir que de manière sous-jacente le calcul des probabilités et consistent à supposer que l'observation de la série à la date t est une fonction du temps t et d'une variable ε_t centrée faisant office d'erreur au modèle, représentant la différence entre la réalité et le modèle proposé :

$$X_t = f(t, \varepsilon_t) \quad (2.1)$$

Les deux modèles de ce type les plus utilisés sont les suivants

- **Le modèle additif.** C'est le modèle classique de décomposition dans le traitement des modèles d'ajustement. La variable X_t s'écrit comme la somme de trois termes :

$$X_t = Z_t + S_t + \varepsilon_t \quad (2.2)$$

Où Z_t représente la tendance (déterministe), S_t la saisonnalité (déterministe aussi) et ε_t les composantes (erreurs au modèle) aléatoire.

- **Le modèle multiplicatif.** La variable X_t s'écrit au terme d'erreur près comme le produit de la tendance et d'une composante de saisonnalité :

$$X_t = Z_t(1 + S_t) (1 + \varepsilon_i) \quad (2.3)$$

L'ajustement est multiplicatif et intervient dans les modèles GARCH (Brockwell et Davis, 2009).

- **Les modèles mixtes** : Il s'agit là des modèles où addition et multiplication sont utilisées. On peut supposer, par exemple, que la composante saisonnière agit de façon multiplicative, alors que les fluctuations irrégulières sont additives. Le modèle est de la forme :

$$X_t = Z_t(1 + S_t) + \varepsilon_t \quad (2.4)$$

Les **modèles stochastiques** sont de même type que les modèles déterministes à ceci près que les variables de bruit ε_t ne sont pas indépendantes et identiquement distribuées (iid) mais possèdent une structure de corrélation non nulle : ε_t est une fonction des valeurs passées plus ou moins lointaines et d'un terme d'erreur η_t .

$$\varepsilon_t = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \eta_t) \quad (2.5)$$

- La classe des modèles de ce type la plus fréquemment utilisée est la classe SARIMA (et de ses sous-modèles ARIMA, ARMA) (Brockwell et al., 2009). Comme vu plus haut, la série chronologique est l'observation d'un processus stochastique : la modélisation porte ici sur la forme du processus (ε_t).
- Les cas particuliers où la relation fonctionnelle g est linéaire est très important et très usité. Il mène aux modèles autorégressifs linéaires, par exemple un modèle d'ordre 2 avec des coefficients autorégressifs (a_1, a_2) est donné par

$$\varepsilon_t = a_1 X_{t-1} + a_2 X_{t-2} + \eta_t \quad (2.6)$$

Où (η_t) est un bruit blanc c'est-à-dire une variable aléatoire de moyenne nulle non corrélée.

Les deux types de modèles ci-dessus induisent des techniques de prévision bien particulières. Schématiquement, on s'intéresse tout d'abord à la tendance et la saisonnalité éventuelle que l'on isole tout d'abord. Une fois un modèle est obtenu, il peut être utilisé pour la prédiction des valeurs futures.

2.2.5. Prévision par les méthodes de lissage exponentiel

On présentera dans ce qui suit trois types de lissage classiques : la méthode de lissage simple puis double et la méthode de Holt-Winters (Gelper et al., 2010).

Soit une série chronologique $X_1, X_2, X_3, \dots, X_{T-2}, X_{T-1}, X_T$. Nous sommes à la période T et nous voulons prédire X_{T+h} , où h est l'horizon de prévision. Pour ce faire, on fera appel à une méthode appelée lissage exponentiel simple. Ce type de lissage permet d'effectuer des prévisions pour des séries chronologiques dont la tendance est constante et sans saisonnalité.

$$X_{T+h} = (1 - \lambda) \sum_{j=0}^{T-1} \lambda^j X_{T-j} \quad (2.7)$$

On peut choisir λ arbitrairement selon que l'on veut une prévision réactive ou rigide ; on peut également choisir la valeur de λ qui minimise la somme des carrés des erreurs de prévision quand on applique la méthode aux données passées. λ est alors la solution du problème de minimisation :

Trouver λ tel que la relation 2.8 soit minimale :

$$\sum_{t=1}^{T-1} \left[X_{t+1} - \lambda \sum_{j=0}^{t-1} (1-\lambda)^j X_{t-j} \right]^2 \quad (2.8)$$

Le lissage exponentiel double généralise l'idée du lissage exponentiel simple au cas où la série peut être ajustée par une droite au voisinage de T . Il existe donc une tendance linéaire : $f_t = a + bt = a_1 + (t - T) a_2$. On cherche les valeurs de a_1 et a_2 qui minimisent :

$$\sum_{j=0}^{T-1} (1-\lambda)^j (X_{T-j} - a_1 + a_2 j)^2 \quad (2.9)$$

Une solution approchée est :

$$A_1(T) = 2 S_1 - S_2 \quad \text{et} \quad A_2(T) = \frac{\lambda}{1-\lambda} (S_1 - S_2) \quad (2.10)$$

avec

$$S_1 = \lambda \sum_{j=0}^{T-1} (1-\lambda)^2 X_{T-j} \quad \text{et} \quad S_2 = \lambda^2 \sum_{j=0}^{T-1} j (1-\lambda)^j X_{T-j} + \lambda S_1 \quad (2.11)$$

La méthode de lissage exponentiel double permet de traiter des séries présentant une tendance linéaire mais sans saisonnalité. On peut également définir des lissages exponentiels généralisés sur le même principe que les techniques décrites dans les sections précédentes permettant de traiter des séries avec saisonnalité. Ainsi la méthode de Holt-Winters a été introduite. Deux méthodes existent, l'une dite non saisonnière et l'autre saisonnière additive. Dans le premier cas, il s'agit d'un lissage exponentiel double avec des relations de mises à jour modifiées qui utilisent des paramètres de réglage permettant de choisir une prévision plus ou moins réactive. Pour la première méthode, on a :

$$A_1(T) = (1-\alpha) [A_1(T-1) + A_2(T-2)] + \alpha X_T \quad (2.12)$$

$$A_2(T) = (1-\gamma) A_2(T-1) + \gamma [A_1(T) - A_1(T-1)] \quad (2.13)$$

Le choix de α et γ est arbitraire dans l'intervalle $[0, 1]$; si on veut une prévision réactive, on les choisira proches de 1.

Dans le second cas, On considère le modèle suivant :

$$X_t = a_1 + a_2 (t - T) + s_t + E_t \quad (2.14)$$

Où s_t est la composante saisonnière et E_t est une variable aléatoire représentant l'erreur, a_1 et a_2 sont des paramètres. A la différence du modèle linéaire, les paramètres a_1 et a_2 changent à chaque instant t . Les formules de mise à jour sont :

$$A_1(T) = (1 - \alpha) [A_1(T - 1) + A_2(T - 2)] + \alpha(X_T - S_{T-\zeta}) \quad (2.15)$$

$$A_2(T) = (1 - \gamma)A_2(T - 1) + \gamma [A_1(T) - A_1(T - 1)] \quad (2.16)$$

$$S_T = (1 - \delta)S_{T-\zeta} + \delta[X_T - A_1(T)] \quad (2.17)$$

2.2.6. Processus stochastique

On appellera ici processus stochastique (à temps discret) une suite X_t de variables aléatoires indicée par le temps. Il faut bien noter que les X_t ne sont pas forcément indépendants les uns des autres et que X_t dépend de t . La notion de processus généralise celle de variable aléatoire : la réalisation d'un processus est une série chronologique (on emploie également le terme de trajectoire).

2.2.6.1. Processus stationnaire

Un processus est dit stationnaire si la loi de tout k -uplet est invariante par translation dans le temps :

$$\forall k, \forall (t_1, t_2, \dots, t_k), \quad \mathcal{L}(X_{t_1}, X_{t_2}, \dots, X_{t_k}) = \mathcal{L}(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h}) \quad (2.18)$$

Cette hypothèse de stabilité temporelle de la loi du processus facilite évidemment bien des analyses. Pratiquement, cette hypothèse implique notamment que la covariance entre les valeurs prises en 2 dates distinctes par le processus ne dépend que de la longueur de la période qui les sépare (et pas de la date initiale). $E(X_t)$, $V(X_t)$ et $E(X_t, X_{t-h})$ ne dépendent pas de t , Soit :

$$E(X_t) = \mu, \quad V(X_t) = \sigma^2 \quad \text{et} \quad E(X_t, X_{t-h}) = \gamma_h \quad (2.19)$$

Un processus est dit faiblement stationnaire si ses moments d'ordre 1 et 2 sont stationnaires :

$$E(X_t) = \mu, \quad V(X_t) = \sigma^2 \quad \text{et} \quad E(X_t, X_{t-h}) = \gamma_h \quad (2.20)$$

- La stationnarité entraîne la faible stationnarité.
- Un processus gaussien faiblement stationnaire est stationnaire.

2.2.6.2. Bruit blanc

Un processus $\{X_t\}$ est un bruit blanc si c'est un échantillon iid tel que :

$$\forall t : \quad E(X_t) = 0 \quad (2.21)$$

$$\forall t \neq s \quad (X_t, X_s) \text{ indépendants} \quad (2.22)$$

2.2.6.3. Fonction d'Auto-corrélation et auto-corrélation partielle

Définition 1 : On étudie la mémoire d'un processus en calculant son auto-corrélation de retard h notée ρ_h qui mesure le lien entre les valeurs du processus à deux dates distantes h .

$$\rho_h = \text{Corr}(X_t, X_{t-h}) = \frac{\text{Cov}(X_t, X_{t-h})}{\sqrt{V(X_t)V(X_{t-h})}} \quad (2.23)$$

La courbe $\rho_h = f(h)$ est appelée autocorrélogramme.

Définition 2 : de même on définit l'auto-corrélation partielle de retard h comme la corrélation entre $(X_t - X_t^*)$ et $(X_{t-h} - X_{t-h}^*)$ où X_t^* désigne la régression de X_t sur les $(h - 1)$ valeurs $\{X_{t-1}, X_{t-2}, \dots, X_{t-h+1}\}$:

$$\tau_h = \text{Corr}(X_t - X_t^*, X_{t-h} - X_{t-h}^*) = \frac{\text{Cov}(X_t - X_t^*, X_{t-h} - X_{t-h}^*)}{\sqrt{V(X_t - X_t^*)V(X_{t-h} - X_{t-h}^*)}} \quad (2.24)$$

avec

$$X_t^* = \sum_{k=1}^{h-1} \alpha_k X_{t-k}, \quad X_{t-h}^* = \sum_{k=1}^{h-1} \beta_k X_{t-k}, \quad (2.25)$$

Où les α_k et les β_k sont les coefficients des régressions.

2.2.6.4. Opérateurs de retard B et de différenciation Δ :

Pour étudier le processus (et donc les séries chronologiques), on définit des opérateurs retard B et de différenciation Δ .

L'opérateur B décale le processus d'une unité de temps vers le passé :

$$B(X_t) = X_{t-1} \quad (2.26)$$

L'opérateur Δ fait la différence entre le processus et sa version décalée d'une unité de temps. Cet opérateur se construit en utilisant l'opérateur précédent :

$$\Delta(X_t) = X_t - X_{t-1} = X_t - B(X_t) = I - B \quad (2.27)$$

Où I est l'opérateur identité.

2.2.7. Modèles pour processus stochastiques

Dans ce paragraphe, nous entamons l'étude détaillée de la partie aléatoire d'une série chronologique. Les processus Autorégressifs (AR), les processus à moyennes mobiles (MA) ainsi que les processus ARMA, ARIMA et SARIMA sont décrits.

2.2.7.1. Modèle autorégressif : Processus autorégressif AR(p)

Les premiers modèles que nous présentons sont les processus autorégressifs, construits à partir de l'idée que l'observation au temps t s'explique linéairement par les observations précédentes.

De façon générale, un processus AR(p) est un processus qui dépend linéairement des p valeurs antérieurs :

$$X_t = \varphi_1 X_{t-1} \dots + \varphi_p X_{t-p} + E_t \quad (2.28)$$

Où $\{E_t\}$ est un bruit blanc. On peut sans restriction de généralité, supprimer le terme constant et obtenir un modèle de la forme :

$$X_t = \sum_1^p \varphi_k X_{k-1} + \varphi_{t-p} X_p + E_t \Leftrightarrow \Theta(B) X_t = E_t \quad (2.29)$$

Où Θ est le polynôme de degré p dont les coefficients sont $(1, -\varphi_1, \dots, -\varphi_p)$

2.2.7.2. Modèle moyenne mobile : Processus moyenne mobile MA(q)

Le terme E_t est souvent présenté comme un "choc" rend compte d'un élément nouveau dans l'évolution du processus. On peut envisager que ces chocs (non-corrélés et d'espérances nulles) aient des effets sur l'évolution du processus non seulement à la date à laquelle ils se produisent mais aussi à des dates ultérieures. Il s'agit d'un processus vérifiant l'équation :

$$X_t = E_t + \theta_1 E_{t-1} + \theta_q E_{t-q} = \Theta(B) E_t + E_t \quad (2.30)$$

Où Θ est le polynôme de degré p dont les coefficients sont $(1, \theta_1, \dots, \theta_q)$. Dans un tel modèle, on suppose de plus que l'influence des chocs passés se manifeste au travers d'une fonction linéaire.

Un tel modèle est appelé moyenne mobile d'ordre q car X_t est une moyenne mobile (en anglais Moving Average) appliquée aux variables aléatoires $E_t, E_{t-1}, \dots, E_{t-q}$. Le terme moyenne est à prendre dans un sens très large dans la mesure où la somme des coefficients θ_k n'est pas nécessairement égale à 1.

2.2.7.3. Modèle autorégressif – Moyenne mobile : Processus ARMA(p,q)

On peut bien évidemment envisager de combiner les deux modèles précédents en introduisant

- Une dépendance du processus vis-à-vis de son passé : modèle AR(p),
- Un effet retardé des chocs : modèle MA(q).

Un tel modèle, appelé autorégressif - moyenne mobile (ARMA), est caractérisé par le paramètre p de la partie autorégressive et le paramètre q de la partie moyenne mobile.

Un processus ARMA(p, q) vérifie l'équation

$$X_t = \varphi_1 X_{t-1} \dots + \varphi_p X_{t-p} + E_t + \theta_1 E_{t-1} + \theta_q E_{t-q} \quad (2.31)$$

Soit

$$\Phi(B)X_t = \Theta(B)E_t \quad (2.32)$$

2.2.7.4. Processus ARIMA et SARIMA

La condition de stationnarité des modèles présentés ici n'est évidemment pas toujours convenable. On peut intégrer certains types de non-stationnarités en élargissant le modèle ARMA. Ces élargissements consistent en une série d'opérations préalables visant à éliminer

la tendance ou la saisonnalité. On a vu que si un processus X admet une tendance polynomiale de degré d , le processus différencié d fois est stationnaire :

$$Y_t = \Delta^d X_t = (I - B)^d X_t \quad (2.33)$$

Le modèle ARIMA revient à appliquer un modèle ARMA sur le processus différencié :

$$Y_t = ARMA(p, q) \Leftrightarrow X_t = ARIMA(p, d, q) \quad (2.34)$$

L'équation d'un modèle $ARIMA(p, d, q)$ est donc donnée par

$$\Phi(B)Y_t = \Theta(B)E_t \Leftrightarrow \Phi(B)\Delta^d X_t = \Theta(B)E_t \quad (2.35)$$

Où Φ et Θ sont deux polynôme de degrés respectifs p et q .

Pour le modèle SARIMA, La saisonnalité est un autre facteur de non-stationnarité. On a vu qu'une façon simple d'éliminer une saisonnalité de période ω consiste à appliquer l'opérateur Δ_ω

$$Z_t = \Delta_\omega X_t = (I - B^\omega)X_t = X_t - X_{t-\omega} \quad (2.36)$$

De façon générale, on peut supposer que l'influence des chocs se transmet entre dates distantes d'un nombre entier de périodes selon un processus $ARIMA(P, D, Q)$:

$$\phi_\omega(B^\omega)\Delta_\omega^D X_t = \Theta_\omega(B^\omega)U_t \quad (2.37)$$

Et que ces chocs eux-mêmes suivent un modèle $ARIMA(P, D, Q)$:

$$\phi(B)\Delta^d U_t = \Theta(B)E_t \quad (2.38)$$

Où $\{E_t\}$ est un bruit blanc. Un tel modèle est noté $SARIMA(p, d, q) \times (P, D, Q)_\omega$ et son équation générale

$$\phi(B)\phi_\omega(B^\omega)\Delta^d \Delta_\omega^D X_t = \Theta(B)\Theta_\omega(B)E_t \quad (2.39)$$

Où ϕ , ϕ_ω , Θ et Θ_ω sont des polynômes de degrés respectifs p , P, q, Q .

2.2.8. Identification et estimation des paramètres

2.2.8.1. Identification du modèle : Méthode Box & Jenkins

Le choix entre les différents modèles présentés ici (AR(p), MA(q), ARMA(p,q), ARIMA(p,d,q), SARIMA, etc.) ne peut généralement se faire à priori. On est le plus souvent réduit à des tâtonnements par un système d'essais/erreurs.

Une méthodologie générale a été proposée par Box & Jenkins qui implique la remise en cause de chaque modèle envisagé jusqu'à un modèle acceptable.

Un modèle est acceptable lorsqu'il prend en compte toute la structure de la partie aléatoire et ne laisse qu'un bruit blanc.

2.2.8.2. Détermination du type et de l'ordre du modèle

On s'intéresse ici à l'identification d'un processus stationnaire : on se limite aux processus ARMA(p,q).

On suppose donc que la saisonnalité et la tendance ont été supprimées. Comme on l'a vu précédemment, cette élimination peut s'opérer au moyen de différentiations de différents types :

$$\Delta^d = (I - B)^d \text{ pour éliminer une tendance polynomiale de degré } d,$$

$$\Delta_\omega = (I - B^\omega) \text{ pour éliminer une saisonnalité de période } \omega.$$

Une fois cette manipulation effectuée, on tente de reconnaître le type de processus auquel on a affaire en étudiant les corrélogrammes. Leurs formes théoriques étant connues, on choisit alors un modèle autorégressif, moyenne mobile ou ARMA. La lecture de corrélogrammes se fait en recherchant les chutes

- Dans le corrélogramme partiel pour un AR(p)
- Dans le corrélogramme pour un MA(q)

2.2.8.3. Estimation des paramètres et prévision

La **Méthode des moments** est une des méthodes les plus utilisées pour un modèle donné. Les auto-corrélations ρ_h et τ_h dépendent des paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$.

$$(\rho_1, \rho_2, \dots, \tau_1, \tau_2, \dots) = F(\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d) \quad (2.40)$$

Les auto-corrélations sont faciles à estimer, L'estimation des paramètres est comme suit :

$$(\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d) = F^{-1}(\rho_1, \rho_2, \dots, \tau_1, \tau_2, \dots) \quad (2.41)$$

Le **Maximum de vraisemblance** est une autre méthode classique d'estimation, elle nécessite de faire des hypothèses sur la loi du bruit blanc $\{E_t\}$. On obtient alors des estimateurs du maximum de vraisemblance des paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$ et de σ .

Une fois les paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$ estimés, il est possible de donner une prédiction de la série à la date t en utilisant l'équation du modèle

$$\phi(B)X_t = \theta(B)E_t \quad (2.42)$$

$$\text{Soit } X_t = \{\varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \theta_1 E_{t-1} + \theta_q E_{t-q}\} + E_t$$

Les X_{t-k} ($1 \leq k \leq p$) sont connus à la date $t - 1$ et la partie restante (E_t) est nulle en espérance.

On prend donc comme prévision de X_t la quantité

$$\hat{X} = \widehat{\varphi_1} X_{t-1} + \dots + \widehat{\varphi_p} X_{t-p} + \widehat{\theta_1} \widehat{E_{t-1}} \dots + \widehat{\theta_q} \widehat{E_{t-q}} \quad (2.43)$$

Où \widehat{E}_t sont définis par

$$\widehat{E}_t = X_t - \widehat{X}_t \quad (2.44)$$

Les \widehat{E}_t sont calculables à la date t .

2.2.8.4. Validation

Dans l'analyse des séries chronologiques par processus, le bruit blanc joue un rôle particulier puisque c'est un processus sans aucune structure. Le test préliminaire concerne ce processus. La première hypothèse testée est :

$$H_0 : "\{X_t\} \text{ est un bruit blanc}" \quad (2.45)$$

Si on accepte cette hypothèse, l'analyse de la série est virtuellement achevée : la série étudiée n'a aucune structure. Les tests du bruit blanc permettant également de « valider » le modèle. Si le processus est bien un ARMA(p,q), le processus résiduel $\widehat{E}_t = X_t - \widehat{X}_t$ doit être un bruit blanc.

Les tests les plus connus sont : le test portemanteau (test « fourretout » en anglais) et le test de (Ljung, 1987).

2.3. Concepts de bases de la charge électrique

2.3.1. Introduction

Dans le paragraphe précédent, nous avons étudié la notion de série chronologique. Dans ce paragraphe, nous nous intéressons aux concepts de base de la charge électrique. Nous commençons par rappeler quelques définitions et un survol rapide des concepts nécessaires à la compréhension du reste de la thèse.

2.3.2. Notion de charge électrique

L'électricité possède un nombre de caractéristiques : elle ne peut être stockée que de manière très limitée (dans des batteries), elle se déplace pratiquement à la vitesse de la lumière.

La bonne gestion des flux d'énergie est donc une tâche complexe. L'offre et la demande d'électricité doivent en effet être toujours en équilibre. Ce n'est pas si simple. La consommation subit de grandes fluctuations (le jour par rapport à la nuit, l'été par rapport à l'hiver, le week-end par rapport aux jours de semaines) et n'est pas toujours prévisible.

La charge électrique varie d'une manière continue des qu'on allume ou on éteint l'électricité. Cet échantillon de variation de la charge est appelée la courbe de la charge. Nous pouvons avoir une courbe de charge journalière, mensuelle et annuelle. Le diagramme (voir figure 2.1) est un exemple de courbe de charge électrique pour deux jours : Le 13/12/2005 et le 6/12/2005.

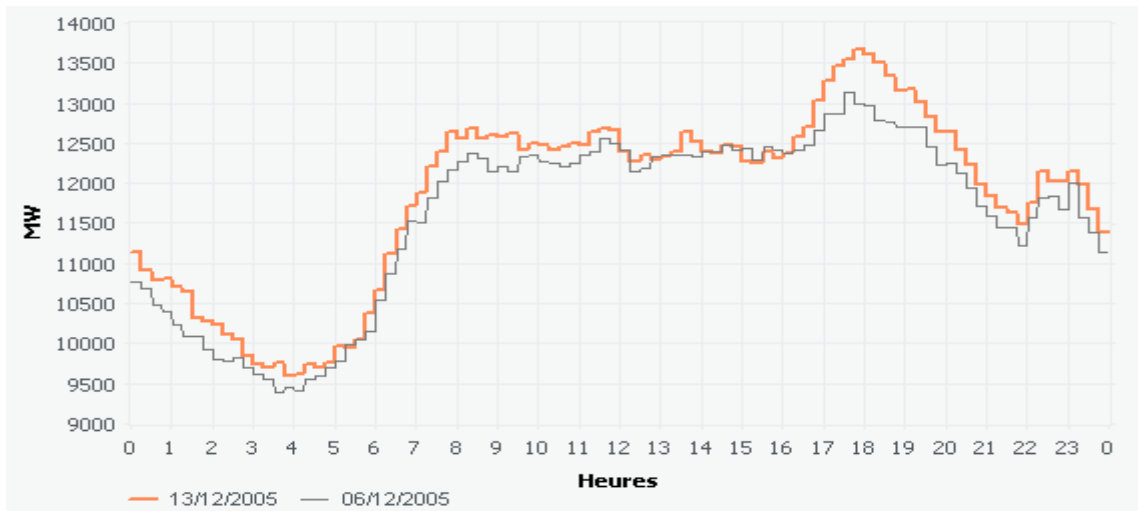


Figure 2.1 - Exemple de charge électrique

Ce graphique représente une courbe de charge en hiver. Cette courbe est le reflet, au quotidien, de nos modes de vie. Que peut-on en déduire ? Il y a peu de consommation la nuit et par conséquent peu de production et de transport. On se réveille dans cet exemple aux alentours de 5 heures : les lumières s'allument un peu partout, le chauffage ainsi que toutes sortes d'appareils électriques (radio, téléviseur, ...) sont mis en service. L'activité dans les bureaux se met peu à peu en route et les usines intensifient leur production. Lorsque la nuit tombe, la charge du réseau diminue à nouveau sensiblement. Notons que la charge d'une journée en été peut être différente.

Contrairement à d'autres formes d'énergie, l'énergie électrique ne peut généralement pas être stockée telle quelle à grande échelle. Chaque Kilowattheure devant donc être produit au moment de sa consommation. Ce délicat équilibre entre l'offre et la demande doit être maintenu en tout temps pour assurer la fiabilité du service électrique.

Par ailleurs, la demande globale d'électricité suit des cycles journaliers, hebdomadaires et saisonniers, basés notamment sur une combinaison de facteurs, dont la succession du jour et de la nuit, le climat, l'activité économique et les habitudes de vie quotidienne des ménages. On distingue trois principaux types de périodes de pointe.

Une **pointe de consommation électrique** est la consommation la plus élevée d'un réseau électrique. Elle dépend de la localisation et de la période étudiée et pose des problèmes particuliers aux gestionnaires de réseaux et aux producteurs d'électricité.

Il y a d'abord les pointes journalières. Elles se produisent souvent en fin de journée un jour de semaine, lorsque les personnes rentrent du travail. La pointe sera plus accentuée dans les réseaux où le chauffage de l'eau et les appareils électroménagers utilisent l'électricité plutôt que le gaz.

Les pointes saisonnières peuvent survenir en été, comme c'est le cas dans la plupart des régions de l'Amérique du Nord ou au Moyen-Orient, par exemple. Dans d'autres régions, comme en France, au Québec et au Nouveau-Brunswick, elle survient invariablement en hiver. Dans les deux cas, les températures extrêmes influent sur la demande de climatisation ou de chauffage électrique des ménages. On parlera de la pointe estivale ou de la pointe hivernale, selon le cas.

2.3.3. La consommation et ses variations

La consommation d'électricité varie en fonction de plusieurs paramètres, dont les deux principaux sont :

- **La période de l'année**, avec trois cycles : un cycle annuel, un cycle hebdomadaire (jours ouvrables et week-end) et un cycle journalier.
- **La température**, dont les variations se traduisent par des modifications de l'utilisation du chauffage électrique en hiver ou de la climatisation en été. La nébulosité (taux de couverture nuageuse) a également une influence sur l'utilisation de l'éclairage et sur le chauffage.

D'autres facteurs ont des effets directement observables sur la consommation. Citons parmi ces facteurs, l'activité économique (creux les week-ends ou lors des vacances), les offres commerciales d'effacement qui incitent les clients à réduire leur puissance électrique consommée, ou le changement d'horaire été-hiver. Des événements exceptionnels peuvent également perturber le profil de la consommation (intempéries, grands événements sportifs), mais leur impact est impossible à prévoir.

2.3.4. Méthode de prévision pour le lendemain

La prévision de la consommation pour le lendemain est d'abord construite à partir d'un **historique des consommations journalières**. Les références les plus courantes, pour un jour ouvrable, sont la courbe de consommation de la veille, du jour correspondant la semaine précédente, du même jour de l'année passée. Anticiper la consommation d'électricité contribue à garantir en permanence l'équilibre production-consommation, et influe directement sur la sûreté du système électrique. Les prévisions de consommation sont réalisées à plusieurs échelles de temps : à long terme, dans le cadre du Bilan Prévisionnel ou des études de développement du réseau, à moyen terme (prévisions annuelles, mensuelles et hebdomadaires) et de la veille pour le lendemain.

2.4. Analyse des données de la charge électrique Algérienne

2.4.1. Données de la charge électrique Algérienne

Les données de la charge électriques, de l'année 2000 à 2004 concernent quatre régions du pays (Oran, Alger, Hassi Messaoud et Ain Amenas). Les régions d'Oran et d'Alger sont situées respectivement au nord ouest et nord centre du pays. Le climat est de type méditerranéen, la population et le développement économique croissent d'une manière très significative. Les deux dernières régions se situent au sud du pays (Sahara) ou baigne un climat chaud pendant toute la période de l'année. La densité de la population est largement inférieure par rapport au nord du pays.



Figure 2.2- Carte géographique de l'Algérie avec les différentes régions (Oran, Alger, Hassi Messaoud et Ain Amenas) en (a) et Carte géographique de l'Algérie en Afrique du Nord (b).

La figure 2.2a illustre la position géographique de chaque région par rapport l'Algérie. La figure 2.2b illustre la position géographique des régions par rapport aux pays limitrophes de l'Afrique du Nord.

La région d'Alger est située au nord de l'Algérie, Elle compte une des plus grandes villes du pays, Alger capitale du pays et situé au bord de la méditerranée. Sa population estimée à 2 882 897 est la plus peuplée d'Algérie. Sa superficie est de 1 190 km² et sa densité est de 2422 habitants/km². Cette région bénéficie d'un climat méditerranéen. Elle est connue par ses longs étés chauds et secs. Les hivers sont doux et humides, la neige est rare mais pas impossible. Les pluies sont abondantes et peuvent être diluviennes. Il fait généralement chaud surtout de la mi-juillet à la mi-août.

La région d'Oran est une subdivision administrative algérienne située au nord-ouest du pays. Elle a pour chef lieu la ville d'Oran située au sud de la méditerranée. Cette région est peuplée de 1 584 607 habitants pour une superficie de 2 114 km² soit une densité de 749,58 habitants/km².

Oran bénéficie d'un climat méditerranéen classique marqué par une sécheresse estivale, des hivers doux, un ciel lumineux et dégagé. Pendant les mois d'été, les précipitations deviennent rares voir inexistantes et le ciel est lumineux et dégagé. L'anticyclone subtropical recouvre la région oranaise pendant près de quatre mois. En revanche la région est bien arrosée pendant l'hiver. Les faibles précipitations (420 mm de pluie) et leur fréquence (72,9 jours par an) sont aussi caractéristiques de ce climat.

La région de Hassi Messaoud est le cœur économique de l'Algérie grâce au pétrole. Sa superficie est de 211 980 km². Sa population est estimée à 633 967 habitants et sa densité est de 2,99 habitants/km². Il fait chaud pendant toute la période de l'année.

Cette région est située à l'extrême Sud Est du pays. Elle est limitrophe de trois pays, la Tunisie au nord-est, la Libye à l'est et le Niger au sud. Avec 284.618 km², c'est la troisième plus grande superficie du pays. Sa population est 42 609 et sa densité est 0.15 habitants/km². Le Climat est désertique avec une pluviométrie de 10 mm/an.

Pour chaque région étudiée, les données brutes associées sont regroupés dans un seul fichier. Les valeurs forment une série chronologique. Nous avons 175392 données (24x4x1827) pour chaque région. (Voir Tableau 2.1).

Période	1 janvier 2000 – 31 décembre 2004
Intervalle de temps	Quart d'heure
Nombre de valeurs	175392

Tableau 2.1 - Les données de la charge électrique (Période, intervalle de temps et nombre de valeurs).

Ces données sont disponibles en format Excel contenant pour chaque quart d'heure de la journée, les valeurs de la demande électrique en Watt que consomment les quatre régions : Oran, Alger, Hassi Messaoud et Ain Amenas. Les données journalières sont recueillies puis transformées dans un seul fichier Matlab.

La demande électrique de la charge électrique algérienne du 01/01/2000 au 31/12/2004 est schématisée en figure 2.3. On remarque une tendance croissante reflétant une activité économique durant cette période.

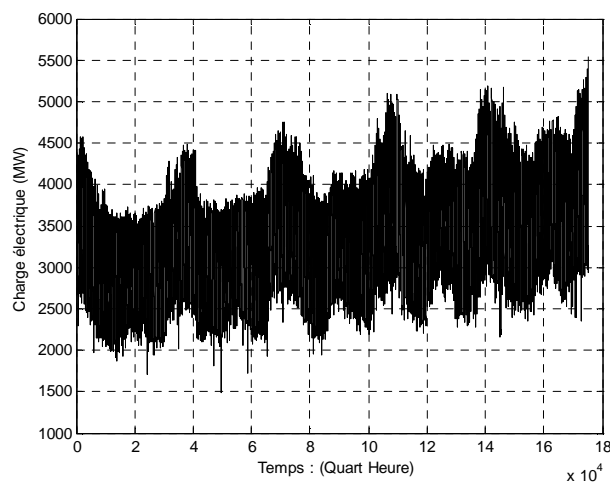


Figure 2.3 - Charge électrique Algérienne pour la période 2000-2004.

Cet accroissement est mis en évidence comme suit. Nous utilisons une équation de forme quadratique :

$$d(t) = a t^2 + b t + c + \varepsilon(t) \quad (2.46)$$

Où t est le temps en heure (début en 2000), $d(t)$ est la tendance au temps t , $\varepsilon(t)$ est une erreur, a , b et c sont les coefficients positifs calculées (Voir tableau 2.2).

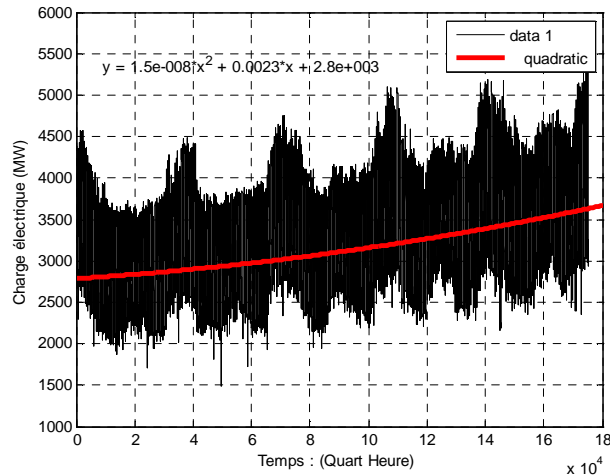


Figure 2.4 Charge avec une courbe de tendance approximative

Coefficients	Valeurs
a	1.5×10^{-8}
b	0.0023
c	2.8×10^3

Tableau 2.2 Coefficients de la courbe de la tendance.

2.4.2. Prétraitement des données de la charge électrique nationale Algérienne

Il est nécessaire alors de réaliser un prétraitement de ces données afin d'obtenir les meilleurs résultats. Le prétraitement consiste donc à :

- Vérifier l'origine des données.
- Traiter les valeurs aberrantes.
- Traiter les valeurs manquantes, les valeurs nulles.

Le nettoyage des données consiste à tenir compte des données manquantes et aberrantes. Ceci se produit lors du fonctionnement incorrect des appareils de mesures. La transformation des données peut se réaliser soit par une normalisation Min Max (2.47) ou une standardisation de type Z-Score (2.48).

$$new = \frac{vact - min}{max - min} \quad (2.47)$$

où new et $vact$ désignent respectivement les valeurs nouvelles et actuelles, min et max sont respectivement la plus petite et la plus grande valeurs.

$$new = \frac{vact - \mu}{\sigma} \quad (2.48)$$

où μ et σ sont respectivement la moyenne et l'écart type.

La transformation des données consiste à faire une normalisation dans l'intervalle 0 et 1. On modifie alors les valeurs de départ par rapport à la valeur minimale et maximale. On aura ainsi des séries de données homogènes qui varient entre 0 et 1.

La préparation des données peut inclure également la réduction des dimensions des données en utilisant une analyse par composante principale qui est une méthode de la famille de l'analyse des données et plus généralement de la statistique multivariée, qui consiste à transformer des variables liées entre elles en de nouvelles variables indépendantes les unes des autres. Ces nouvelles variables sont nommées composantes principales ou axes. Elle permet au praticien de réduire l'information en un nombre de composantes plus limité que le nombre initial de variables (Tipping et Bishop, 1999).

2.4.3. Caractéristiques de la charge électrique Algérienne

Dans la terminologie utilisée en énergie électrique, le côté demande et souvent désigné par le terme charge électrique. Lorsqu'on allume un téléviseur ou un ordinateur nous avons dans l'esprit que l'électricité est présent. A chaque sollicitation de l'électricité, La charge change d'une manière continue formant une courbe de charge. On a donc la courbe de charge journalière, hebdomadaire, saisonnière et annuelle.

La figure 2.5 donne un exemple de courbe de charge journalière de travail en été et en hiver. Le profil de la consommation de la charge est caractérisé par 4 points :

- Le creux de nuit, qui correspond au minimum de la consommation sur les 24 heures de la journée,
- La pointe du matin,
- Le creux de l'après midi
- La pointe du soir.

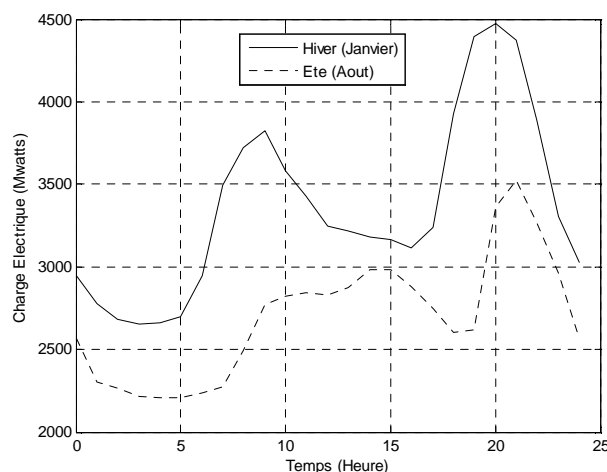


Figure 2.5 - Charge électrique d'une journée de travail en été et en hiver.

On peut noter également une différence nette entre la courbe de la charge d'un jour d'hiver et d'un jour d'été. Les pics de consommations sont assez élevés en hiver par rapport à l'été. Un

jour de travail en hiver (Samedi) exhibe un pic assez élevé à 20 heures, pour un jour d'été le pic est atteint à 21 heures (Figure 2.5)

La figure 2.6 est un exemple de charge hebdomadaire. Notons, une différence nette entre la courbe de la charge électrique d'un jour de travail (du Samedi au Mercredi) et un jour de weekend (Jeudi et vendredi). Ceci s'explique par la diminution de l'activité économique et la prière du vendredi. Notons qu'en Algérie les Week-ends sont le jeudi et le vendredi. La charge électrique peut être ventilée en groupes, appelés type de jours. Chaque groupe ayant des caractéristiques communes.

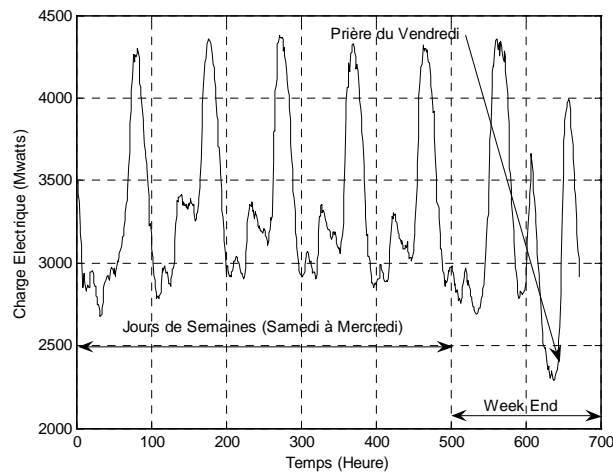


Figure 2.6 - Charge électrique hebdomadaire.

Le profil de consommation sur une journée (samedi 18 mars 2000) est illustré à la figure 2.7 pour les quatre régions. Notons les différences entre les pics de matinée et de soirée (courbe en vert pour Oran et bleu pour Alger) Ces pics sont assez élevés pour les villes du nord par rapport à celles du sud (courbe en rouge pour Hassi Messaoud et bleu ciel pour Ain Amenas).

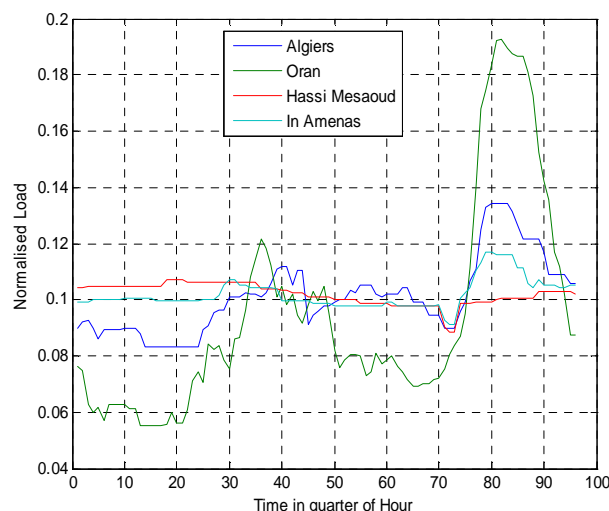


Figure 2.7 - Charge électrique par région pour un jour particulier (Samedi 18 Mars 2000).

Sur les courbes des villes du nord (vert pour Oran et bleu pour Alger), notons la présence habituelle des trois pics de la journée contrairement à celle du sud. Ceci est dû aux habitudes

des consommateurs et à l'activité économique en constante croissance qui distingue les régions du nord par rapport à celles du sud du pays.

2.5. Conclusion et contributions

Dans le cadre de ce chapitre, nous avons tout d'abord mis en évidence la notion de charge électrique comme une série temporelle. Nous avons insisté sur la phase préparation des données qui est primordiale avant tout type de traitement.

Nous avons ensuite présenté la description, le prétraitement et les caractéristiques de la charge électrique algérienne. Cette étude a mis en relief le caractère cyclique de la consommation. On peut identifier un cycle annuel, un cycle hebdomadaire et un cycle journalier.

Des différences existent entre les consommations identiques et stables des jours de semaines et de Week-end (Jeudi et Vendredi où la consommation diminue) ainsi qu'un jour en hiver et en été. Ceci nous amène à vouloir identifier les profils de consommation appelés type de jours.

D'autres paramètres peuvent influencer la consommation d'électricité : La météorologie (la température et la nébulosité), les fêtes religieuses et nationales, le ramadhan et les événements exceptionnels (finale de la coupe du monde, éclipse totale du soleil). Toutefois, nous nous tiendrons aux paramètres cycliques qui peuvent être eux même une série temporelle, et ceci dans la limite des données disponibles et des contraintes d'utilisation.

Dans le prochain chapitre, nous allons analyser l'état de l'art scientifique relatif à l'identification de types de jours. Nous identifierons, alors, les principales approches fondées sur la classification automatique – *clustering*–.

3

Classification non-supervisée : Etat de l'art.

Sommaire

3.1 Introduction	26
3.2 Concepts de base de la classification automatique	26
3.2.1 Les données	26
3.2.2 Les groupes d'objets	26
3.2.3 Partitionnement d'un groupe d'objets	29
3.2.4 Partitionnement dur ou stricte	29
3.2.5 Partitionnement flou	29
3.3 Algorithmes classiques de classification	30
3.3.1 Algorithme de classification hiérarchique	30
3.3.2 Les Nuées dynamiques	32
3.3.2.1 L'algorithme des K-moyennes	32
3.3.2.2 L'algorithme des C-moyennes floues	32
3.3.3 Approche à bases de réseaux de neurones : les cartes auto-organisées de Kohonen	36
3.4 Approche de classification à deux niveaux	36
3.5 Evaluation et critère de validité	37
3.5.1 Indices propres aux cartes-organisées de Kohonen	37
3.5.1.1 Erreur de quantification	37
3.5.1.2 Taux d'erreurs topologiques	38
3.5.2 Indices propres à la classification dure	38
3.5.2.1 Erreur Quadratique Moyenne	38
3.5.2.2 Indice de Dunn	39
3.5.2.3 Indice de Davies-Bouldin	39
3.5.3 Indices propres à la classification floue	39
3.5.3.1 Coefficient de Partition	39
3.5.3.2 Entropie de classification	40
3.5.3.3 Indice de Partition	40
3.5.3.4 Indice de Xie et Beni	40
3.6 Résumé et Conclusion	

3.1. Introduction

La classification non supervisée est une méthode de segmentation des données de la charge électrique Algérienne. Le but de ce chapitre est d'introduire les concepts fondamentaux pour une bonne compréhension des chapitres suivants. Nous commencerons par un rappel des définitions et concepts utilisés dans l'analyse par regroupement. Ensuite nous donnons un survol de l'ensemble des approches utilisés en classification en général pour la classification non supervisée et plus particulièrement les approches utilisées dans ce travail. Nous terminerons ce chapitre par l'investigation de la validité des classes obtenues avec l'introduction des indices de validité.

3.2. Concepts de base de la classification automatique – *Clustering* -

La classification non supervisée ou classification automatique (*clustering*) est une étape importante de l'analyse de données, elle consiste à identifier des groupes d'objets ou individus similaires (clusters) à partir d'un ensemble de données sans en connaître au préalable sa structure. Elle ne doit pas être confondue avec la classification supervisée ou classement (classification) qui consiste à déterminer les règles qui ont permis de séparer un ensemble d'individus en classes connues a priori (Jain et Dubes, 1988 ; Everitt et al., 2001 ; Xu et Wunsch, 2009).

3.2.1. Les données – *Data* -

Les techniques de partitionnement peuvent être appliquées aux données quantitatives (numériques) ou qualitatives (par catégories). Dans cette thèse, il est question de données quantitatives. Les données sont des observations obtenues par expérience. Un ensemble de n échantillons est noté $X = \{X_1, X_2, \dots, X_n\}$ où Chaque élément de $X_i = (x_{11}, x_{12}, \dots, x_{1d})^T$ est un vecteur représentant le i ème objet et x_{ii} est un attribut de X_i , d étant le nombre d'attributs de l'objet X_i .

3.2.2. Les groupes d'objets – *Clusters* -

Il existe une multitude de définition pour définir la notion de groupes d'objets. Généralement, un groupe d'objets est un ensemble d'objets similaire ou identiques au sein du même groupe mais différents par rapport aux objets des autres groupes.

Les données peuvent constituer un groupe d'objets de différentes formes géométriques (Voir figure 3.1) sphériques, linéaire, etc. Un groupe d'objet peut être bien séparé ou enchevêtré les uns avec les autres (Figure 3.2).

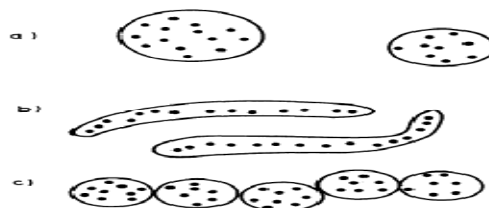


Figure 3.1 Quelques formes de classes : (a) En forme de boules. (b) Formes étirées. (c) Compacte mais pas bien séparées

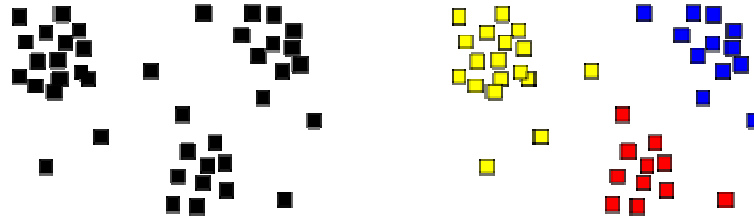


Figure 3.2 - Exemple de classification automatique (trois classes détectées jaune, bleu et rouge).

Il est commun de définir le concept de similarité à l'aide de la notion duale de dissimilarité, on dit de deux individus qu'ils sont d'autant plus similaires qu'ils sont proches au sens d'une mesure de dissimilarité.

Nous rappelons ci-dessous la définition générale d'une mesure de dissimilarité avant de considérer le cas des métriques et des ultramétriques qui sont deux types de mesures particulières.

Définition 3.2.1. (Mesure de dissimilarité) On appelle indice ou mesure de dissimilarité sur un ensemble, une application $d : \Omega \times \Omega \rightarrow R_+$ qui vérifie les propriétés suivantes pour tout couple

$(x, y) \in \Omega \times \Omega$:

$d(x, y) = d(y, x)$ (symétrie)

$d(x, y) = 0 \iff x = y$ (séparabilité)

Définition 3.2.2. (Métrique) On appelle métrique sur un ensemble, une application $d : \Omega \times \Omega \rightarrow R_+$

qui vérifie les propriétés suivantes pour tout couple $(x, y) \in \Omega \times \Omega$

$d(x, y) = d(y, x)$ (symétrie)

$d(x, y) = 0 \iff x = y$ (séparabilité)

$d(x, y) \leq d(x, z) + d(z, y)$ (inégalité triangulaire)

Définition 3.2.3. (Ultramétrique) On appelle ultramétrique sur un ensemble, une application

$d : \Omega \times \Omega \rightarrow R_+$ qui vérifie les propriétés suivantes pour tout couple $(x, y) \in \Omega \times \Omega$

$d(x, y) = d(y, x)$ (symétrie)

$d(x, y) = 0 \iff x = y$ (séparabilité)

$d(x, y) \leq \max\{d(x, z), d(z, y)\}$ (inégalité ultramétrique)

L'homogénéité des individus regroupés au sein d'un groupe est souvent évaluée à l'aide d'un critère statistique appelée **variance** dont la définition est rappelée ci-dessous.

Définition 3.2.4. (Variance) On définit la variance $V(C_i)$ d'un groupe d'objets C_i ainsi :

$$V(C_i) = \frac{1}{N_i} \sum_{x_j \in C_i} d^2(x_j - \mu_i) \quad (3.1)$$

où N_i et μ_i sont respectivement le nombre d'objets et le centroïde du groupe C_i .

Dans le contexte de la classification automatique, on distingue généralement la variance intra-classe V_{intra} , que l'on souhaite minimiser, de la **variance inter-classe** V_{inter} , que l'on cherche à maximiser :

$$V_{intra} = \frac{1}{N} \sum_{C_i \in C} N_i \times V(C_i)^2 \quad (3.2)$$

$$V_{inter} = \frac{1}{N} \sum_{C_i \in C} N_i \times (\mu_i - \mu)^2 \quad (3.3)$$

où N_i et μ_i sont respectivement le nombre d'objets et le centroïde du groupe C_i , et de manière analogue, N et μ désignent respectivement le nombre d'objets et le centroïde de . La première évalue l'homogénéité moyenne des groupes d'une partition et la seconde permet de quantifier la différence entre les groupes. La **formule de König-Huyghens** (Jain et Dubes, 1988) permet de relier la variance intra-classe et inter-classe à la variance totale $V_{totale} = V(\Omega)$:

$$V_{totale} = V_{intra} + V_{inter} \quad (3.4)$$

Une mesure de la distance ou métrique est utilisée pour évaluer la similitude entre observations d'une même classe. On distingue plusieurs types de métriques, la plus connue étant la distance euclidienne. L'indice de proximité entre l'échantillon d'ordre i et d'ordre k est noté par $d(i, k)$. Soit une matrice d'échantillons $[x_{ij}]$ où x_{ij} est la j ème caractéristique de l'échantillon i tel que $1 \leq i \leq n$ et $1 \leq j \leq m$ avec n le nombre d'échantillons et m le nombre de caractéristiques.. Les différentes mesures de dissimilarité sont décrites par le tableau 3.1.

Distance	Equation
Euclidienne	$\sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} = [(x_i - x_j)^t (x_i - x_j)]$
Euclidienne élevé au carré	$\sum_{k=1}^n (x_{ik} - x_{jk})^2$
Mahalanobis	$\sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^T}$
Manhattan ou City block	$\sum_{k=1}^n x_{ik} - x_{jk} $
Chebychev	$\max_{1 \leq k \leq n} x_{ik} - x_{jk} $
Cosine	$1 - \frac{\sum_{k=1}^n x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^n x_{ik}^2 x_{jk}^2}}$
Minkowski	$\sum_{k=1}^n (x_{ik} - x_{jk} ^t)^{1/t}$

Tableau 3.1 Mesure de dissimilarité pour le calcul de la distance des centres des classes.

3.2.3. Partition d'un groupe d'objets – Partition -

Un groupe d'objet est un sous-ensemble d'un ensemble de données. Le but de la classification est de trouver c groupes d'objets ou classes, noté (X_1, X_2, \dots, X_c) d'un ensemble X de données. On utilise une matrice de partition notée $U = [u_{ij}]$, où $i = 1, \dots, n$ et $j = 1, \dots, c$ pour exprimer le degré d'appartenance d'un objet quelconque x_j de X à la classe X_j .

3.2.4. Partitionnement dur ou strict - *Hard partition* -

Un partitionnement dur ou strict peut être défini comme une famille de sous-ensembles $\{X_i \mid 1 \leq i \leq c\}$ tel que $u_{ij} = 1$ si $x_j \in X_i$ sinon $u_{ij} = 0$. Chaque objet appartient à un et un seul groupe.

Les propriétés suivantes sont vérifiées :

$$X_i \neq \emptyset \text{ for } i = 1, 2, \dots, c \quad (3.5)$$

$$X_i \cap X_j = \emptyset \quad i = 1, 2, \dots, c, j = 1, 2, \dots, c \text{ and } i \neq j \quad (3.6)$$

$$\bigcup_{i=1}^c X_i = X \quad (3.7)$$

L'équation (3.5) signifie qu'il n'existe pas de classes vides, l'équation (3.6) signifie qu'aucun élément n'appartient à plus d'une classe et l'équation (3.7) que l'union de toutes les classes forme l'univers de toutes les classes.

Les éléments de la matrice de partition sont définis comme suit :

$$u_{ij} \in \{0, 1\} \quad 1 \leq i \leq n, \quad 1 \leq j \leq c \quad (3.8)$$

$$\sum_{j=1}^c u_{ij} = 1, \quad 1 \leq i \leq n, \quad (3.9)$$

$$1 \leq \sum_{i=1}^n u_{ij} = 1 \quad 1 \leq j \leq c \quad (3.10)$$

L'équation (3.8) signifie que chaque objet peut appartenir ou pas à une classe. La contrainte (3.9) implique que chaque objet appartient à une et une seule classe. La contrainte (3.10) implique que chaque classe contient au moins un objet.

3.2.5. Partitionnement flou - *Fuzzy partition* -

Une partition floue peut être vue comme une généralisation du partitionnement dur, il permet à u_{ij} d'atteindre des valeurs réelles dans $[0, 1]$. Chaque objet appartient à plusieurs groupes d'objets avec des degrés d'appartenance. La classification stricte vue précédemment peut être vue comme un cas particulier de la classification floue où les échantillons ont une valeur 1 lorsque à la classe à laquelle ils appartiennent et 0 aux autres classes. Souvent dans la classification floue, il y a présence d'entrecroisement entre classes. Les éléments de la matrice de partition sont définis comme suit :

$$u_{ij} \in [0,1], \quad 1 \leq i \leq n, \quad 1 \leq j \leq c \quad (3.11)$$

$$\sum_{j=1}^c u_{ij} = 1, \quad 1 \leq i \leq n, \quad (3.12)$$

$$0 < \sum_{i=1}^n u_{ij} < n \quad 1 \leq j \leq c \quad (3.13)$$

La matrice notée $U = [u_{ij}]$, où $i = 1, \dots, n$ et $j = 1, \dots, c$ contient les valeurs d'appartenance de chaque à une classe (0 ou 1), la matrice a n lignes et c colonnes. La somme de tous les éléments d'une colonne doit être égale à 1. La somme de chaque ligne doit être inférieure à 1 et supérieure à zéro.

3.3. Algorithme classiques de classification

On distingue deux grandes familles de techniques de classification : la classification hiérarchique et la classification non hiérarchique appelé aussi nuées dynamiques (Jain et Dubes, 1988 ; Everitt et al., 2001 ; Xu et Wunsch, 2009).

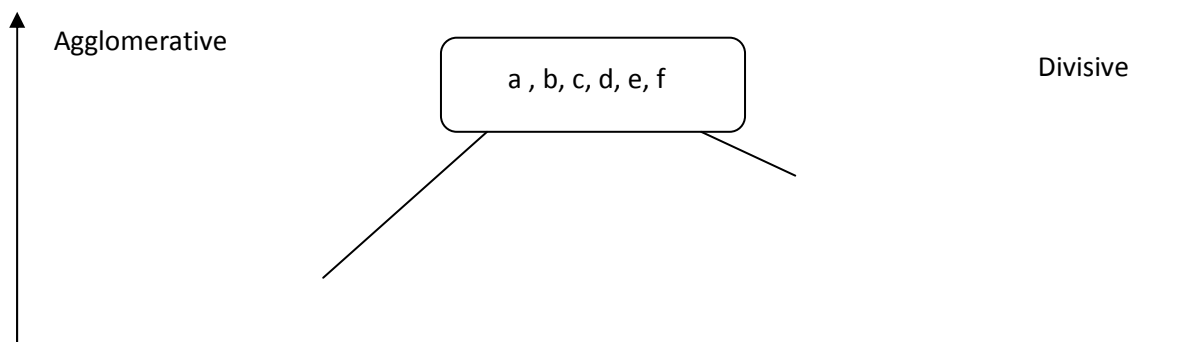
3.3.1. Algorithme de classification hiérarchique

La classification hiérarchique crée une hiérarchie entre classes, cette hiérarchie peut être représentée par une structure arborescente appelée dendrogramme.

On distingue deux types d'approches de classification hiérarchique, les méthodes descendantes – *divisive* – et les méthodes ascendantes – *agglomérative* –.

La méthode descendante procède par divisions successives de l'ensemble des observations jusqu'à obtenir une partition formée de singletons. La méthode ascendante, quand à elle commence par la partition de l'ensemble des observations et procède par fusion progressive des parties jusqu'à obtention de la partition la plus grossière.

La figure 3.3 illustre un exemple de la classification hiérarchique sur l'ensemble $\{a,b,c,d,e,f\}$ à l'aide des deux méthodes ascendantes et descendantes .



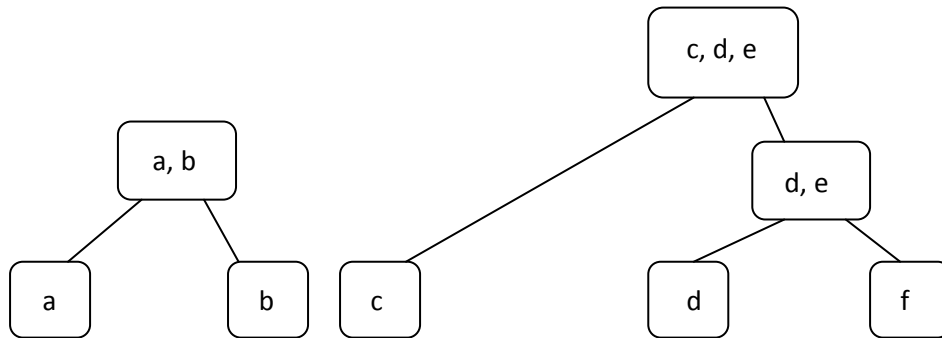


Figure 3.3 - Classification ascendante hiérarchique et descendante hiérarchique.

Les méthodes descendantes sont pratiquement inutilisées. Différents indices d'agrégation de groupes ont été proposés pour la classification ascendante hiérarchique qui est la plus utilisée.

- L'indice du saut minimum – *single linkage* - est défini comme la distance minimale qui sépare deux éléments issus de groupes différents.
- L'indice du saut maximum – *complete linkage* – correspond à la distance maximale qui sépare deux éléments issus de groupes différents.
- L'indice du saut moyen est l'espérance de la distance qui sépare deux éléments issus de groupes différents.
- La distance entre les centroïdes des groupes qui se calcule au plus en temps linéaire $O(N)$ contrairement aux indices précédents dont la complexité est quadratique $\theta(N^2)$.
- L'indice de Ward est défini comme l'augmentation de la variance intra-classe résultant de la fusion des deux groupes considérés.

Nous ne détaillerons pas d'avantage ces méthodes qui sont trop coûteuses pour être utilisées sur les volumes de données manipulés aujourd'hui. Le lecteur intéressé est invité à consulter (Xu et Wunsch, 2009) pour avoir plus de détails.

3.3.2. Les nuées dynamiques

Dans le cas des méthodes de type « nuées dynamiques », chaque groupe est représenté par un prototype, encore appelé centre, et chaque objet est affecté au groupe dont il est le plus proche. Nous commençons par introduire l'algorithme des K-moyennes avant de présenter une extension aux classifications floues (Xu et Wunsch, 2009).

3.3.2.1. L'algorithme des K-moyennes - *K-means* -

L'algorithme des K-moyennes est l'un des algorithmes les plus connus de la classification dure (Forgy, 1965 ; MacQueen, 1967 ; Duda et al., 2001 ; Theodorakis et Koutroumbas, 2006).

Il consiste à choisir aléatoirement des centres initiaux et améliorer la partition obtenue de manière itérative jusqu'à obtention d'une partition optimale en minimisant le critère suivant selon une procédure d'optimisation itérative.

$$J(U, M) = \sum_{j=1}^n \sum_{i=1}^c u_{ij} \|x_j - m_i\|^2 \quad (3.14)$$

où n est le nombre total d'objets de l'ensemble des données $= \{x_1, x_2, \dots, x_n\}$, c est le nombre de classes de l'ensemble des centres des classes $V = \{m_1, m_2, \dots, m_c\}$, u_{ij} est un élément de la matrice de partition, $\|x_i - m_j\|$ désigne la norme euclidienne entre x_i et m_j .

L'algorithme est composé des étapes suivantes :

Etape 1 : Initialiser une K-partition aléatoire. Calculer la matrice des classes prototypes $M = [m_1, \dots, m_K]$;

Etape 2 : Affecter chaque objet de l'ensemble des données au plus proche cluster C_i i.e., $x_j \in C_l$, si $\|x_j - m_l\| < \|x_j - m_i\|$ pour $j = 1, \dots, N, i \neq l$, et $i = 1, \dots, K$;

Etape 3 : Recalculer la matrice du cluster prototype

$$m_i = \frac{1}{N_i} \sum_{x_j \in C_i} x_j \quad (3.15)$$

Etape 4 : Répéter les étapes 2 et 3 jusqu'à pas de changement pour chaque cluster.

L'algorithme des K-moyenne est facile à mettre en œuvre mais nécessite la spécification du nombre de classes et sa convergence dépend des conditions initiales.

3.3.2.2. L'algorithme des C-moyennes floues -Fuzzy C-means : FCM -

(Zadeh, 1965) eut l'idée de définir les fonctions d'appartenance pour toutes les classes, la classification flou a été appliqué dans différents domaines.

Dans la littérature de la classification flou, L'algorithme des C-moyennes floues (Fuzzy C-means (FCM)) et ses variantes, sont les méthodes les plus connues (Hoppner et al., 1999 ; Bezdek, 1981).

Cette approche est semblable au K-moyenne, sauf que la partition engendrée par les centres est floue, c'est-à-dire que le degré d'appartenance u_{ij} varie dans l'intervalle $[0,1]$ au lieu d'être un élément de $\{0,1\}$.

L'algorithme de classification C-moyenne flou consiste à minimiser la fonction objective

$$J(U, M) = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m \|x_i - m_j\|^2 \quad (3.16)$$

où n est le nombre total d'objets de l'ensemble des données $X = \{x_1, x_2, \dots, x_n\}$, c est le nombre de classes de l'ensemble des centres des classes $V = \{m_1, m_2, \dots, m_c\}$, u_{ij} est un élément de la matrice de partition, $\|x_i - m_j\|$ désigne la norme euclidienne entre x_i et m_j et $m > 1$ est un exposant qui fixe le niveau de flou coefficient. Les centres des classes et les fonctions d'appartenances de la fonction objective s'expriment par :

$$m_i = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \quad 1 \leq i \leq c \quad (3.17)$$

$$u_{ij} = \left[\sum_{k=1}^c \left(\frac{\|x_i - m_j\|^2}{\|x_i - m_k\|^2} \right)^{1/(m-1)} \right]^{-1} \quad 1 \leq j \leq c, 1 \leq i \leq n \quad (3.18)$$

L'algorithme FCM s'énonce comme suit :

Étape 1 Soit X un ensemble de donnée, choisir le nombre de classes c tel que $1 < c < n$, l'exposant $m > 1$, l'erreur de tolérance $\varepsilon > 0$. Initialiser la matrice de partition U

Étape 2 Calculer les centres des classes m_i $i = 1, \dots, c$ en utilisant l'équation 3.17

Étape 3 Mettre à jour la matrice de partition selon équation 3.18

Étape 4 Si l'amélioration de $J(U, M)$ est inférieure à un seuil, on arrête sinon on retourne à l'étape 2.

L'algorithme des C-moyennes floues généralise donc celui du K-moyenne en étant beaucoup moins sensible au bruit dans les stimuli grâce aux partages de l'appartenance entre les différentes classes.

3.3.3. Approche à base de réseaux de neurones : Les cartes auto-organisées de Kohonen.

L'algorithme de cartes auto-organisées (Self-Organizing Map (SOM)) a été introduit par Kohonen (Kononen, 1990) au début des années 80 pour modéliser un phénomène couramment observé dans le cerveau : la formation de cartes. L'objectif des cartes de Kohonen est de représenter des données d'entrées multidimensionnelles en prototypes vecteurs qui peuvent être visualisés dans une carte à une ou deux dimensions. Les cartes de Kohonen sont constituées de deux couches. La couche d'entrée où les données sont présentées et la couche topologique composée de neurones selon une géométrie prédéfinie. Chaque unité de la carte de Kohonen est appelé neurone et est totalement connecté aux neurones d'entrées. Le vecteur poids de ces connexions forme le référent associé au neurone.

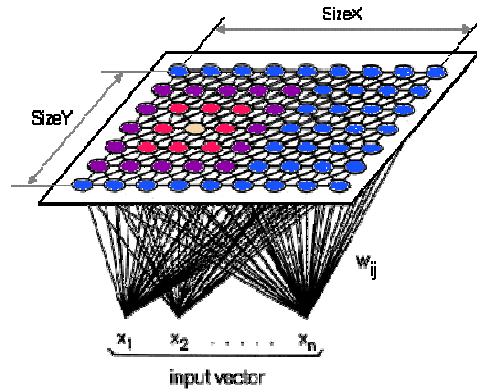


Figure 3.4 - Structure de la carte de Kohonen.

Pendant la phase d'apprentissage, le processus d'auto-organisation permet de concentrer l'adaptation des poids de connexion sur la région de la carte la plus active. Cette région d'activité est choisie comme étant le voisinage associé au neurone dont l'état est le plus actif. Le critère de sélection du neurone le plus actif est de chercher celui dont le vecteur poids est le plus proche au sens de la distance euclidienne de la forme présentée.

Différents algorithmes d'apprentissage ont été proposés pour l'adaptation des poids de la carte, le voisinage V_i d'un neurone i est composé des neurones de la carte topologique qui se situent à l'intérieur de la zone d'influence. C'est le choix de la fonction h qui permet d'introduire des zones d'influence autour de chaque de chaque neurone. La fonction de voisinage h peut être de la forme :

$$h_{ji}(t) = \eta(t) \exp\left(\frac{-\|r_j - r_i\|}{2\sigma^2(t)}\right) \quad (3.19)$$

Où r_j and r_i représentent les coordonnées des neurones correspondants sur la carte et $\sigma(t)$ est une fonction monotone et décroissante. On peut choisir $\sigma(t)$ comme :

$$\sigma(t) = \sigma_0 \exp(-t/\tau) \quad (3.20)$$

Où σ_0 est une valeur initiale et τ est un temps constant (Haykin, 1999).

La fonction de voisinage est une fonction gaussienne introduit pour chaque neurone de la carte un voisinage global. La taille de ce voisinage est limité par $\sigma(t)$ de la gaussienne.

La version de l'algorithme d'apprentissage de ce modèle se déroule essentiellement en trois phases :

La phase d'initialisation où des valeurs aléatoires sont affectées aux poids des connexions (référents ou prototypes) de chaque neurone de la carte. La phase de compétition pendant laquelle, pour toute forme d'entrée x_i , un neurone J de voisinage $V_{b(i)}$ est sélectionné comme gagnant. Ce neurone est dont le vecteur poids est le plus proche au sens de la distance euclidienne :

$$J = \arg_j \min(x_i - w_j) \quad (3.21)$$

La phase d'adaptation où les poids de chaque neurone de la carte sont mis à jour selon les règles d'adaptation suivantes :

$$w_j(t+1) = w_j(t) + h_{jj}(t)(x - w_j(t)) \quad (3.22)$$

Ce processus d'adaptation est répété jusqu'à stabilisation de l'auto-organisation. Une version batch de cet algorithme a été proposée : les vecteurs poids ne sont mis à jour qu'après la présentation de toutes les formes d'entrées et on remplace alors le prototype des neurones par le barycentre pondéré à l'aide de la fonction de voisinage des formes d'entrées qui les ont activés.

L'algorithme des cartes-auto organisés s'énonce comme suit :

Étape 1 : Déterminer la topologie de la carte auto-organisée. Initialiser aléatoirement les vecteurs poids $w_j(0)$ pour $j = 1, \dots, K$;

Étape 2 : Présenter un échantillon d'entrée x au réseau. Choisir le neurone gagnant J ayant la plus petite distance euclidienne à x , i.e.

$$J = \arg \min(x - w_j) \quad (3.23)$$

Étape 3 : Calculer le taux d'apprentissage courant et la taille du voisinage

Étape 4 : Mettre à jour les vecteurs poids des neurones dans le voisinage de selon l'équation eq. 3.22.

Étape 5 : Répéter les étapes 2 à 4 le changement de la position du neurone soit proche d'un nombre positif choisi.

3.4. Approche de classification à deux niveaux

A l'inverse des méthodes directes qui opèrent directement sur les données, l'approche à deux niveaux combine les cartes auto-organisatrices de Kohonen avec un algorithme de classification partitionné (K-moyenne ou C- moyenne floue). Cette approche est décrite sur la Figure 3.5.

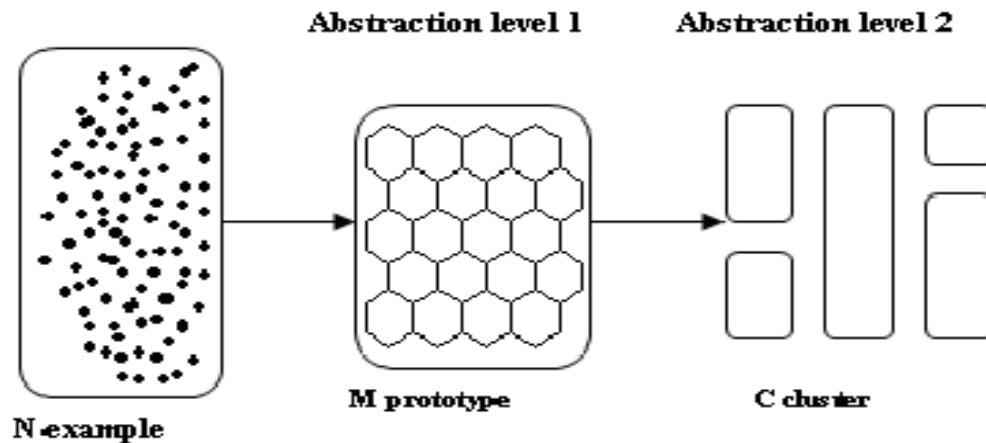


Figure 3.5 -Une première abstraction est obtenue en créant un ensemble de prototypes en utilisant les cartes auto-organisées. Le second niveau d'abstraction classe ensuite les prototypes obtenus.

Une première abstraction est obtenue en créant un ensemble prototype en utilisant le plus souvent les cartes auto-organisées de Kohonen. Le second niveau classe ensuite les prototypes obtenus.

(Vesanto et Alhoniemi, 2000) utilisèrent une approche à deux niveaux combinant les cartes auto-organisatrices SOM et l'algorithme des K-moyennes. Une première abstraction est obtenue en créant un ensemble prototype en utilisant le plus souvent les cartes auto-organisées de Kohonen. Le second niveau classe ensuite les prototypes obtenus à l'aide des K-moyennes.

Wu et Chow (2004) utilisèrent une approche à deux niveaux en combinant les cartes auto-organisatrices de Kohonen SOM et la classification hiérarchique ascendante à la place de l'algorithme des K-moyenne.

L'avantage d'une telle approche est la réduction du temps de calcul et un gain d'espace mémoire appréciable, en particulier lorsque la quantité de donnée à classer est très importante (Vesanto et Alhoniemi, 2000 ; Guérif, 2006).

3.5. Evaluation et critères de validité

La validation des classes obtenues par les algorithmes de classification automatique permet de répondre aux questions suivantes : y'a-t-il un partitionnement optimal pour les données traitées et combien y-t-il de classes optimales ? (Jain and Dubes, 1993 ; Gordon, 1998 ; Halkidi et al., 2002).

Si le nombre de classes n'est pas connu a priori, un indice de validité des classes peut aider à trouver le nombre optimal de classes. La méthode de validation s'effectue comme suit : On fixe au départ les paramètres d'entrées de l'algorithme de classification. On exécute l'algorithme en variant le nombre de classes de 2 jusqu'à une valeur maximale c_{max} (Halkidi et al., 2002).

Un indice de validité tient compte de deux aspects : la compacité et la séparabilité des classes. La compacité est un indicateur de la variation des données au sein de la classe alors que la séparabilité est un indicateur sur le niveau de séparation des classes entre-elles.

De nombreux indices de validité ont été développés pour évaluer la qualité des partitions avec comme objectif principal un partitionnement optimal (Halkidi et al., 2002).

Pour la classification stricte ou dure, les indices bien connus dans la littérature sont l'indice de Dunn (Dunn, 1974), l'indice de Calinski-Harabasz (Calinski et Harabasz, 1974) et l'indice de Davis-Bouldin (Davies and Bouldin, 1979). Pour la classification floue, les indices les plus connus sont le coefficient de partition (Bezdek, 1974a), l'entropie de partition (Bezdek, 1974b), l'indice de partition (Bensaid et al., 1966) et l'indice de Xie et Beni (Xie and Beni, 1991).

Dans le cas d'une classification dure, les classes obtenues sont compactes et bien séparés contrairement à la classification floue où les classes peuvent s'enchevêtrer. Dans les sections suivantes, nous allons présenter des indices de qualité pour l'algorithme de Kohonen et quelques indices de validités qui sont pertinents pour la classification dure et floue.

3.5.1. Indices propres aux cartes auto-organisées.

De nombreux indices de qualité ont été développés pour les cartes auto-organisées, nous introduisons ici deux types d'erreur : l'erreur de quantification et le taux d'erreur topologique (Kohonen 2001).

3.5.1.1. Erreur de quantification

Les cartes auto-organisées font partie des méthodes de quantification vectorielle et il semble donc naturel de les évaluer à l'aide de l'erreur de quantification moyenne que l'on définit ainsi :

$$Q_{err} = \frac{1}{N} \times \sum_{i=1}^N \|x_i - w_{b(i)}\| \quad (3.24)$$

Où $b(i)$ est l'indice du prototype le plus proche de l'observation x_i .

3.5.1.2. Taux d'erreur topologiques

Les cartes auto-organisées sont aussi une méthode de projection de données multidimensionnelles sur un espace de faible dimension et le taux d'erreurs topologiques permet de quantifier la conservation de la topologie locale de l'espace des observations par la carte. On considère qu'il y'a une erreur topologique chaque fois que les deux prototypes les plus proches d'une observation ne sont pas voisins sur la carte. Le taux d'erreur topologique peut se définir ainsi :

$$T_{err} = 1 - \sum_{i=1}^N \left(\frac{1}{N(b(i))} \right) \left(\arg_{j \neq i} \min \|x - w_j\| \right) \quad (3.25)$$

où $\left(1 \mid N(b(i))\right)$ est la fonction indicatrice de l'ensemble des voisins du prototype le plus proche de l'observation x_i et w_j le vecteur poids.

3.5.2. Indices propres à la classification dure

3.5.2.1. Erreur Quadratique Moyenne

L'erreur quadratique moyenne (*Mean Squared Error MSE*) est une mesure de compacité très répandue, elle est notamment équivalente à la fonction de coût de l'algorithme des K-moyennes présenté au paragraphe 3.3.2.1

$$MSE = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^c c_{ij} \times \|x_i - w_j\|^2 \quad (3.26)$$

Où c est le nombre de groupes, w_j centre de groupe C_j et où c_{ij} indique si $x_i \in C_j$.

Lorsqu'on étend cette mesure au cas des partitions floues, on retrouve (à un coefficient multiplicateur près) la fonction de coût optimisée par l'algorithme des K-moyennes floues donnée par l'équation (3.16) :

$$FMSE = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^c (\mu_j(x_i))^f c_{ij} \times \|x_i - w_j\|^2 \quad (3.27)$$

3.5.2.2. Indice de Dunn

Dans le cas de la classification dure, l'indice de Dunn (Dunn, 1974) tient compte à la fois de la compacité et de la séparabilité des groupes : la valeur de cet indice est d'autant plus faible que les groupes sont compacts et bien séparés. Notons que la complexité de l'indice de Dunn devient prohibitive dès qu'on manipule de grands ensembles d'objets ; il est conséquent rarement utilisé.

$$I_{Dunn} = \frac{\min\{D_{min}(C_i, C_j) : i \neq j\}}{\max\{S_{max}(C_i)\}} \quad (3.28)$$

Où $D_{min}(C_i, C_j)$ est la distance minimale qui sépare un objet du groupe C_i d'un objet du groupe C_j et où $S_{max}(C_i)$ est la distance maximale qui sépare deux objets du groupe C_i :

$$D_{min}(C_i, C_j) = \min\{\|x - y\| : x \in C_i \text{ et } y \in C_j\} \quad (3.29)$$

$$S_{max}(C_i) = \max\{\|x - y\| : (x, y) \in C_i \times C_i\} \quad (3.30)$$

3.5.2.3. Indice de Davies-Bouldin

Dans le cas d'une classification dure, l'indice de Davies-Bouldin (Davies et Bouldin, 1979) tient compte à la fois de la compacité et de la séparabilité des groupes : la valeur de cet indice est d'autant plus faible que les groupes sont compacts et bien séparés. La compacité signifie que les individus d'une classes devaient être très proches les uns par rapport aux autres. La séparation elle signifie que chaque classe doit être nettement séparée des autres.

Cet indice dont la complexité est $\theta (K \times (N + K))$ favorise les groupes hypersphériques. Il est particulièrement bien adapté pour une utilisation avec la méthode des K-moyennes.

$$I_{DB} = \frac{1}{K} \sum_{k=1}^K \max_{l \neq k} \left\{ \frac{S_c(C_k) + S_c(C_l)}{(D_{ce}(C_k, C_l))} \right\} \quad (3.31)$$

Où $S_c(C_i)$ est la distance entre un objet du groupe C_i et son centre, et où $D_{ce}(C_i, C_j)$ est la distance qui sépare les centres des groupes C_i et C_j :

$$S_c(C_i) = \frac{1}{N_i} \sum_{i=1}^{N_i} \|x - w_i\| \quad (3.32)$$

$$D_{ce}(C_i, C_j) = \|w_i - w_j\| \quad (3.33)$$

3.5.3. Indices propres à la classification floue

3.5.3.1. Coefficient de Partition (Partition Coefficient (PC))

Bezdek a conçu cet indice de validité pour mesurer le niveau d'enchevêtrement entre classes (Bezdek, 1974a).

$$PC = \frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n u_{ij}^2 \quad (3.34)$$

L'inconvénient du coefficient de partition est sa décroissance monotone avec le paramètre c .

3.5.3.2. Entropie de classification (Classification Entropy (CE))

L'entropie de classification (Bezdek, 1974b) est un autre indice de la partition de la classe similaire à l'indice du coefficient de partition.

$$CE = - \frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n u_{ij} \log(u_{ij}) \quad (3.35)$$

3.5.3.3. Indice de Partition (Partition index (SC))

C'est un indice qui tient compte de la compacité et la séparation des classes (Bensaid, et al., 1996).

$$SC = \sum_{i=1}^c \frac{\sum_{j=1}^N (u_{ij})^m \|x_j - v_i\|^2}{N_i \sum_{k=1}^c \|v_k - v_i\|^2} \quad (3.36)$$

3.5.3.4. Indice Xie et Beni (Xie and Beni's index (XB))

Dans le cas d'une classification floue, il est fréquent d'utiliser l'indice Xie et Beni (Xie et Beni, 1991) pour tenir compte de la compacité et la séparabilité des groupes. Il quantifie le rapport de la variation totale des classes et la séparation des classes .

$$XB(c) = \frac{\sum_{i=1}^c \sum_{j=1}^N (u_{ij})^m \|x_j - v_i\|^2}{N \min_{i,j} \|x_j - v_i\|^2} \quad (3.37)$$

3.6. Résumé et Conclusion

Ce chapitre a introduit les concepts de base des algorithmes de classification non supervisée, les approches K-moyennes, C-moyennes floues, les cartes auto-organisées ainsi que les indices de validités correspondantes à ces méthodes.

L'algorithme des K-moyennes, facile à implémenter souffre du choix du nombre de classes et des conditions initiales pour sa convergence vers la solution finale. L'algorithme des K-moyennes floues, extension des K-moyennes étant beaucoup moins sensible au bruit dans les stimuli grâce aux partages de l'appartenance entre les différentes classes. Ceci résulte dans bien des situations de rester pris dans des minimums locaux de la fonction objective. L'algorithme est donc également beaucoup moins sensible à une mauvaise initialisation des centres. L'algorithme des cartes auto-organisées offre, quand a lui, une visualisation des classes obtenues.

Des indices de qualités et de validités pertinents sont présentés pour évaluer respectivement les cartes auto-organisées et les classifications dures et floues. L'erreur de quantification et le taux d'erreur topologiques sont retenues pour la première, les indices de Dunn, Davies-Bouldin sont présentées pour la classification dure et les indices Coefficient de partition, Entropie de classification, indice de partition et Xie et Beni pour la classification floue.

Dans le cas d'une classification dure, les classes obtenues sont compactes et bien séparés contrairement à la classification floue ou les classes peuvent s'enchevêtrés.

Une combinaison entre les cartes-auto organisées et une méthode partitionnée (K-moyenne ou C- moyenne floue) est utile pour pallier aux inconvénients des méthodes partitionnées dues à leur sensibilité aux conditions initiales et la difficulté du choix du nombre de classes (Vesanto et Alhoniemi, 2000). Ainsi les données sont classées avec les cartes auto-organisées dans un premier temps en prototypes qui seront par la suite traités par une approche de partitionnement (K moyennes ou C-moyennes floues). Ceci permet une réduction du temps de calcul et un gain d'espace mémoire appréciable (Vesanto et Alhoniemi, 2000).

Dans le chapitre suivant, nous allons nous intéresser à l'identification de types de jours de la charge électrique nationale en utilisant les approches présentées dans ce chapitre.

4

Identification par type de jour de la charge électrique nationale Algérienne

Sommaire

4.1 Introduction	43
4.2 Approche de segmentation de la charge électrique	43
4.3 Identification par type de jour de la charge électrique nationale Algérienne avec les les cartes de Kohonen	44
4.4 Identification à deux niveaux par type de jour de la charge électrique nationale	
Algérienne : Approche à deux niveaux : SOM et K-moyenne	48
4.5 Conclusion et contributions	52

4.1. Introduction

Dans ce chapitre, nous allons décrire l'identification de type de jour pour la charge électrique nationale Algérienne.

Après une définition du concept de la segmentation des données de la charge électrique, nous axerons en particulier sur l'identification par type de jour.

Nous proposons une analyse en deux axes :

- La première concerne l'identification de type de jour de la charge électrique nationale Algérienne en utilisant uniquement l'approche classique des cartes auto-organisées SOM.
- La seconde, une méthode hybride à deux phases combinant SOM et une classification dure : K-moyenne en utilisant l'indice de validité Davies-Bouldin.

Le but de cet analyse est de pouvoir mettre en évidence les avantages et les inconvénients de chacune de ces approches et de pouvoir proposer celle qui nous semble la plus adaptée.

4.2. Approche de segmentation de la charge électrique

La segmentation des données de la charge électrique a pour but le partage de l'ensemble des données en plusieurs ensembles distincts modélisés séparément. Cette approche rappelle le fameux adage diviser pour régner. Cette segmentation est avantageuse si les ensembles sont suffisamment différents. Les données de la charge électrique peuvent être segmentées par :

- Le type de jour.
- L'heure du jour.

La segmentation par type de jour a fait l'objet de beaucoup de travaux de recherche (Muller et Schatzel, 1999 ; Muller et Petrisch, 1998 ; Ho et al., 1990 ; Bretschneider et al., 1999 ; Hsu et Yang, 1991). Par exemple, Muller et Schatzel ont identifiés cinq classes primaires :

- Les journées d'été.
- Jours d'été précoces et tardifs.
- Jour de printemps et d'automne
- Les jours d'hiver précoces
- Jours d'hiver

Chacune de ces classes primaires peut comporter des classes secondaires. Le niveau de segmentation pour le type de jour est subjectif et dépend dans une large mesure du prévisionniste.

Comme l'a souligné (Hubele et cheng, 1990), l'application d'un modèle de prévision de la charge séparée pour chaque saison (été, automne, hiver, été) a l'avantage de ne pas inclure l'information saisonnière dans le modèle.

Une segmentation plus poussée de la charge par jour de semaine (par exemple le dimanche d'été, le dimanche d'hiver, lundi d'été, etc.) réduit encore la quantité d'information qu'un

modèle doit incorporer. De telles approches ont été mise en œuvre avec succès par (Srinivasan et al., 1999) et (Mastorocostas et al., 1999).

Pour un modèle unique pour toutes les données, l'information du type de jour est souvent incorporée comme entrée supplémentaire. (Par exemple Chen et al., 1992 ; Lertpalangsunti et Chan, 1998).

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

- L'entretien avec les opérateurs de la société (Exemple, Ho *et al.*, 1990)
- Les algorithmes de classification automatique (Par exemple (Bretschneider *et al.*, 1999) (voir aussi chapitre 3)
- Les cartes auto-organisées de Kohonen (par exemple (Hsu et Yang, 1991) et (Pelikan *et al.*, 1996). (Voir paragraphe 4.3).

La sélection par heure du jour est une autre méthode de segmentation de la charge électrique. Des modèles de prédiction ont été développés et sont classés en deux catégories : les modèles séquentiels développés par Choueiki et Papadakis (Choueiki *et al.*, 1997 ; Papadakis *et al.*, 1999) et les modèles parallèles traités par (Hill, 1998 ; Gupta, 1985).

Le premier type de modèle modélise la charge horaire comme une série unique alors que la seconde approche modélise chaque heure de la journée par une série. Notons également une l'approche qui combine les deux, Cette approche dite Multi-timeScale (MTS) (Gupta, 1985; Khotanzad *et al.*, 1996 ; Murray *et al.*, 2000).

4.3. Identification par type de jour de la charge électrique nationale Algérienne avec les cartes de Kohonen

Les cartes auto-organisatrices de Kohonen (Kohonen, 1990) sont idéales pour l'identification de type de jours puisque le nombre de types de jours n'est pas connu à priori. Les cartes de Kohonen, dans ce cas peuvent être appliquées de plusieurs manières (Hsu et Yang, 1991 ; Muller et Petrisch, 1998 ; Bretschneider *et al.*, 1999). Cependant, nous n'avons pas noté de différences significatives au niveau des résultats. L'algorithme utilisé par Hsu et Yang a été choisi pour notre application (Hsu et Yang, 1991).

L'approche par Kohonen consiste à faire une projection des données initiales sur un espace discret et régulier de faible dimension. Le réseau SOM est composé de deux couches : la couche d'entrée contenant les données et la couche de sortie représentée par un treillis de neurone selon une géométrie prédéfinie. Chaque neurone de la couche topologique est totalement connecté aux neurones de la couche d'entrée. Le vecteur poids de ces connexions forme le référent ou le prototype associé à ce neurone.

La mise en œuvre de cette carte auto-organisée est identique aux travaux de Hsu et Yang [Hsu et Yang, 1991]. Le réseau est composé d'une grille de nœuds en sorties connectés aux nœuds d'entrées via un ensemble de poids. Quand le vecteur d'ordre k $P_k \in R^{l \times n}$, le réseau calcule l'activation de chaque nœud par P_k :

$$a_{i,j,k} = W_{i,j} P_k \quad (4.1)$$

Où $a_{i,j,k}$ et $W_{i,j}$ sont respectivement l'activation et le poids ($\in R^{lxn}$) connectés à P_k au nœud i et j respectivement.

P_k est dit mappé sur le nœud avec la plus grande activation. Après plusieurs vecteurs d'entrées sont présentés, des entrées similaires sont mappés aux mêmes nœuds ou nœuds adjacents, C'est-à-dire avec le plus petit voisinage. Un voisinage de taille N_c au environ du nœud i, j est défini comme les nœuds $i \pm N_c$ N_c au nœud $j \pm N_c$. P_k pour notre étude est construit en deux étapes. La courbe de la charge journalière est extraite Initialement de chaque jour pour donner un ensemble de courbes de charge qui a une valeur minimale nulle et une valeur maximale égale à un (Hsu et Yang, 1991).

$$Y'(i)_k = \frac{Y(i)_k - \min Y_k}{\max Y_k - \min Y_k} \quad i = 1, \dots, 24 \quad (4.2)$$

Où $Y'(i)_k$ et $Y(i)_k$ sont les éléments d'ordre i de la courbe de la charge $Y'_k \in R^{1 \times 24}$ et la charge actuelle $Y_k \in R^{1 \times 24}$ du jour k respectivement. Les courbes de la charge électrique sont normalisées avec une longueur d'une unité.

$$P(i)_k = \frac{Y^i(i)_k}{\left(\sum_{j=1}^{24} Y_k^2\right)^{1/2}} \quad i = 1, \dots, 24 \quad (4.3)$$

Où $P(i)_k$ est le i ème élément de P_k . Les poids sont initialisés comme suit :

$$W_{i,j} = \left\| \left[\left(\mu_p(1) \right), \dots, \left(\mu_p(24) \right) \right] + 5u \left[\left(\rho_p(1) \right), \dots, \left(\rho_p(24) \right) \right] \right\| \quad (4.4)$$

Où μ_p et ρ_p sont la moyenne et l'écart type de $P(i)$ pour tout k , u est un nombre aléatoire uniformément distribué dans l'intervalle -0.5 à 0.5 et $W_{i,j}$ est normalisé à un (Hsu et Yang, 1991). Les poids ne sont pas initialement aléatoirement mais initialisés aux alentours de la moyenne des entrées. Durant l'entraînement, les entrées sont présentés un par un et les poids du nœud déclenché (le nœud pour lequel les entrées est mappé) ainsi que les nœuds des ses voisins sont mis à jour selon la formule :

$$W_{i,j}(m+1) = W_{i,j}(m) + \alpha(m) [P_k - W_{i,j}(m)] \quad (4.5)$$

Où α est le gain d'activation avec $0 < \alpha < 1$ et m , le nombre d'itération. Cela a pour effet d'accroître l'activation du nœud déclenché ainsi que ses voisins. Dans une simple itération, toutes les entrées sont présentés et les poids sont mis à jour. Après plusieurs itérations, la taille du voisinage est décrétementé de un et ainsi de suite jusqu'à atteindre la valeur nulle.

L'algorithme 1 rappelle les grandes lignes de cette approche.

Algorithme 1 Approche SOM

1. Fixer les paramètres de la carte SOM : N_c , la taille du voisinage initiale, la valeur de l'adaptation du gain, le nombre d'itération et la taille de la grille.
2. Présenter en entrée des données normalisées au réseau.
3. Déterminer le neurone gagnant : celui qui a la plus petite distance.

4. Mettre à jour les vecteurs poids.
5. Répéter les étapes 2 à 4 jusqu'à stabilisation de l'auto-organisation.

Les tests effectués pour l'identification de types de jours concernent la consommation électrique horaire Algérienne pour les années allant de 2000 à 2004 (Benabbas et al., 2008). Les cartes auto-organisatrices ont été appliquées avec les paramètres suivants : N_c , la taille du voisinage initiale de valeur 4, l'adaptation du gain de valeur $\alpha = 0.002$, le nombre d'itération $m = 10$ et la taille de la grille de dimension 18×18 . La Figure 4.1 montre les nœuds déclenchés identifiés pour les années allant de 2000 à 2004.

A noter que les paramètres ci-dessus ont été choisis empiriquement après une batterie de tests portant sur l'apprentissage des Cartes de Kohonen, sauf exception de la taille de la carte qui a été sciemment exagéré pour une meilleure visualisation du comportement de la charge et des shifts été-hiver, jours de travail jours fériés etc.

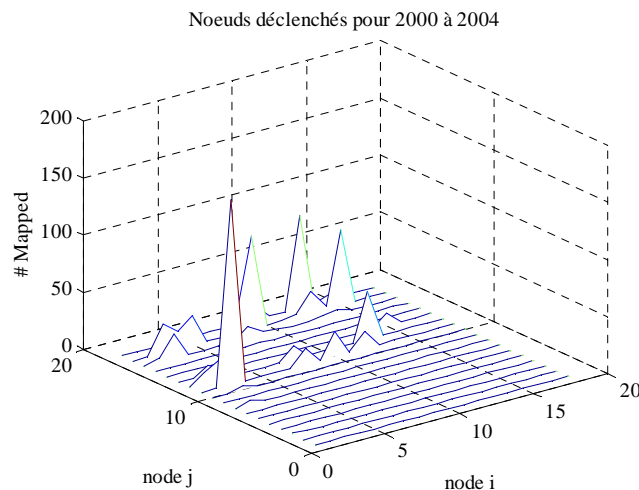


Figure 4.1 - Nœuds déclenchés pour les jours de l'année 2000 à 2004.

On remarque que les nœuds déclenchés se situent pour le nœud i entre 0 et 17 et le nœud j entre 10 et 20.

Nous analysons à présent l'identification hebdomadaire (Figure 4.2) et saisonnière des types de jours (Figure 4.3). Les nœuds déclenchés pour les jours de semaine (de Samedi à Mercredi) occupent la même position au niveau de la carte avec un pic prédominant à la position $i = 2$ et $j = 11$ alors que les jours de Week-ends jeudi et vendredi occupent un emplacement totalement différent. Pour le vendredi, la différence est nette à cause de la diminution d'activité et la prière du vendredi.

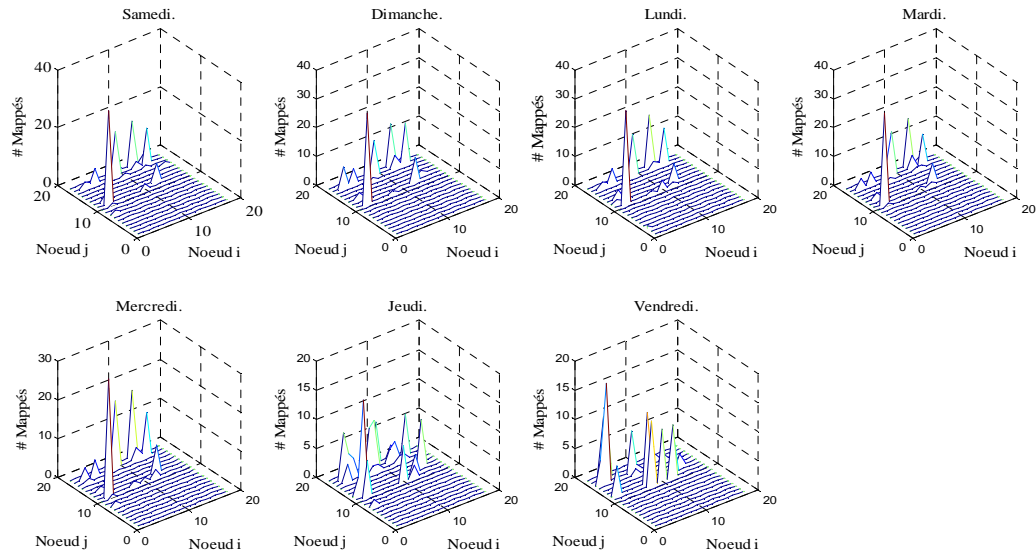


Figure 4.2 - Identification des types de jours pour une semaine (Samedi à Vendredi).

La comparaison saisonnière donne un changement de direction des pics selon les mois de l'année. Les nœuds déclenchés des mois d'hiver occupent la partie droite en haut de la carte. Les nœuds déclenchés se déplacent ensuite de la droite vers la gauche pour toute la période d'hiver (Décembre à Février). Les mois du printemps sont bien identifiés et occupent le milieu gauche de la carte. Les nœuds déclenchés des mois de septembre jusqu'à octobre se déplacent du milieu de la carte vers le haut de la carte. Les jours de ces mois se comportent comme des jours d'hiver et d'été. Le nombre de cluster déterminé visuellement est égal à 6 (voir figure 4.1).

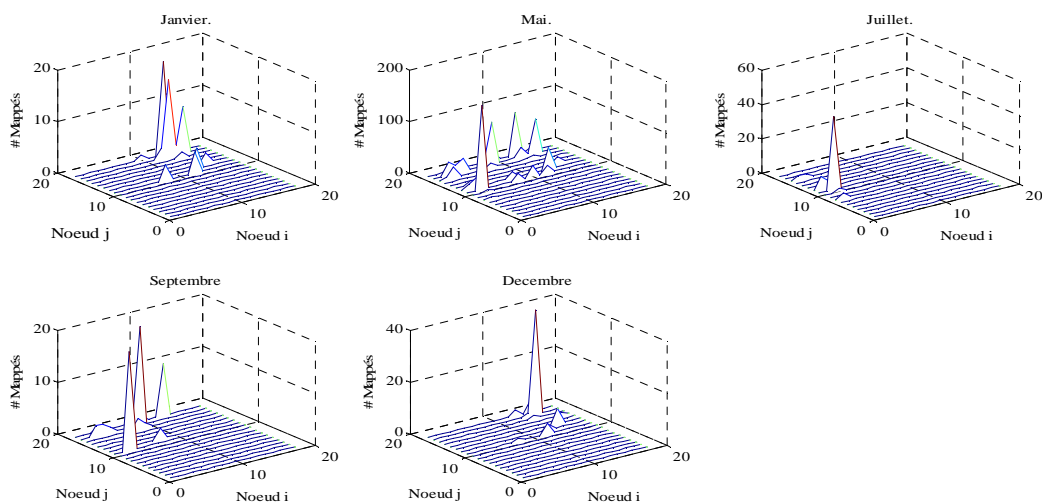


Figure 4.3 - Identification de type de jours saisonnière (Janvier à Décembre).

4.4. Identification à deux niveaux par type de jour de la charge électrique nationale Algérienne : Approche à deux niveaux : SOM et K-moyenne.

Une carte-auto-organisée peut être vue comme une méthode de classification automatique dont résulte une partition de l'espace d'observations qui comporte autant de parties qu'il y a de neurones. Il est souvent souhaitable de diminuer le nombre de classes pour en faciliter

l'analyse. Plusieurs méthodes de découpages ont ainsi été proposées (Vesanto et Alhionemi, 2000).

La méthode des K-moyenne est une autre méthode de classification. Son principe consiste à choisir arbitrairement une partition. Ensuite, les exemples sont examinés un à un. Si un exemple devient plus proche du centre d'une classe autre que la sienne, il est déplacé vers cette nouvelle classe. On recalcule, ensuite les centres des nouvelles classes et on réaffecte les exemples aux partitions, et ainsi de suite jusqu'à obtenir une partition stable.

Le critère à minimiser dans ce cas est défini par :

$$J(U, M) = \sum_{j=1}^n \sum_{i=1}^c u_{ij} \|x_j - m_i\|^2 \quad (4.6)$$

où n est le nombre total d'objets de l'ensemble des données $= \{x_1, x_2, \dots, x_n\}$, c est le nombre de classes de l'ensemble des centres des classes $V = \{m_1, m_2, \dots, m_c\}$, u_{ij} est un élément de la matrice de partition, $\|x_i - m_j\|$ désigne la norme euclidienne entre x_i et m_j .

L'algorithme initial nécessite de fixer à priori le nombre c de classes souhaitées. Néanmoins, (Vesanto et Alhionemi, 2000) ont proposé de déterminer automatiquement une valeur de c en retenant la partition qui minimise l'indice de Davies-Bouldin (Davies et Bouldin, 1979) défini par :

$$I_{DB} = \frac{1}{K} \sum_{k=1}^K \max_{l \neq k} \left\{ \frac{S_c(C_k) + S_c(C_l)}{D_{ce}(C_k, C_l)} \right\} \quad (4.7)$$

Où $S_c(C_i)$ est la distance moyenne entre un objet dur groupe C_i et son centre, et où $D_{ce}(C_k, C_l)$ est la distance qui sépare les centres des groupes C_i et C_j et K le nombre de classes.

$$S_c(C_i) = \frac{1}{N_i} \sum_{i=1}^{N_i} \|x - w_i\| \quad (4.8)$$

$$D_{ce}(C_i, C_j) = \|w_i - w_j\| \quad (4.9)$$

La méthode des K-moyenne associée à l'indice de Davies-Bouldin recherche une partition de l'espace des observations dont les différentes parties sont compacte et séparées.

Nous avons également appliqué l'algorithme SOM sur les mêmes données, la carte SOM obtenue est découpée à l'aide de la méthode des K-moyennes et le nombre de classes a été sélectionné à l'aide de l'indice de Davies-Bouldin. L'algorithme 2 rappelle les grandes lignes de cette approche

Algorithme 2 Approche à deux phases SOM et K-moyenne

1. Définir les paramètres pour l'architecture SOM

2. Application de l'algorithme SOM pour générer les prototypes
3. Segmentation des prototypes de SOM avec l'algorithme des K-moyenne
4. Exécuter l'algorithme K-moyenne en variant le nombre de cluster de 2 à nbmax.
5. Choisir K qui minimise l'indice de Davies-Bouldin
6. Visualisation et interprétation des résultats.

L'analyse par l'algorithme des K-moyenne est appliquée sur les sorties obtenues par les cartes auto-organisatrices sur le codebook. La sélection du nombre adéquat de cluster est accompli en exécutant plusieurs fois l'algorithme pour chaque valeur de k .

Le tableau 4.1 montre l'indice de Davies-Bouldin DB et le nombre optimal de classes obtenus pour l'Algérie. Le nombre de classes optimal est 6. Cette valeur est obtenue pour la plus petite valeur de l'indice de Davies-Bouldin qui est 0.75574. L'évolution de l'indice DB en fonction du nombre de classes est illustrée sur le graphique 4.4a. La figure 4.4b montre la segmentation de la carte SOM en six classes.

Pays	Algérie
classe	6
DB indice	0.75574

Tableau 4.1 - Nombre de classe et indice de Davies-Bouldin.

Pour plus de détails sur les classes obtenues, une étude quantitative a été mise en œuvre.

Ainsi, le tableau 4.2 indique le nombre d'items par classes pour l'Algérie. Notons par exemple la prédominance des classes C2 et C3 par rapport aux autres en terme de nombre de jours. Ces résultats sont reproduits sous forme d'histogrammes sur la figure 4.4c.

Classe	C1	C2	C3	C4	C5	C6
Items	72	64	187	183	124	101

Tableau 4.2 - Nombre d'items par classes pour l'Algérie.

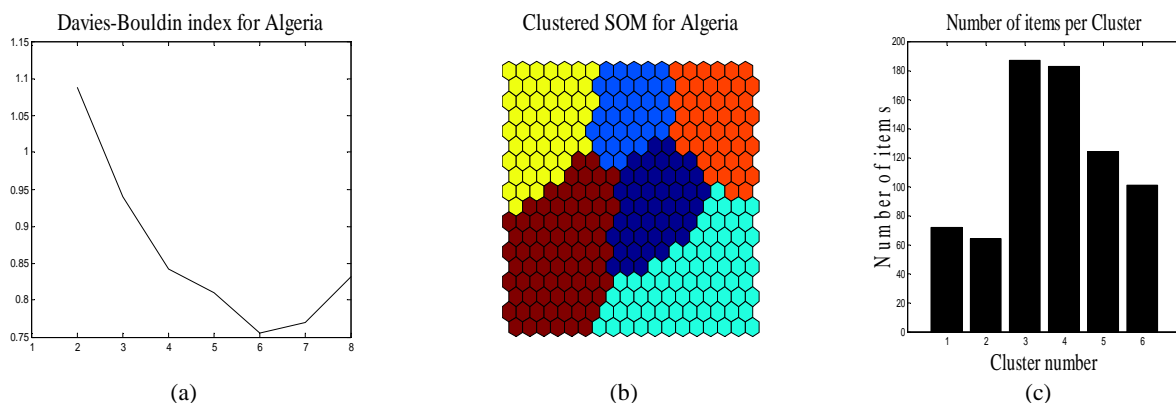


Figure 4.4 – Evolution de l'indice DB minimum 0.75574 atteint pour 6 classes (a), Segmentation de SOM avec k-moyenne en 6 classes (b) et Nombre d'items par classe pour l'Algérie (c).

Les tableaux 4.3 et 4.4 montrent respectivement les distributions hebdomadaires et mensuelles par classes pour l’Algérie. Dans le tableau 4.2, la classe C2 contient un nombre important de Vendredi, ce qui laisse supposer que cette classe représente des week-ends.

Jour	Sam	Dim	Lun	Mar	Mer	Jeu	Ven
C1	10	10	09	13	09	21	00
C2	01	02	00	03	02	18	38
C3	32	35	32	32	32	21	03
C4	30	31	30	31	31	26	04
C5	10	09	11	09	11	14	60
C6	21	17	22	16	20	05	00

Tableau 4.3 - Distribution hebdomadaire par classes pour l’Algérie.

Mois	J	F	M	A	M	J	J	A	S	O	N	D
C1	02	22	37	06	00	00	00	00	03	02	00	00
C2	00	00	00	15	21	09	05	09	05	00	00	00
C3	48	23	06	00	00	00	00	00	10	31	15	54
C4	00	00	00	00	14	51	57	53	08	00	00	00
C5	12	12	09	04	00	00	00	00	06	28	45	08
C6	00	00	10	35	27	00	00	00	28	01	00	00

Tableau 4.4 - Distribution mensuelle par classes pour l’Algérie.

Les figures 4.5 et 4.6 montrent respectivement sous forme d’histogrammes la distribution hebdomadaire et mensuelle par classes pour l’Algérie.

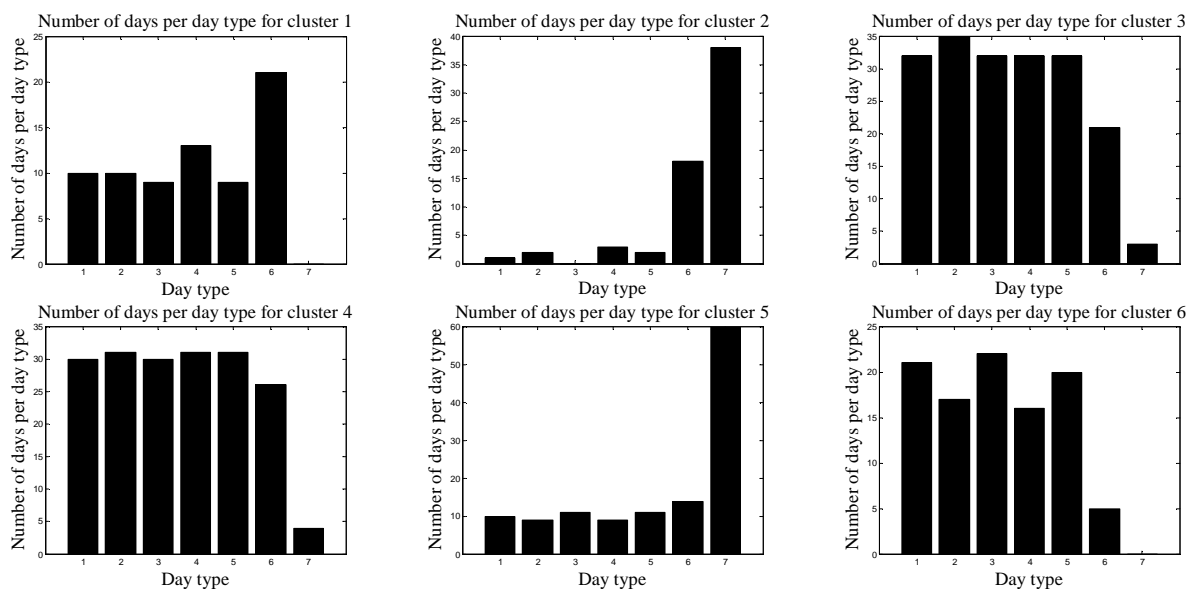


Figure 4.5 – Distribution hebdomadaire des classes de l’Algérie.

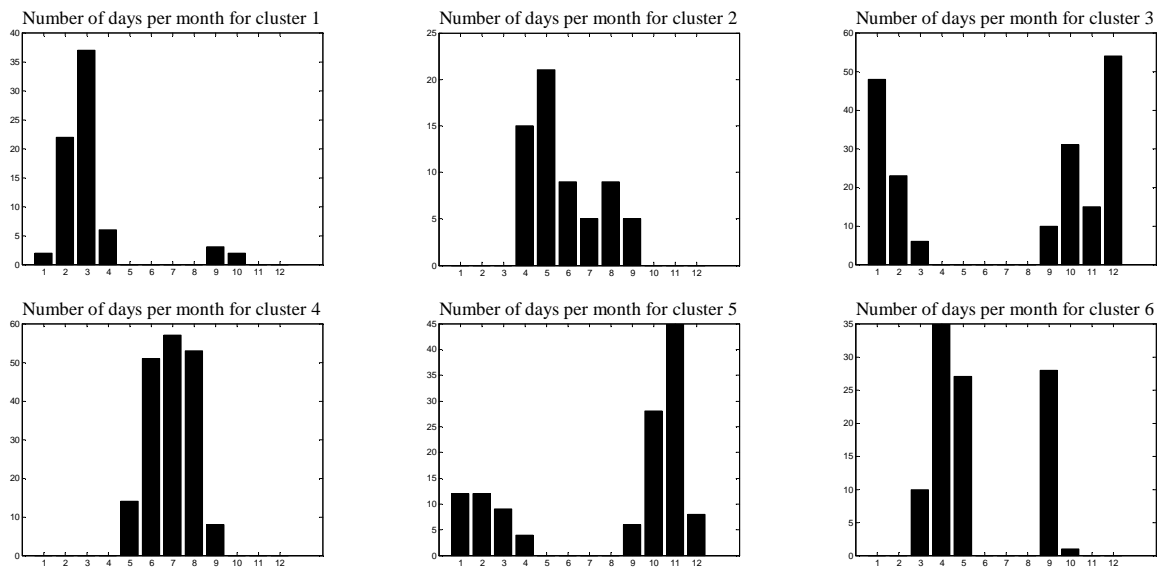


Figure 4.6 – Distribution mensuelle des classes de l'Algérie.

4.5. Conclusion et contributions

Dans ce chapitre, nous avons étudié l'identification de types de jours pour la consommation électrique algérienne avec des méthodes non supervisées.

L'analyse par les cartes auto-organisatrices de Kohonen permet une identification visuelle des différentes classes. Cette approche qualitative a permis de capturer les différences entre jours de travail et weekend, les jours d'été et d'hiver ainsi que les jours de vacances.

L'algorithme des K-moyenne combiné aux cartes de Kohonen avec l'indice de validité de Davies-Bouldin a permis de déterminer le nombre de classes égale à 6. Une analyse quantitative a permis de donner plus d'informations sur les classes obtenues. Cette approche s'applique dans le cas où les classes sont hyper-sphériques compactes et bien séparées.

Cependant, l'étude de la charge électrique globale n'est pas significative par rapport aux spécificités de chaque région. De plus, l'application de SOM est tributaire d'un bon choix des paramètres permettant de définir la carte SOM. Aussi, il importe de faire un bon choix des paramètres de la carte de Kohonen.

Le chapitre suivant proposera une étude de l'identification des jours par rapport aux régions, les approches à deux phases présentées utiliseront SOM et K-moyenne d'une part et SOM et C-moyenne flou, d'autre part en tenant compte de critères de qualités pour le choix de la carte auto-organisées et des critères de validités adéquats pour le choix des paramètres de K-moyenne et C-moyenne flou. Une comparaison entre ces deux approches sera envisagée.

5

Identification par type de jour de la charge électrique régionale Algérienne

Sommaire

5.1 Introduction	54
5.2 Choix des paramètres et critères de qualités de SOM	54
5.3 Identification à deux niveaux par type de jour de la charge électrique régionale Algérienne.....	56
5.3.1 Approche à deux niveaux avec la classification dure	56
5.3.2 Approche à deux niveaux avec la classification floue	60
5.4 Conclusion et contributions	65

5.1. Introduction

Dans ce chapitre, nous allons détailler la mise en œuvre des différentes approches hybrides de la classification des types de jour de la charge électrique régionale en tenant compte d'une meilleure qualité de la carte de Kohonen et un choix optimal des indices de validité pour la classification dure et floue. Nous proposons une analyse en trois axes :

- La première étude concerne le bon choix des paramètres (Dimension de la carte, etc..) et les critères de qualités retenues pour l'application des cartes auto-organisatrices de Kohonen.
- Mise en œuvre de la méthode à deux phases combinant SOM et K-moyenne en utilisant l'indice de validité Davies-Bouldin.
- Mise en œuvre de la méthode à deux phases combinant SOM et C-moyenne flou avec les indices de validité flou.

Le but de cet analyse est de pouvoir mettre en évidence les avantages et les inconvénients de chacune de ces approches et de pouvoir proposer celle qui nous semble la plus adaptée et enfin identifier clairement les classes de jours de la charge électrique.

5.2. Choix des paramètres et critères de qualités de SOM.

L'approche des cartes de Kohonen est une projection des données initiales sur un espace discret et régulier de faible dimension (2D ou 1D). Le réseau SOM est composé de deux couches. La couche d'entrée formée des données d'entrées et la couche de sortie qui est un treillis de neurones selon géométrie prédéfinie. Chaque neurone de la couche topologique est totalement connecté aux neurones de la couche d'entrée. Le vecteur poids de ces connexions forme le référent ou le prototype associé à ce neurone (Kohonen, 1990).

La forme de la maille à la base du réseau de neurones est la plus souvent rectangulaire ou hexagonale. Un maillage rectangulaire est particulièrement adapté pour des applications nécessitant la visualisation des données traitées.

Pour déterminer les dimensions de la carte - nombre de neurones en largeur, noté x , et en hauteur, noté y , on utilise une heuristique faisant intervenir le nombre C d'individus utilisés pour l'apprentissage et les deux plus grandes valeurs propres, V_1 et V_2 , de la matrice de covariance de la base d'apprentissage (Vesanto, 2000).

- Le nombre total de neurones N_n de la carte est approximé à $N_n = 5 \times C^{0.54321}$
- En posant $R = \sqrt{V_1 V_2}$, on résout alors pour x et y le système d'équations : $xy = N_n$ et $x/y = R$

Les tests effectués pour l'identification de types de jours de la consommation électrique concernent deux villes, Alger situé au nord du pays et Hassi Messaoud dans le sud pour les années 2003 et 2004.

Ainsi le tableau 5.1 résume les paramètres utilisés par SOM :

Paramètre	Alger	Hassi Messoud
taille	19x7	17x8
lattice	hexa	hexa
shape	sheet	sheet
training	batch	batch

Tableau 5.1 - Paramètres de SOM.

L'algorithme 3 rappelle les grandes lignes de l'algorithme SOM appliqué :

Algorithme 3 Approche SOM

5. Déterminer la topologie optimale de la carte SOM.
 6. Présenter en entrée des données au réseau.
 7. Déterminer le neurone gagnant : celui qui a la plus petite distance.
 8. Mettre à jour les vecteurs poids.
 6. Répéter les étapes 2 à 4 jusqu'à stabilisation de l'auto-organisation.
-

Deux critères de qualités sont retenus pour l'application de SOM, ce sont :

- Q_e : Erreur de quantification moyenne, elle mesure la résolution de SOM et est calculé la distance euclidienne entre le vecteur d'entrée et le BMU (Best Matching Unit) du vecteur référent
- T_e : Taux d'erreur topologique, quantifie la préservation de la topologie des données. On considère qu'il y'a une erreur topologique chaque fois que les deux prototypes les plus proches d'une observation ne sont pas voisins sur la carte.

Le tableau 5.2 indique les valeurs respectives de Q_e et T_e pour Alger et Hassi Messaoud.

Mesures de qualités	Algiers	Hassi Messaoud
Q_e	0.3236	0.2936
T_e	0.0328	0.0533

Tableau 5.2 - Mesures de qualités de SOM.

5.3. Identification à deux niveaux par type de jour de la charge électrique régionale algérienne.

Nous allons étudier dans un premier temps l'identification de types de jours pour la charge électrique régionale en utilisant la classification dure puis dans un second temps l'identification avec l'analyse floue, conformément à l'approche à deux niveaux (Section 4.3).

5.3.1. Approche à deux niveaux avec la classification dure.

L'application directe des algorithmes de classification automatique sur les données nécessite un temps énorme d'exécution. Pour réduire ce temps, l'approche à deux phases a été proposée : elle consiste à appliquer en premier les cartes de Kohonen ce qui permet de générer les prototypes. Ensuite cette carte est segmentée avec un algorithme partitionné K-means.

L'algorithme des K-means est appliqué sur les prototypes. L'exécution de cet algorithme nécessite la connaissance du nombre de classes. Pour ce faire, nous avons utilisé des indices de validité propre à la classification dure.

De nombreux indices de validités peuvent être appliquées à la classification stricte ou dure. Citons les plus connus : l'indice de Davies-Bouldin (Davies et Bouldin, 1979), l'indice de Dunn (Dunn, 1979) et l'indice Calinski-Harabasz (Calinski et Harabasz, 1974). L'indice de Davies-Bouldin qui tient compte de la séparabilité et de la compacité des classes est choisi pour sa popularité et son temps rapide d'exécution (Jain et Dubes, 1988).

Pour trouver le nombre optimal de classes, deux propriétés importantes sont utilisées : la compacité et la séparation.

Le nombre optimal de classes est obtenu pour la plus petite valeur de l'indice de Davies-Bouldin qui s'exprime comme suit :

$$\frac{1}{C} \sum_{k=1}^C \min_{l \neq k} \frac{(S_c(Q_k) + S_c(Q_l))}{d_{ce}(Q_k, Q_l)} \quad (5.1)$$

Où C est le nombre de classes, $S_c(C_i)$ est la distance entre un objet du groupe C_i et son centre, et $D_{ce}(C_i, C_j)$ est la distance qui sépare les centres des groupes C_i et C_j :

$$S_c(C_i) = \frac{1}{N_i} \sum_{i=1}^{N_i} \|x - w_i\| \quad (5.2)$$

$$D_{ce}(C_i, C_j) = \|w_i - w_j\| \quad (5.3)$$

Cet indice est convenable pour l'évaluation des classes qui sont sphériques et compactes.

L'algorithme 4 rappelle les grandes lignes :

Algorithme 4 Approche à deux phases SOM et K-means

1. Déterminer la topologie de la carte SOM.
2. Application de SOM pour générer les prototypes.
3. Segmentation de SOM avec l'algorithme K-moyenne.
4. Exécuter l'algorithme K-moyenne en variant le nombre de cluster de 2 à nbmax
5. Choisir K qui minimise l'indice de Davies-Bouldin
6. Interprétation et visualisation des résultats.

Les tests effectués pour l'identification de types de jours concernent la consommation électrique horaire de l'Alger et Hassi Messaoud pour les années 2003 et 2004 (Benabbas et al., 2008).

Le tableau 5.3 montre l'indice de Davies-Bouldin DB et le nombre optimal de classes obtenus respectivement pour Alger et Hassi Messaoud. Le nombre de classes optimal est respectivement 4 pour Alger et 3 pour Hassi Messaoud. Ces valeurs sont obtenues pour les valeurs minimales de l'indice de Davies-Bouldin qui sont respectivement 0.8540 pour Alger et 0.7929 pour Hassi Messaoud. L'évolution de l'indice DB en fonction du nombre de classes pour les villes d'Alger et Hassi Messaoud est illustrée respectivement par les graphiques 5.1a et 5.2a.

Région	Alger	Hassi Messaoud
classe	4	3
DB indice	0.8540	0.7929

Tableau 5.3 - Nombre de classe et indice de Davies-Bouldin.

Une étude quantitative a permis de fournir plus de détails sur les classes obtenues pour chaque ville. Ainsi, les tableaux 5.4 et 5.5 montrent respectivement le nombre d'items par classes pour les villes d'Alger et Hassi Messaoud. Notons par exemple pour Alger, la prédominance de la classe C1 par rapport aux autres en termes de nombre de jours.

Classe	C1	C2	C3	C4
Items	274	151	222	84

Tableau 5.4 - Nombre d'items par classe pour Alger.

Classe	C1	C2	C3
Items	163	284	284

Tableau 5.5 - Nombre d'items par classe pour Hassi Messaoud.

fonction du nombre de classes. L'indice minimal de Davies-Bouldin DB de valeur 0.8540 est obtenue pour quatre classes. La figure 5.1b montre la segmentation de la carte SOM en quatre classes. La figure 5.1c est un histogramme indiquant le nombre d'items ou jours contenus dans chaque classe.

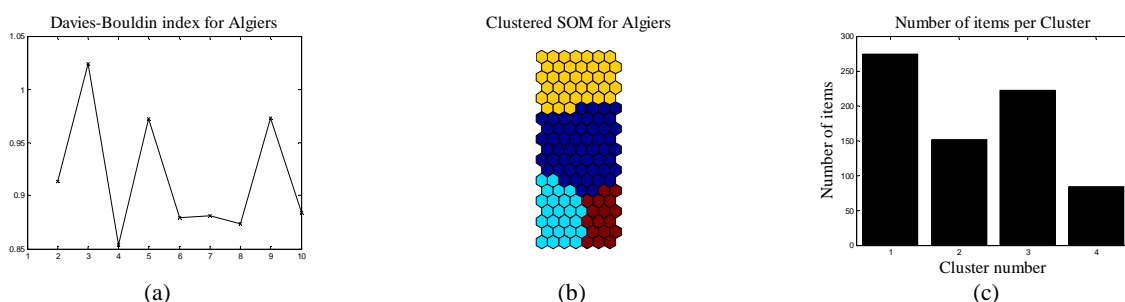


Figure 5.1 – Evolution de l'indice DB minimum 0.8540 atteint pour 4 classes (a), Segmentation de SOM avec k-means en 4 classes (b) et Nombre d'items par classe pour Alger (c).

L'évolution de l'indice de Davies-Bouldin en fonction du nombre de classes pour la ville de Hassi Messaoud est donnée en figure 5.2a. L'indice minimal 0.7929 est obtenu pour trois classes. La figure 5.2b montre la segmentation de la carte SOM en trois classes. La figure 5.2c est un histogramme indiquant le nombre d'items ou jours contenus dans chaque classe.

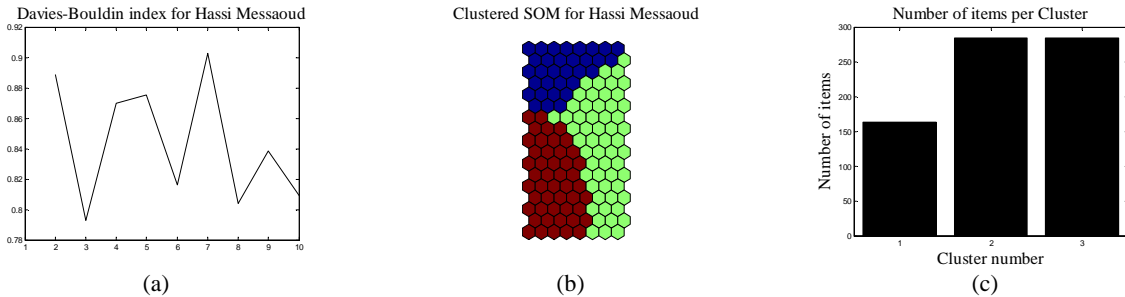


Figure 5.2 – Evolution de l'indice *DB* minimum 0.7929 atteint pour 3 classes (a), Segmentation de SOM avec k-means en 3 classes(b) et Nombre d'items par classe pour Hassi Messaoud (c). la ville d Alger.

Les tableaux 5.6 et 5.7 montrent respectivement les distributions hebdomadaires et mensuelles par classes pour Alger. Dans le tableau 5.6, la classe C3 contient un nombre important de Vendredi, ce qui laisse supposer que cette classe représente des weekends.

Jour	Sam	Dim	Lun	Mar	Mer	Jeu	Ven
C1	38	35	43	48	38	42	30
C2	25	24	22	24	25	21	10
C3	31	30	25	18	26	33	59
C4	10	15	14	14	16	9	6

Tableau 5.6 - Distribution hebdomadaire par classe pour Alger.

Mois	J	F	M	A	M	J	J	A	S	O	N	D
C1	21	24	33	32	18	18	5	7	22	19	36	39
C2	0	0	0	0	0	21	56	55	18	1	0	0
C3	4	9	23	27	44	21	1	0	17	41	24	11
C4	37	24	6	1	0	0	0	0	3	1	0	12

Tableau 5.7 - Distribution mensuelle par classe pour Alger

nombre important de jours en juin, juillet, aout et quelques jours de septembre donc des jours d'été.

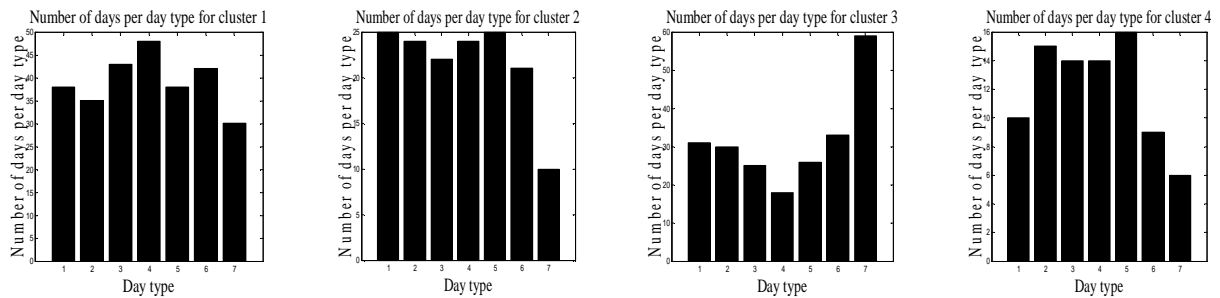


Figure 5.3 – Distribution hebdomadaire des classes de la ville d’Alger.

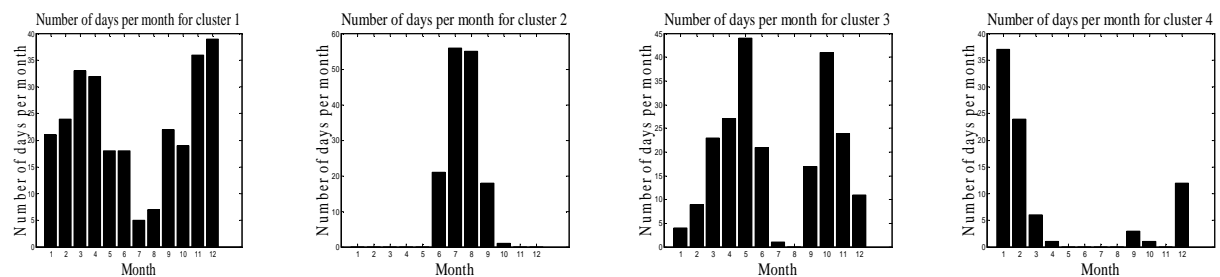


Figure 5.4 – Distribution mensuelle pour Alger.

Les tableaux 5.8 et 5.9 montrent respectivement les distributions hebdomadaires et mensuelles par classes pour Hassi Messoud. Dans le tableau 5.8, la classe C1 contient un nombre important de Vendredi, ce qui laisse supposer que cette classe représente des weekends pour la ville de Hassi Messaoud.

Jour	Sam	Dim	Lun	Mar	Mer	Jeu	Ven
C1	20	22	17	17	19	26	42
C2	48	47	39	40	44	37	29
C3	36	35	38	47	42	42	34

Tableau 5.8 - Distribution hebdomadaire nar classes pour Hassi Messaoud.

Mois	J	F	M	A	M	J	J	A	S	O	N	D
C1	1	7	32	37	29	14	5	2	5	23	7	1
C2	44	30	21	21	26	18	7	1	17	23	31	45
C3	17	20	9	2	7	28	50	59	38	16	22	16

Tableau 5.9 - Distribution mensuelle par classes pour Hassi Messaoud.

Les figures 5.5 et 5.6 montrent respectivement sous forme d’histogrammes la distribution hebdomadaire et mensuelle par classes pour Hassi Messaoud.

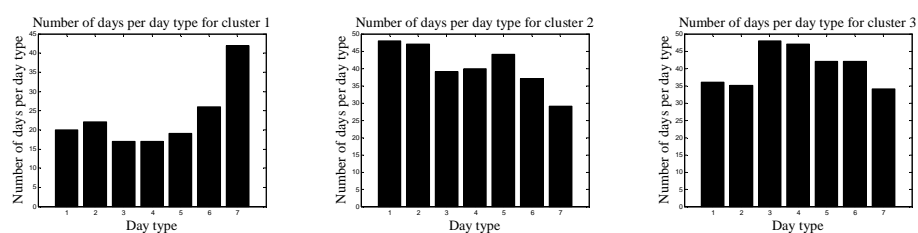


Figure 5.5 – Distribution hebdomadaire des classes de la ville de Hassi Messaoud.

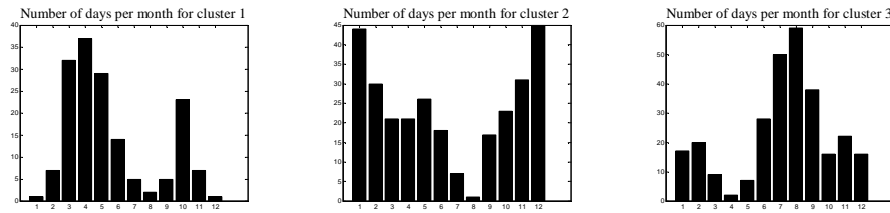


Figure 5.6 – Distribution mensuelle pour Hassi Messaoud.

5.3.2. Approche à deux niveaux avec la classification floue.

L'application directe des algorithmes de classification automatique sur les données nécessite un temps énorme d'exécution. Pour réduire ce temps, l'approche à deux phases a été proposée : elle consiste à appliquer en premier les cartes de Kohonen ce qui permet de générer les prototypes. Ensuite cette carte est segmentée avec un algorithme partitionné C-moyenne flou.

L'application de l'algorithme des fuzzy C-means est affecté par deux paramètres : le nombre de classes c et le degré de flou m . Les indices de validités seront utilisés pour déterminer le nombre optimal de classes c (Halkidi et al., 2002).

Les indices de validités retenues sont le coefficient de partition (PC) (Bezdek, 1981), l'entropie de partition ou classification (PE) (Bezdek, 1981), l'indice de partition (SC) (Bensaid, 1996) et l'indice de Xie et Beni (XB) (Xie et Beni, 1991) présentées dans la section 3.5.3.

L'analyse FCM est exécuté pour ces indices, le degré de flou m est fixé à 2, le maximum d'itérations à 100. Ces valeurs ont été choisies après plusieurs simulations. La précision est fixée à $1e-5$.

L'algorithme 5 rappelle les grandes lignes :

Algorithme 5 Approche à deux phases SOM et Fuzzy C-means

1. Déterminer la topologie de la carte SOM.
 2. Déterminer le neurone gagnant : celui qui a la plus petite distance.
 3. Mettre à jour les vecteurs poids.
 4. Répéter les étapes 2 à 4 jusqu'à stabilisation de l'auto-organisation.
 5. Initialiser le degré flou m et choix optimal du nombre de classes K avec les indices flous
 6. Segmentation des prototypes de la carte SOM avec l'algorithme C-moyenne flou
 7. Interprétation des résultats (Visualisation des classes, Matrices des centres de classes)
-

Nous avons mené des expériences à l'aide des indices de validité pour Alger et Hassi Messaoud. Le tableau 5.10 présente les résultats des valeurs optimales des quatre indices de validités pour c variant de 2 à 13. L'indice du coefficient de partition PC et l'indice de partition PE suggèrent respectivement $c = 2$ et $c = 3$ alors que l'indice de partition SC et l'indice de Xie et Beni XB donnent les mêmes valeurs 4 pour Alger et 3 pour Hassi Messaoud.

La courbe de l'indice de PC et PE en fonction du nombre de classes c est respectivement une courbe monotone croissante et décroissante. Ces indices tiennent compte uniquement des données. Les indices XB et SC tenant compte à la fois des données et des propriétés d'appartenance floue sont retenues (Xie et Beni, 1991).

Finalement, l'indice de XB qui évalue à la fois les propriétés d'appartenance floue et les données est retenu. La valeur minimale de XB indique le nombre optimal de classes. Ainsi, le nombre optimal de classes est 4 pour Alger et 3 pour Hassi Messaoud (Voir tableau 5.10)

DataSet	c	PC	PE	SC	XB
Alger	4	2	2	4	4
Hassi Messaoud	3	3	3	3	3

Tableau 5.10 – Valeurs de c préférées.

Les valeurs d'appartenance dans chaque classe indiquent la probabilité qu'un jour appartient à classé dans cette classe. Un échantillon de 741 jours et leurs valeurs d'appartenance pour les quatre classes de la ville d'Alger est donné dans la table 5.11. Pour un jour quelconque, la classe ayant la plus grande valeur d'appartenance est la classe représentative de ce jour. Par exemple, pour le premier jour correspondant au 1 janvier 2003, les valeurs d'appartenance des quatre classes sont : 0,0812, 0,0353, 0,2641 and **0,6192**. Notons que la somme de ces valeurs est proche de l'unité. La classe représentative de ce jour est donc la quatrième classe. De la même manière, tous les jours ont été traités et classés. Les tableaux 5.11 et 5.12 montrent respectivement en gras les valeurs d'appartenance des jours et leurs classes représentatives pour Alger et Hassi Messaoud.

Item number	C1	C2	C3	C4
1	0.0812	0.0353	0.2641	0.6192
2	0.1012	0.0388	0.3516	0.5083
3	0.0691	0.0341	0.2005	0.6960
4	0.2288	0.0535	0.5690	0.1485
5	0.7004	0.0675	0.1813	0.0505
.....				
.....				
727	0.3198	0.4956	0.1209	0.0635
728	0.4382	0.4055	0.1080	0.0481
729	0.5729	0.3079	0.0864	0.0326
730	0.4382	0.4055	0.1080	0.0481
731	0.6134	0.2451	0.1013	0.0401

Tableau 5.11 – Résultats de la classification floue pour les quatre classes d'Alger.

Item number	C1	C2	C3
1	0.0695	0.1092	0.8211
2	0.0412	0.0476	0.9110
3	0.2916	0.1350	0.5733
4	0.0319	0.0658	0.9022
5	0.1279	0.1038	0.7682
6	0.0379	0.0938	0.8681
.....			
.....			
726	0.0700	0.2241	0.7058
727	0.0522	0.8123	0.1353
728	0.0737	0.3281	0.5980
729	0.0438	0.1481	0.8079
730	0.0379	0.0938	0.8681
731	0.0773	0.7364	0.1861

Tableau 5.12 - Résultats de la classification floue des trois classes de Hassi Messaoud.

Les figures 5.7 et 5.8 montrent respectivement les résultats obtenus par l'analyse floue pour Alger et Hassi Messaoud. Les figures 5.7a et 5.8a montrent respectivement pour chaque ville l'évolution des valeurs d'appartenance en fonction des jours. L'axe horizontal indique le prototypes obtenus par SOM et l'axe vertical représente les valeurs d'appartenance correspondantes. La visualisation des classes est obtenue en utilisant l'analyse en composantes principales ACP (Section 2.2).

Les résultats sont présentées en 3D en figure 5.7b pour Alger et 5.8b pour Hassi Messaoud et 2D en figure 5.7c pour Alger et 5.8c pour Hassi Messaoud. X et Y sont respectivement la première et la seconde composante principale, l'axe Z étant les valeurs d'appartenance.

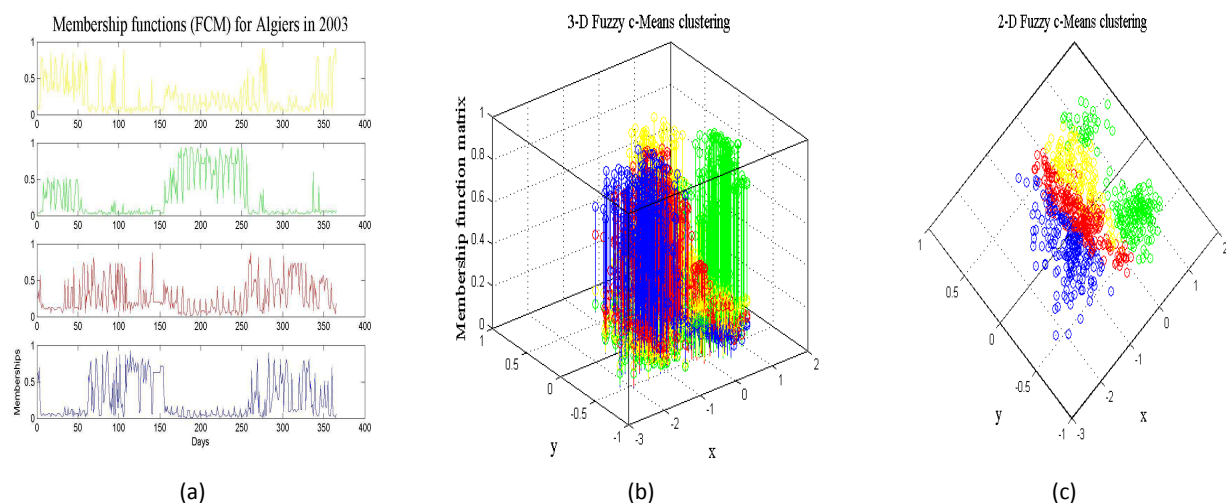


Figure. 5.7. Résultats graphiques de la classification floue pour Alger. (a) On distingue quatre classes de haut en bas (C1 à C4). L'axe horizontal de chaque sous-figure représente les prototypes obtenus par SOM et l'axe vertical de chaque sous-figure représente les valeurs d'appartenance. (b) et (c) donne la visualisation des classes en utilisant PCA respectivement en 3D et 2D. X et Y sont la première et seconde composante principale.

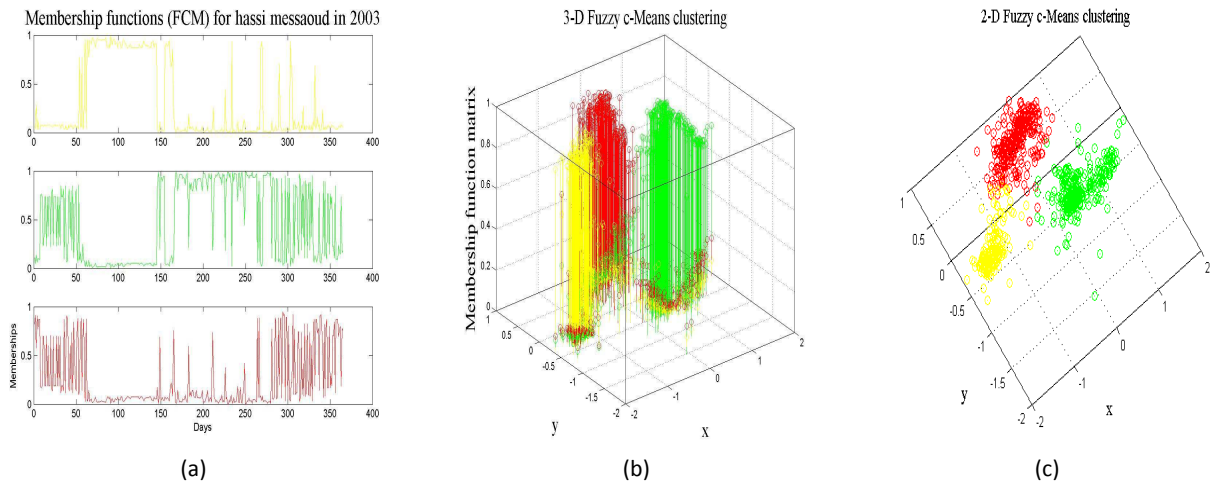


Figure 5.8. Résultats graphiques de la classification floue pour Hassi Messaoud. (a) On distingue Trois classes de haut en bas (C1 à C3). L’axe horizontal de chaque sous-figure représente les prototypes obtenus par SOM et l’axe vertical de chaque sous-figure représente les valeurs d’appartenances. (b) et (c) donne la visualisation des classes en utilisant PCA respectivement en 3D et 2D. X et Y sont la première et seconde composante principale.

Le jour qui a la valeur d’appartenance la plus élevée dans un groupe est le jour représentatif de ce groupe. Les jours représentatifs de chaque classe d’Alger et de Hassi Messaoud sont décrits respectivement dans les Tableaux 5.13 et 5.14. Les Tableaux 5.15 et 5.16 montrent respectivement les classes et périodes associées avec chaque classe coloré et visualisé dans les figures 5.7b et 5.8b.

Les tableaux 5.17 et 5.18 indiquent en gras pour chaque région et pour chaque type de jour le degré d’appartenance par rapport aux classes. Par exemple, pour la région d’Alger le jour correspondant au 01 juillet 2003 appartient avec 90,55 % à la classe C2 (Eté), 5,55 % à la classe C1 (Hiver), 2,6 % à la classe C3 (Automne) et 1,73% à la classe C4 (Printemps). Le jour du 31 janvier 2003, un jour d’hiver appartient à 84,14% à la classe C1 (Hiver), 8,87% à la classe C2 (Eté), 5,26% à la classe C3 (Automne) et 1,73% à la classe C4 (Printemps).

De la même manière, pour la région de Hassi Messaoud, le jour correspondant au 31 janvier 2003 appartient avec 69,27 % à la classe C3 (Printemps), 23,73 % à la classe C2 (Eté), 0,07 % à la classe C1 (Automne). Le jour du 20 décembre 2003 appartient à 69,69 % à la classe C3 (Printemps), 23,14% à la classe C2 (Eté), 0,07 % à la classe C1 (Automne).

Jour représentatif	30 /12/2003	01/07/ 2003	21/10/2003	27/03/2003
Degré d’appartenance	0.9099	0.9055	0.5202	0.926

Tableau 5.13 – Jour représentatif et degré d’appartenance par classe pour Alger.

Jour représentatif	01/04/1 2003	27/05/ 2003	21/10/2003
Degré d’appartenance	0.9899	0.9881	0.9462

Tableau 5.14 - Jour représentatif et degré d’appartenance par classe pour Hassi Messaoud

Saison	Classes	Overlapping	Période
Hiver	C1 (Jaune)	C2,C3,C4	Dec, Jan, Fev
Eté	C2 (Vert)	C1, C3, C4	Juillet , Aout
Automne	C3 (Rouge)	C1, C2, C4	Sept, Oct, Nov
Printemps	C4 (Bleue)	C1, C2, C3	Mars, Avril, Mai, Juin

Tableau 5.15 - Classe identifiée pour Alger.

Saison	Classes	Overlapping	Période
Automne	C1 (Jaune)	C2, C3	Feb, Mars, Avril
Eté	C2 (Vert)	C1, C3	Mai, Juin, Juillet, Aout
Printemps	C3 (rouge)	C1, C2	Sept, Oct, Nov, Dec, Jan

Tableau 5.16 - Classe identifiée pour Hassi Messaoud.

Day	C1	C2	C3	C4
01/07/2003	0.0555	0.9055	0.026	0.013
31/01/2003	0.8414	0.0887	0.0526	0.0173
22/01/2003	0.7928	0.0799	0.0953	0.032
21/10/2003	0.1645	0.056	0.5202	0.2594
10/08/2003	0.208	0.6406	0.1072	0.0442
10/05/2003	0.102	0.0497	0.3682	0.4802

Tableau 5.17 – Valeurs d'appartenance par classe et par jour pour Alger.

Day	C1	C2	C3
31/01/2003	0.07	0.2373	0.6927
20/12/2003	0.0717	0.2314	0.6969

Tableau 5.18 – Valeurs d'appartenance par classe et par jour pour Hassi Messaoud.

5.4. Conclusion et contributions

A l'issue des approches proposées, nous dressons le bilan suivant :

Au premier abord, l'identification des types de jours de la charge électrique régionale Algérienne utilisant les cartes auto-organisée de Kohonen SOM est une approche qualitative permettant une bonne visualisation des classes moyennant un bon choix des paramètres de SOM. Son principal inconvénient est l'absence d'informations quantitatives.

La segmentation des cartes de Kohonen avec l'algorithme des K-Means en utilisant l'indice de validité de Davies-Bouldin DB permet une meilleure définition des frontières des classes et permet une analyse quantitative très riche. Cette approche est utile pour des classes

hypersphériques, compactes et séparées. Seulement, cette approche ne reflète pas le contexte réel de la charge électrique algérienne.

L'approche SOM combiné avec l'approche floue avec l'indice de validité flou XB fournit des informations plus détaillées et s'adapte bien au contexte de la charge électrique régionale algérienne. SOM combiné à FCM est particulièrement utile quand les bordures des classes ne sont pas clairement séparées et ambiguës. Cette classification pourra être utilisée pour la prédiction de la charge électrique afin de réduire l'espace des entrées utilisé pour chaque classe et de ce fait, implémenter des modèles de prédictions respectifs à ces dernières.

Le chapitre suivant illustre l'apport de l'identification de types de jours pour la prédiction de la charge électrique Algérienne.

6

La prédiction à court terme de la charge électrique Algérienne

Sommaire

6.1 Introduction	68
6.2 Prévion de la charge électrique	68
6.2.1 Types de prévision de charge	68
6.2.2 Facteurs influençant la charge électrique	69
6.2.3 Les méthodes de prévision de charge	70
6.2.3.1 Méthode de régression	70
6.2.3.2 Système expert	70
6.2.3.3 Méthode basée sur les séries chroniques stochastiques	71
6.2.3.4 Les techniques non-linéaires de la charge électrique	73
6.3 Méthode basée sur les systèmes experts flous	73
6.3.1 Système expert flou	73
6.3.2 Moteurs d'inférences pour système flou	75
6.3.3 Génération automatique de la base des règles	77
6.4 Méthode basée sur les réseaux de neurones	78
6.4.1 Méthode des réseaux de neurones, pourquoi ?	78
6.4.2 Théorie des réseaux de neurones artificiels	79
6.4.2.1 Modèle des réseaux de neurones (RNA)	79
6.4.2.2 Apprentissage	82
6.4.2.3 Réseau rétro-propagation	83
6.5 Cas d'application : Approche de prédiction par système expert flou	85
6.5.1 Méthodologie de conception du système	85
6.5.1 Préparation des données initiales	86
6.5.2 Choix de la base	86
6.5.3 Architecture du système projeté	86
6.5.4 Apprentissage	87
6.5.5 Validation	88
6.5.6 Test	88
6.5.7 Résultats de calcul	88
6.6 Conclusion	95

6.1 Introduction

La prévision des consommations joue un rôle très important dans la détermination des stratégies d'exploitation optimale, de planification de réseaux électriques. Dans le contexte de la libération du marché de l'énergie, la prévision de la consommation joue aussi un rôle essentiel dans la gestion optimale des consommations, suivant différents objectifs (minimiser la facture, réduire les pics de consommation, contribuer aux services systèmes) (Haykin, 1999).

Dans ce chapitre, nous étudions les éléments de base de la prévision de la charge électrique. Nous décrirons les différents types de prévision de charge, les facteurs influençant la charge et les différentes méthodes utilisées dans le contexte linéaire et non-linéaire de la prédiction à court terme et surtout l'utilisation des types de jours et classes obtenus dans une stratégie de prédiction.

Dans ce contexte, nous présenterons en fin de ce chapitre une étude de cas de la prédiction de la charge électrique Algérienne. Cet exemple de prévision utilise un système expert basé sur la logique floue, avec génération automatique de règles d'inférence. Cette approche tient compte de la notion de classification élaborée aux chapitres 4 et 5.

6.2 Prévision de la charge électrique

6.2.1 Types de prévision de charge

Afin de d'établir un modèle de prévision de charge, il est nécessaire d'identifier les besoins auxquels on souhaite répondre. Selon l'horizon de prévision on distingue les types de prévision :

- La prévision à long terme :

L'horizon de prévision est compris dans une période entre un an et dix ans. Ce temps est nécessaire pour planifier, construire des centrales de transport et de distribution de l'électricité.

- La prévision à moyen terme :

L'horizon de prévision à moyen terme est une période entre un mois et une année. Ce type de prévision est utilisé pour déterminer les équipements des centrales et les ouvrages des réseaux qui faut installer ou pour établir les contrats du marché électrique.

- La prévision à court terme

La prévision à court terme peut être divisée en prévision à très court terme et prévision à court terme. La première comprend une période de 30 minutes en temps réel à partir du moment où on réalise la prévision de la consommation. Le but de ce type de prévision est de couvrir la distribution, de la manière la plus économique possible. On effectue un suivi de la charge à partir des prévisions faites précédemment dans un passé proche (24 heures avant).

La prévision à court terme est prévue entre une heure et une semaine dans le futur; durant cette période, la compagnie d'électricité doit connaître le plan d'exploitation des centrales, le

plan de maintenance ou les stratégies de gestion des charges pour minimiser les coûts. Les meilleures sources d'information pour la prévision sont l'historique des consommations et de la température des années précédentes.

Dans notre étude, nous nous intéressons à la prévision de la charge à court terme. C'est une fonction essentielle du système de gestion d'énergie pour le dispatching, la planification, l'analyse du système de contrôle des réseaux électriques. La prévision pour le jour ou les jours suivants est nécessaire pour demander d'adapter le plan de production ou les stratégies de gestion des charges afin de minimiser le coût d'exploitation en respectant les contraintes de fonctionnement et de fiabilité des réseaux électrique.

6.2.2 Facteurs influençant la charge électrique

La prévision de charge électrique dépend de différents facteurs qu'il va falloir identifier. Ces facteurs peuvent être :

Les facteurs économiques

Les spécificités économiques d'un secteur peuvent affecter le profil d'utilisation de la charge électrique. Ceci concerne le type de client, les conditions démographiques ou les activités industrielles.

Les facteurs temporels

Les facteurs temporels incluent l'aspect saisonnier, hebdomadaire, et les jours de vacances. Les exemples pour l'effet saisonnier incluent l'effet des heures de jour dans une saison, qui affecte le modèle de charge. La charge industrielle des jours de la semaine sera plus importante que celle des jours de week-end. Les vacances également ont un effet sur le modèle de charge.

Les effets climatiques

La température est le facteur climatique qui influence le plus la prévision de la charge électrique. Ainsi, une variation de ce facteur pourrait affecter la quantité de puissance requise pour le chauffage en hiver et la climatisation en été. L'humidité particulièrement dans les aires chaudes et humides, la précipitation, les orages, et l'intensité de vent et de lumière du jour peuvent également influencer la charge électrique.

Des perturbations aléatoires

les aciéries, grands consommateurs d'énergie électrique peuvent causer des variations brusques de la charge. En outre, certains événements, comme les émissions de télé à grande audience populaires influencent la charge électrique.

Facteurs des prix

Un profil d'utilisation de la charge varie d'une région à une autre. Par exemple, le profil d'utilisation pour des secteurs ruraux est différent de celui des zones urbaines. Le profil

d'utilisation peut également dépendre du type de consommateur. Par exemple, le profil d'utilisation résidentiel peut être différent des consommateurs commerciaux et industriels.

6.2.3 Les méthodes de prévision de charge

Dans la bibliographie, on peut trouver un large éventail de méthodes de prévision de la consommation électrique. La classification est basée sur certaines caractéristiques, comme par exemple le type de modèle de charge, le type de données à fournir au modèle, le temps requis de calcul informatique, l'algorithme de prévision ainsi que la disponibilité de résultats expérimentaux.

Certains auteurs proposent une typologie selon le type de modèle de charge utilisé. La répartition est essentiellement fondée sur deux modèles de base : le modèle de charge maximale et la forme de la courbe de charge. D'autres auteurs (Dash et al., 1997) regroupent les méthodes selon les modèles en séries chronologiques et les modèles de régression.

6.2.3.1 Méthode de régression

Il s'agit de l'une des techniques statistiques qui reflètent le rapport entre la consommation électrique et d'autres facteurs tels que le temps, le jour et le type de consommation. La méthode de régression s'efforce de déterminer les valeurs prises par une variable en tenant compte de l'influence des autres variables. On dira alors qu'il y a une variable dépendante et plusieurs variables indépendantes entre lesquelles on recherche une corrélation. L'objectif de cette analyse est généralement de déterminer quelles sont les variables indépendantes qui ont le plus d'influence sur la variable dépendante (Hann et al., 2009).

6.2.3.2 Système expert

Cette méthode incorpore des règles et des procédures employées par les experts. Les systèmes experts sont des modèles heuristiques, qui peuvent habituellement tenir compte des facteurs quantitatifs et qualitatifs (Dash et al., 1997).

On a proposé beaucoup de variantes de ce type depuis les années 80. Une approche typique est d'essayer d'imiter le raisonnement d'un opérateur humain. Une méthode possible pour qu'un expert humain crée la prévision est de rechercher dans la base de données historique le jour qui correspond au mieux au jour cible en tenant compte des facteurs qui le caractérisent. Les valeurs de la charge enregistrées correspondant à ce jour sont alors prises comme base de la prévision. Un système expert peut être une version automatisée de ce genre de processus de recherche.

D'autre part, le système expert peut affiner ses résultats en tenant compte plus explicitement des facteurs externes et des profils d'utilisation quotidiens. Un exemple sera donné pour la prédiction de la charge électrique dans ce qui suit.

6.2.3.3 Méthode basée sur les séries chronologiques stochastiques

Box et Jenkins (Box et Jenkins, 1970) développa les bases de l'analyse par séries chronologiques. Ces techniques ont été utilisées dans plusieurs applications de prédiction de la charge électrique (Di Caprio et. al., 1985 ; Rajukar et Newill, 1985 ; Ho et. al.,1990 ; Murray, 1996 ; Infield et Hill, 1998 ; Charytoniuk et. al.,1999).

Box et Jenkins modélise une série chronique comme une fonction des valeurs antérieures, erreurs de rendement du passé et entrées externes. Cependant, la série chronique à prédire doit être stationnaire (Box et Jenkins, 1970).

La forme générale du modèle de Box-Jenkins :

$$A(q)x(k) = \frac{B(q)}{F(q)} u(k - \tau) + \frac{C(q)}{D(q)} \varepsilon(k) \quad (6.1)$$

Où $x(k)$ est a série chronologique stationnaire à modeler au temps k

$u(k)$ est une entrée externe avec le retard entre l'entrée et $x(k)$

$\varepsilon(k)$ est l'erreur dans le modèle au temps k

A, B, C, D, F sont des polynômes en fonction du retard q^{-1} , A peut prendre la forme suivante :

$$A(q)x(k) = (1 - a_1q^{-1} - a_2q^{-2} \dots \dots - a_{na}q^{-na} x(k) \quad (6.2)$$

Où a_1, a_2, \dots, a_{na} sont es coefficients du polynôme et na est l'ordre. B, C, D et F peuvent être développés de la même façon avec respectivement les coefficients b_1, b_2, \dots, b_{nb} , c_1, c_2, \dots, c_{nc} , d_1, d_2, \dots, d_{nd} et les ordres respectifs sont nb, nc, nd . Les autres modèles proposés par Box et Jenkins (1970) peuvent être dérivés directement de l'équation 6.1 en remplaçant F et D par 1 comme indiqué sur le tableau 6.1.

Nom du modèle	A	B	C
AR	$1 - a_1q^{-1} \dots \dots \dots a_{na}q^{-na}$	0	1
ARX	$1 - a_1q^{-1} \dots \dots \dots a_{na}q^{-na}$	$b_1 + b_2q^{-1} \dots \dots \dots + b_{nb}q^{-nb}$	1
MA	1	0	$1 + c_1q^{-1} \dots \dots \dots c_{nc}q^{-nc}$
MAX	1	$b_1 + b_2q^{-1} \dots \dots \dots + b_{nb}q^{-nb}$	$1 + c_1q^{-1} \dots \dots \dots c_{nc}q^{-nc}$
$ARMA$	$1 - a_1q^{-1} \dots \dots \dots a_{na}q^{-na}$	0	$1 + c_1q^{-1} \dots \dots \dots c_{nc}q^{-nc}$
$ARMAX$	$1 - a_1q^{-1} \dots \dots \dots a_{na}q^{-na}$	$b_1 + 2q^{-1} \dots \dots \dots + b_{nb}q^{-nb}$	$1 + c_1q^{-1} \dots \dots \dots c_{nc}q^{-nc}$

Tableau 6.1 Les différents modèles de Box-Jenkins et polynômes associés ($F = 1$, $D = 1$)

De plus si la série stationnaire à prédire a été produite par différenciation de la série originale, le modèle est appelé modèle intégré et la lettre I est ajouté au nom du modèle. Par exemple, le modèle $ARMAX$ devient $ARIMAX$. L'équation 6.1 devient (Bowerman O'Connell, 1987) :

$$A(q)\nabla_s^D \nabla^d (x(k)) = B(q)u(k - \tau) + C(q)\varepsilon(k) \quad (6.3)$$

Box et Jenkins a aussi examiné le cas ou la série à prédire présente une saisonnalité modélise une série chronique comme une fonction des valeurs antérieures, erreurs de rendement du

passé et entrées externes. Cependant, la série chronique à prédire doit être stationnaire (Box et Jenkins, 1970).

$$A_s(q)A_{ns}(q)\nabla_s^D\nabla^d(x(k)) = B(q)u(k - \tau) + C_s(q)C_{ns}(q)\varepsilon(k) \quad (6.4)$$

Où A_s et A_{ns} est l'opérateur saisonnier AR et C_s et C_{ns} est l'opérateur saisonnier MA.

Le modèle décrit par l'équation 6.4 est appelé *SARIMAX* (Seasonal Auto Regressive Integrated Moving Average eXogenous).

L'approche de Box et Jenkins procède en quatre étapes :

- Transformer les données pour rendre la série stationnaire
- Etablir une hypothèse initiale concernant les ordres du polynôme.
- Estimer les paramètres du modèle
- Diagnostique de test du modèle

Box et Jenkins (1970) suggèrent une transformation de la série chronique en série stationnaires. De nombreuses techniques ont été élaborées mais la technique à utiliser dépend de la série à transformer.

Citons parmi ces techniques : la différenciation utilisée par Box et Jenkins (voir chapitre 2) qui utilise la notion d'opérateur de retard et l'opérateur de différence. La transformé de Box-Cox (Box and Cox, 1964), élimination d'une tendance linéaire (Ljung, 1987), élimination d'une tendance polynomiale (Brockwell et Davis, 1987).

Après avoir rendu la série stationnaire, on identifie les termes AR et MA nécessaires pour corriger les auto-corrélations partielles. Cette analyse est basée sur l'examen des fonctions d'auto-corrélation simple et auto-corrélation partielle (voir chapitre 2). On estime par la suite

La prochaine étape consiste à faire l'estimation des paramètres des modèles (estimation des coefficients) en utilisant les techniques présentées au chapitre 2 : maximum likelihood estimation (Harvey, 1981) et least squares estimation (Harvey, 1981). La dernière phase consiste à diagnostiquer et tester le modèle.

La méthode Box et Jenkins a été utilisée dans de nombreux travaux scientifiques pour la prédiction à court terme de la charge électrique (Kodogiannis et Agnostakis, 1999 ; Mbamalu and El-awary, 1993 ; Moghram et Rahman, 1989 ; Barakat et. al., 1990 ; Elkateb et. al., 1998).

Cependant, comme mentionné par (Mohamad et. al., 1996), la technique de Box et Jenkins présente l'inconvénient d'utiliser une grande base de donnée pour l'apprentissage et le procédé de différenciation peut générer des erreurs.

Rajurka et Newill (1985) utilisèrent un modèle *ARMAX* dans lequel les coefficients du modèles sont variables. Darbellay et Slama (2000) comparèrent un modèle *ARIMAX* avec un réseau de neurone. La donnée température est une entrée exogène pour les deux modèles. Les résultats obtenus par le modèle *ARIMAX* sont meilleurs, notons que la série chronique à prédire dans ce cas est la charge journalière totale et non pas la charge horaire.

6.2.3.4 Les techniques non-linéaires de la charge électrique

Un système est dit linéaire s'il satisfait le principe de superposition (Sinha, 1991). La superposition est composée de deux idées : l'additivité et l'homogénéité (Sinha, 1991). L'additivité signifie que la sortie du système linéaire est la somme de deux entrées est égale à la somme des sorties pris chacun individuellement.

$$f(u_1 + u_2) = f(u_1) + f(u_2) \quad (6.5)$$

Où f est une fonction linéaire et u_1 et u_2 sont les entrées.

L'homogénéité est définie par la relation suivante :

$$f(cu_1) = cf(u_1) \quad (6.6)$$

Où c est une constante.

Une simple exception est l'équation de la droite affine.

$$f(u(t)) = au(t) + c \quad (6.7)$$

Où c et a sont des constantes.

Un système non-linéaire où le principe de superposition ne s'applique pas. Il y'a deux facteurs présent dans la prédiction à court terme de la charge électrique :

- La relation no-linéaire entre charge électrique et la température
- La non-stationnarité de la charge due à la tendance.

On peut citer parmi les techniques les plus populaires, l'approche système expert flou et les réseaux de neurones.

6.3 Méthode basée sur les systèmes experts flous

6.3.1 Système expert flou

Le terme flou (fuzzy on anglais) a été introduit par Zadeh (1965), où une nouvelle discipline mathématique, la logique flou basé sur la théorie des ensembles flous a été créée. Le but de cette théorie est de tolérer les concepts imprécis et inexacts en les représentants par des ensembles flous.

La théorie des ensembles flous (Voir chapitre 2) autorise le concept de **l'appartenance partielle** d'un objet ou une variable à un ensemble flou, donc, permet une transition graduelle de l'appartenance complète ou non appartenance. Ainsi, dans la logique floue, un objet ou une variable dans un domaine peut partiellement appartenir à plusieurs sous ensemble flou du même domaine simultanément.

En utilisant une fonction d'appartenance μ , on définit un ensemble flou F sur un domaine U comme suit :

$$\mu_F(x) : U \rightarrow [0,1] \quad (6.8)$$

Soit C un ensemble strict et x un élément de C tel que $x \in C$. Le degré d'appartenance de x à C est égale à 1 si x appartient totalement à C et à 0 s'il n'appartient pas à C .

$$\mu_C(x) = \begin{cases} 1 & \text{si } x \in C \\ 0 & \text{si } x \notin C \end{cases} \quad (6.9)$$

Si l'ensemble F est flou, Le degré d'appartenance de x à C est égale à 1 si x appartient totalement à C et à 0 s'il n'appartient pas à C . Cependant, si x appartient partiellement à F alors le degré d'appartenance prend des valeurs entre 0 et 1.

$$\mu_F(x) = \begin{cases} 1 & \text{si } x \in C \text{ (complètement)} \\ (0,1) & \text{si } x \in C \text{ (partiellement)} \\ 0 & \text{si } x \notin C \text{ (non membre)} \end{cases} \quad (6.10)$$

Pour expliquer l'importance des ensembles flous et la fonction d'appartenance, considérant la loi de Boyle, par exemple. Il dit que la pression (P) d'une masse de gaz donnée a une relation inverse avec le volume (V) du gaz, si la température (T) est constante.

En utilisant les règles de la logique floue, la loi de Boyle peut être déterminée comme suit :

Règle-1 : **SI** pression est **Elevée** ET température est **Constante** **Alors** volume est **Petit**

Règle-2 : **SI** pression est **Moyen** ET température est **Constante** **Alors** volume est **Moyen**

Règle-3 : **SI** pression est **Faible** ET température est **Constante** **Alors** volume est **Elevée**

Ces règles peuvent être résumées par la relation :

$$PV = C \quad (6.11)$$

Les trois règles IF-THEN sont suffisantes pour modéliser les observations de Boyle. La pression, la température et le volume sont les variables et les termes Petit, Moyen et Elevée sont les étiquettes. Ainsi les règles de la logique floue (IF-THEN) décrivent les expériences ou observations d'un système quelconque dans la modélisation de ce système avec les ensembles flou.

Les systèmes experts flous ont une relation direct avec les concepts flous, tel que les ensembles flous, variables linguistiques, et la logique floue. Les systèmes flous sont les seuls ayant la possibilité de traiter en plus les données numériques, et les connaissances linguistiques. D'un point de vue mathématique, un système expert flou est un mappage non linéaire d'un vecteur de caractéristiques d'entré (donnée) en une sortie type scalaire (Ajoy et Popvitch, 2005).

Le diagramme d'un système flou est montré dans la figure 6.1. Le système a comme entrée une valeur **précise** (X), cette dernière est **fuzzifiée** (transformée en degré d'appartenance dans l'ensemble flou d'entrée), en suite, elle est transmise au moteur d'inférence flou. En utilisant les règles floues SI-ALORS enregistrées dans la base des règles, le moteur d'inférence produit une valeur floue qui va être **défuzzifiée** donnant le résultat (Y) pour être utilisable.

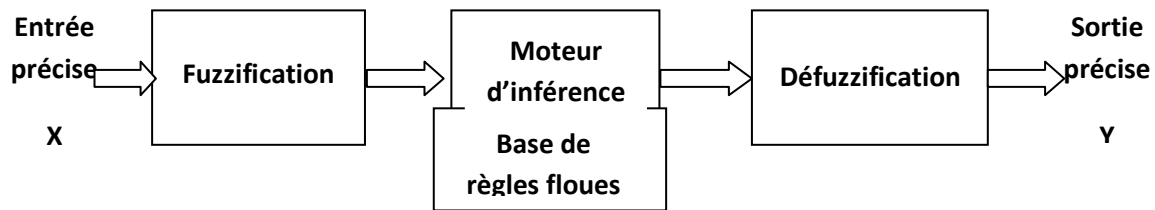


Figure - 6.1 Diagramme d'un système expert basé sur la logique floue

Les ensembles flous peuvent être utilisés dans un système logique flou de plusieurs manières :

- Dans la description du système.
- Dans la spécification des paramètres du système.
- Dans la représentation des entrées, sorties et états du système.

6.3.2 Moteurs d'inférences pour système flou :

Un **moteur d'inférence** pour un système donné est le noyau de ce système flou. En pratique, les deux moteurs d'inférence suivants sont les plus utilisés :

- Le moteur d'inférence Mamdani (Mamdani, 1977)
- Le moteur d'inférence Takagi-Sugeno (Kosko, 1997)

La principale différence réside dans la conséquence des règles floues. Le système Mamdani utilise les ensembles flous comme règles floues conséquentes alors que le moteur d'inférence Takagi-Sugeno (*TS*) emploie des fonctions linéaires pour les variables d'entrées comme règle conséquente. Nous présenterons dans ce qui suit la méthode de Mamdani qui a les avantages suivants :

- Efficacité dans les calculs.
- Efficace pour les techniques linéaires, d'optimisation et adaptatives.
- Elle s'adapte bien pour l'analyse mathématique.

Mamdani (1977) a proposé le premier moteur d'inférence pour objectif de contrôler la combinaison, locomotive à vapeur, et la casserole, basé sur un ensemble de règles linguistiques de contrôle construite comme une connaissance d'un expert humain.

Appliqué sur la loi de Boyle, les règles linguistiques suivantes peuvent être écrites comme suit :

Règle-1 : SI la pression est ELEVÉE et la température est CONSTANTE alors le volume est FAIBLE

Règle-2 : SI la pression est MOYENNE et la température est CONSTANTE alors le volume est MOYEN

Règle-3 : SI la pression est FAIBLE et la température est CONSTANTE alors le volume est ÉLEVÉ

Ces règles sont dites règles **floues de type Mamdani** (premières introduites et utilisées par Mamdani). Les ensembles flous sont utilisés pour décrire les variables d'entrées et de sorties

du système. L'essentielle caractéristique de ces règles c'est que les deux parties d'une règle de ce type ont une nature floue (imprécise). Inférer revient au processus de génération de l'ensemble flou de sortie quand l'ensemble des entrées et l'ensemble de règles sont disponibles. les règles floues de type Mamdani ont souvent une sortie floue qui n'est pas compatible avec le signal réel, quand une valeur précise est demandée. L'ensemble flou de sortie doit être défuzzifié, sinon il ne sera pas compatible avec le signal de sortie. La défuzzification est le processus de transformation d'un ensemble flou à une valeur numérique. On peut utiliser à cet effet la méthode du **centre de gravité** noté CDG :

Soit un ensemble flou F représenté comme suit :

$$F = \{\mu_F(x_1)/x_1, \mu_F(x_2)/x_2, \dots, \mu_F(x_p)/x_p\} \quad (6.12)$$

Le CDG permet de calculer les coordonnées de x , centre de gravité de l'ensemble flou F comme suit :

$$x' = \left(\sum_{i=1}^p \mu_F(x_i) \cdot x_i \right) / \left(\sum_{i=1}^p \mu_F(x_i) \right) \quad (6.13)$$

Inférer un modèle flou de type-Mamdani comprend quatre étapes de base :

- Etape 1 : calculer le degré d'accomplissement de chaque règle par

$$\beta^l = \mu_{G_1^1}(x_1) \wedge \mu_{G_2^1}(x_2) \wedge \dots \wedge \mu_{G_n^1}(x_n), \quad (6.14)$$

Où \wedge est l'opérateur min

- Etape 2 : calculer l'ensemble de conséquences de chaque règle par

$$F_1^l = \beta^l \wedge F_1^l \quad (6.15)$$

- Etape 3 : agréger tout les conséquences des ensemble flous par

$$F_{aggr}' = \bigcup_{l=1}^M F_1^l = (F_1^1 \cup F_1^2 \cup \dots \cup F_1^M) \quad (6.16)$$

- Etape 4 : défuzzifier l'ensemble agrégé F_{aggr}' en utilisant la méthode CDG.

6.3.3 Génération automatique de la base des règles

Le moteur d'inférence requiert une base de règles complètes et cohérentes. Les règles floues de type (Mamdani, 1977) sont générées par l'algorithme suivant :

Soit la série temporelle $X = \{X_1, X_2, X_3, \dots, X_q\}$ pour $t = \{1, 2, 3, \dots, q\}$, les règles floues de type Mamdani sont générées comme suit :

- **Etape 1.** Partition des données de la série temporelle en forme XIO
 $XI(t) = [X\{t - (D - 1)d\}, X\{t - (D - 2)d\}, \dots, X\{t - d\}, X\{t\}]$,
 $XO(t) = [X(t + L)]$,
- **Etape 2.** Division de l'intervalle du domaine $[X_{min}, X_{max}]$ en $(n - 1) = 2N$ régions floues en chevauchement.

$$X_{min} = \text{Min}(X), X_{max} = \text{Max}(X).$$

Assigner à chaque région une Fonction d'appartenance gaussienne FAG, notée G_1, G_2, \dots, G_n .

- **Etape 3.** Calculer le segment $S = (X_{max} - X_{min})/n - 1$ afin que la moyenne des FAGs soit :
 $c_1 = X_{min}, c_2 = c_1 + S, \dots, c_r = c_{r-1} + S = c_1 + (r - 1)S, \dots, c_n = c_1 + (n - 1)S = X_{max}$
 Et leurs variances :

$$\text{Sigma}_{G1} = \text{Sigma}_{Gn} = \text{Sigma}_a,$$

$$\text{Sigma}_{G2} = \text{Sigma}_{G3} = \dots = \text{Sigma}_{Gn-1} = \text{Sigma}_b.$$

Sélectionner deux valeurs convenable pour Sigma_a et Sigma_b , afin que deux régions floues adjacentes soient partiellement chevauchées.

- **Etape 4.** Fuzzifier tous les valeurs d'entrées et de sortie
 Pour chaque entrée X_{ki} ou sortie Y_k calculer le degré d'appartenance dans chaque région Gaussienne.

$$0 < \mu_{G_j}(X_{ki}) = f(X_{ki}) = \exp\left(-0.5 \cdot \frac{(X_{ki} - c_j)^2}{(\text{sigma}_{G_j})^2}\right) < 1.$$

Pour $i = 1, k = 1, \text{ et } j = 1, 2, 3, \dots, n$

- **Etape 5.** Arranger tous les degrés d'appartenance pour chaque variable X_{ki} , ou Y_k ($k = 1, 2, \dots, M$, et $i = 1, 2, 3, 4, \dots, \text{ etc.}$) dans un vecteur de dimension $(n \times 1)$
 Trier tous les vecteurs côté à côté séquentiellement pour obtenir la Mu-matrice de taille $\{n \times (i_{max} + 1)\}$.
 Pour un système à 4 entrées et une sortie la matrice sera de taille $(n \times 5)$.
- **Etape 6.** Sélectionner la valeur maximale du degré d'appartenance dans chaque colonne et sauvegarder l'indice r de la ligne en question comme suit

$$0 < \mu_{G_r}(X_{ki}) = \max\left(\mu_{G_j}(X_{ki})\right) \leq 1, j = 1, 2, 3, \dots, n \text{ et } 1 \leq r \leq n.$$

Créer la base des règles floues à partir des valeurs de r ainsi leur degré d'appartenance correspondant.

- **Etape 7.** Créez la liste des règles et résolvez les problèmes de conflits (s'il y en a) en utilisant le degré de la règle. De même, supprimez les règles redondantes de la liste des règles (s'il y en a)

Où :

D : c'est le nombre d'entrées.

d : le nombre d'unités de temps par échantillon.

L : Nombre de pas en avant.

n : nombre d'ensembles flous.

σ_a, σ_b , sont des paramètres de variance utilisés par la fonction de fuzzification $(f(X_{ki}) = \exp(-0,5 \cdot \frac{(X_{ki}-c_j)^2}{(\sigma_{G_j})^2}))$, et comme il n'y a pas une théorie pour les déterminer, il faut faire plusieurs itérations de test pour les optimiser.

6.4 Méthode basée sur les réseaux de neurones

6.4.1 Méthode des réseaux de neurones, pourquoi ?

Les méthodes statistiques classiques de prévision de charge présentent l'avantage de leur simplicité. Elles peuvent, en outre, tenir compte de l'influence des différentes conditions météorologiques. Cependant, ces méthodes présentent quelques inconvénients :

- Il est difficile de tenir compte de la non-linéarité entre les entrées et les sorties ainsi que des facteurs aléatoires
- Elles manquent de flexibilité pour les variations de charge rapides ...
- La prévision, en particulier des charges résidentielles est très compliquée. Le choix de la méthode des réseaux de neurones peut être utilisé lorsque :
- Il n'existe pas de modèle mathématique précis de la charge
- La charge est une fonction des différents facteurs (passé, calendrier, météo, autres...)
- La fonction de prévision est complexe et inconnue, et la relation non linéaire
- Les méthodes conventionnelles classiques comme la méthode de régression ou d'interpolation, peuvent ne pas fournir dans ce cas une précision suffisante. En plus, ces méthodes complexes avec des moyens de calcul très lourd peuvent converger lentement et peuvent même diverger dans certains cas. Elles ne sont donc pas bien adaptées à applications en temps réel.

L'utilisation de la méthode des réseaux de neurones (RNA) en particulier pour la prédiction de la charge électrique permet d'éviter les inconvénients rencontrés par les méthodes classiques. Cette méthode est commode pour prendre en compte les différents facteurs non linéaires et aléatoires. Les réseaux de neurones fonctionnent comme une boîte noire dont le fonctionnement sera déterminé par la phase d'apprentissage, c'est pourquoi, le calcul est très rapide. Cette méthode peut en étant confronté à différents types de situations évoluer (grâce à l'auto apprentissage) vers un modèle adaptatif

6.4.2 Théorie des réseaux de neurones artificiels

L'étude des réseaux de neurones est une voie prometteuse de l'intelligence artificielle, qui a des applications dans de nombreux domaines :

- Industrie : contrôle qualité, diagnostic de pannes, corrélations entre les données fournies par différents capteurs, analyse de signature ou d'écriture manuscrite.

- Finance : prévision et modélisation du marché (cours de monnaies), sélection d'investissements, attribution de crédits.
- Télécommunications et informatique : analyse du signal, élimination du bruit, reconnaissance de formes (bruits, images, paroles), compression de données.
- Environnement : Evaluation des risques, analyse chimique, prévisions et modélisation météorologiques, gestion des ressources.

6.4.2.1 Modèle des réseaux de neurones (RNA)

Les modèles de réseaux de neurones artificiels (RNA) sont, à l'origine, une imitation du fonctionnement du cerveau. Le cerveau d'humain contient environ 10 milliards de neurones et chacun est connecté à environ 10000 autres. De manière schématique, un neurone peut être représenté par la Figure 6.2.

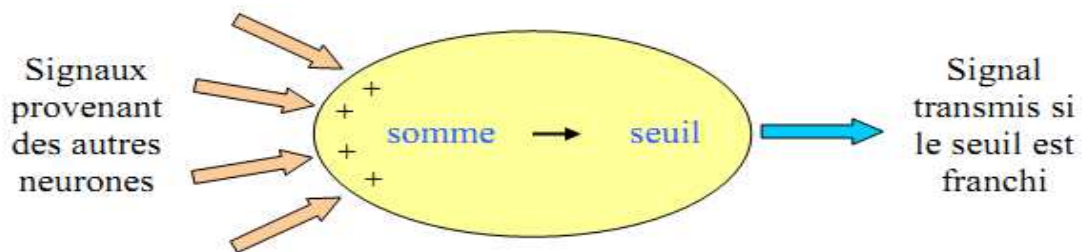


Figure 6.2 – Schématisation d'un neurone biologique

Il fait la somme de toutes les informations qu'il reçoit et émet un signal si cette somme est suffisamment élevée.

Si l'on ramène la contribution d'un neurone au cerveau tout entier, on constate que chaque neurone effectue un travail très simple par rapport au résultat obtenu. En effet, les neurones réalisent des opérations basiques, et pourtant, lorsque l'on en fait fonctionner 10 milliards ensemble, on peut créer une entité intelligente.

Cependant, un cerveau ne sait rien faire s'il n'a pas fait un apprentissage. Il a besoin d'informations venant de l'extérieur. C'est pour cela qu'il est relié aux différents organes du corps. Par exemple, il reçoit des images, des sons, des douleurs.

Grâce à ces informations il est capable de faire son apprentissage. Par exemple, lorsqu'une action a provoqué une douleur, il doit changer l'organisation des neurones afin de ne pas répéter la même erreur.

L'objectif des réseaux de neurones artificiels est donc de modéliser le fonctionnement des neurones réels, mais aussi de permettre un apprentissage.

Un neurone formel peut être représenté comme dans la Figure 6.3.

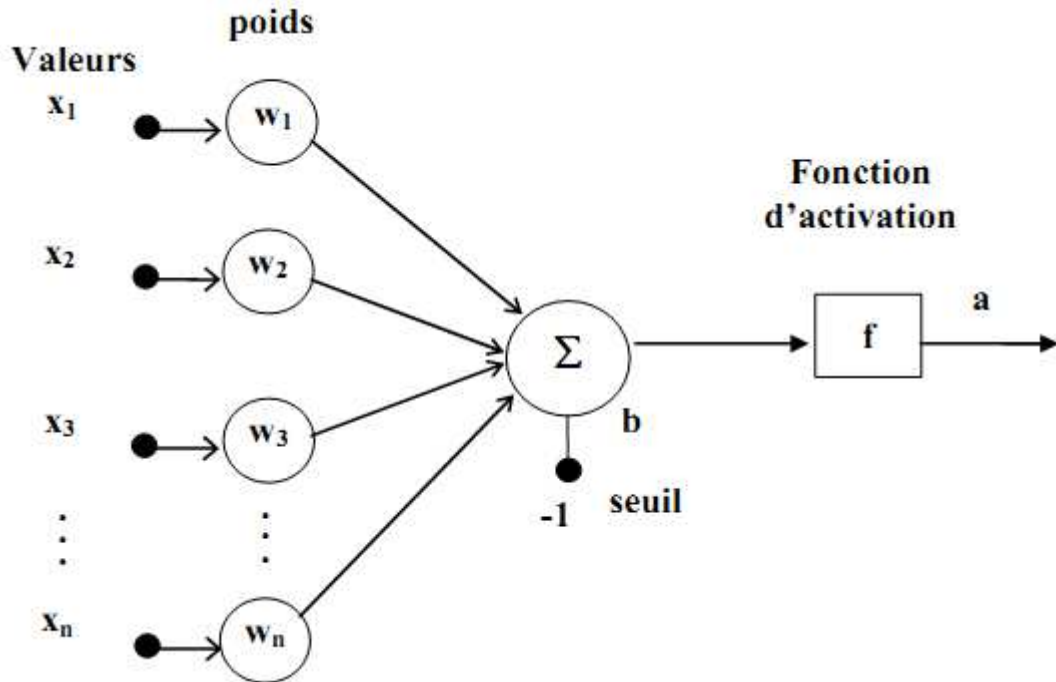


Figure 6.3 – Modèle d'un neurone artificiel

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) , 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 l'équation (6.13) (Haykin, 1999):

$$a = f \left(\sum_{j=1}^n w_j x_j - b_j \right) \quad (6.17)$$

Dans un réseau de neurones artificiel, les neurones artificiels 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 aucune, une ou plusieurs couches intermédiaires (cachées). Le modèle d'une couche des neurones est présenté sur la Figure 6.4.

Un poids est associé à chacune des connexions. L'indice i (lignes) désigne toujours le numéro de neurone sur la couche, et l'indice j (colonnes) spécifie le numéro de l'entrée. Le terme w_j désigne le poids de la connexion qui relie le neurone i à son entrée j . L'ensemble des poids d'une couche forme donc une matrice W de dimension $n \times S$ (n entrées et S neurones).

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{s1} & \cdots & w_{sn} \end{pmatrix} \quad (6.18)$$

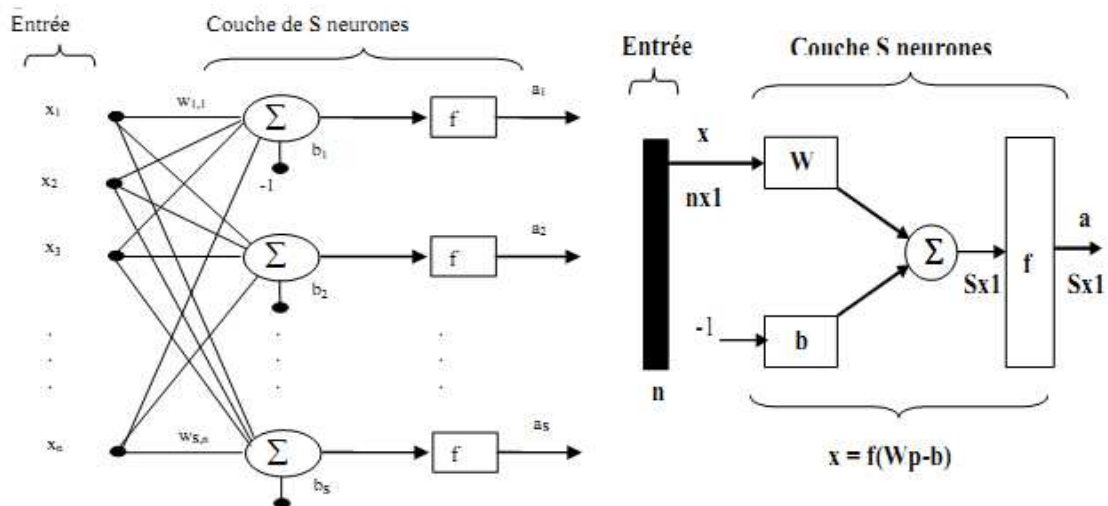


Figure 6.4 – Modèle d'une couche de n neurones (à gauche) et Représentation matricielle d'une couche à S neurones (à droite).

La fonction d'activation (ou fonction de seuil, ou encore fonction de transfert) (Tableau 6.2) sert à introduire une non-linéarité dans le fonctionnement du neurone. Le résultat après transfert est la sortie du neurone. Cet enchaînement "non-linéaire" représente finalement les propriétés "physiques" du neurone.

Nom de la fonction	Relation d'entrée/sortie	Nom Matlab
Seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$	hardlim
Seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$	hardlims
Linéaire	$a = n$	purelin
Linéaire saturée	$a = 0$ si $n < 0$ $a = 1$ si $0 \leq n \leq 1$ $a = 0$ si $n > 1$	satlin
Linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$	satlins
Linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$	poslin
Sigmoïde	$a = \frac{1}{1 + e^{-n}}$	logsig
Tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	tansig
Compétitive	$a = 0$ si n maximum $a = 1$ autrement	compet

Tableau 6.2 – Fonction de transfert.

6.4.2.2 Apprentissage

La notion d'apprentissage est claire et intuitive pour les humains ou les animaux. C'est une procédure cognitive qui doit faire en sorte que l'individu réalise, de manière autonome, une tâche donnée.

Pour un RNA, l'apprentissage peut être considéré comme le problème de la mise à jour des poids des connexions au sein du réseau, afin de réussir la tâche qui lui est demandée. L'apprentissage est la caractéristique principale des RNA et il peut se faire de différentes manières et selon différentes règles. Les procédures d'apprentissage peuvent se subdiviser en deux types :

Apprentissage supervisé : Dans ce type d'apprentissage, le réseau s'adapte par comparaison entre le résultat qu'il a calculé, en fonction des entrées fournies, et la réponse attendue en sortie.

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.

Apprentissage non supervisé : Dans ce type, l'apprentissage est basé sur des probabilités. Le réseau va se modifier en fonction des régularités statistiques de l'entrée et établir des catégories, en attribuant et en optimisant une valeur de qualité, aux catégories reconnues.

Les méthodes d'apprentissage supervisé sont largement utilisées. Cependant d'autres types de réseaux de neurones peuvent être obtenus par des techniques d'apprentissage non supervisé. Des réseaux non supervisés peuvent être utilisés, parfois, pour identifier certaines caractéristiques dans un groupe de données.

Il existe diverses structures de réseau de neurones qui sont plus ou moins performantes selon l'application à laquelle ils sont destinés. La classification plus reconnue fait la différence entre les Réseaux "FEED-FORWARD" (ce sont des réseaux dans lesquels l'information se propage de couche en couche sans retour en arrière possible) et les réseaux "FEED-BACK" appelés aussi "réseaux récurrents" (ce sont des réseaux dans lesquels il y a retour en arrière de l'information).

Dans notre étude nous utiliserons un modèle largement répandu appelé le multicouche (PMC Perceptron Multi-Couches) (Cybenko, 1989). Dans ce modèle, dans une couche, chaque neurone est connecté à tous les neurones de la couche précédente et de la couche suivante (excepté pour les couches d'entrées et de sorties) et il n'y a pas de connexions entre les neurones d'une même couche. Les fonctions d'activation utilisées dans ce type de réseaux sont principalement les fonctions à seuil ou sigmoïdes.

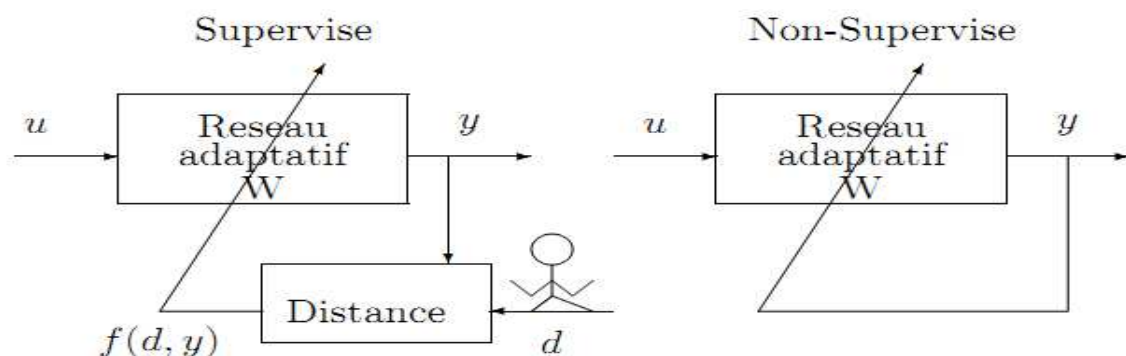


Figure 6.5 – Types d'apprentissage en RNA.

6.4.2.3 Réseau rétro-propagation

Le réseau rétro-propagation est conçu pour qu'il fonctionne comme réseau multicouche, avec propagation vers l'avant, en utilisant la fonction d'apprentissage supervisée. Cette méthode est caractérisée par le processus d'apprentissage. Ce processus est effectué par un entraînement contrôlé par un agent externe qui détermine la sortie du réseau devant une entrée déterminée.

Les réseaux de propagation en arrière possèdent la capacité d'auto-adapter les poids des neurones des couches intermédiaires pour appréhender la relation qui existe entre un ensemble de patrons d'entrée et ses sorties correspondantes. Une fois en fonctionnement, ils ont la capacité de généralisation (fournir des sorties satisfaisantes à des entrées que le système n'a jamais vues dans sa phase d'entraînement). Une architecture simple de la structure d'un réseau multicouche du type rétro-propagation est présentée dans la figure 6.6 ci-dessous.

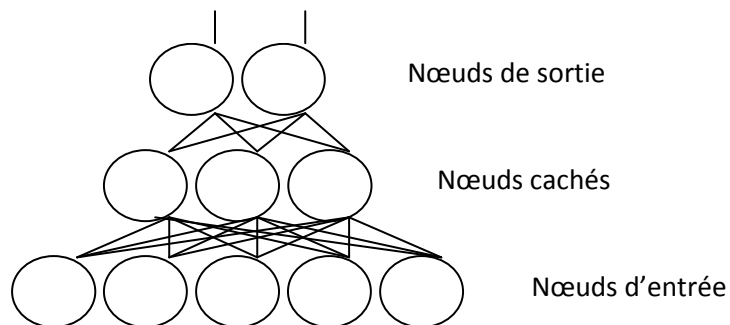


Figure 6.6 – Architecture d'un réseau rétro-propagation.

La procédure d'entraînement des réseaux d'apprentissage supervisé du type rétro-propagation suit les étapes suivantes :

- Sélection des données d'entrée pour le réseau
- Choix de l'architecture du réseau
- Entraînement du réseau
- Simulation du réseau afin d'obtenir la réponse à de nouvelles entrées.

La Figure 6.7 illustre ces différentes étapes :

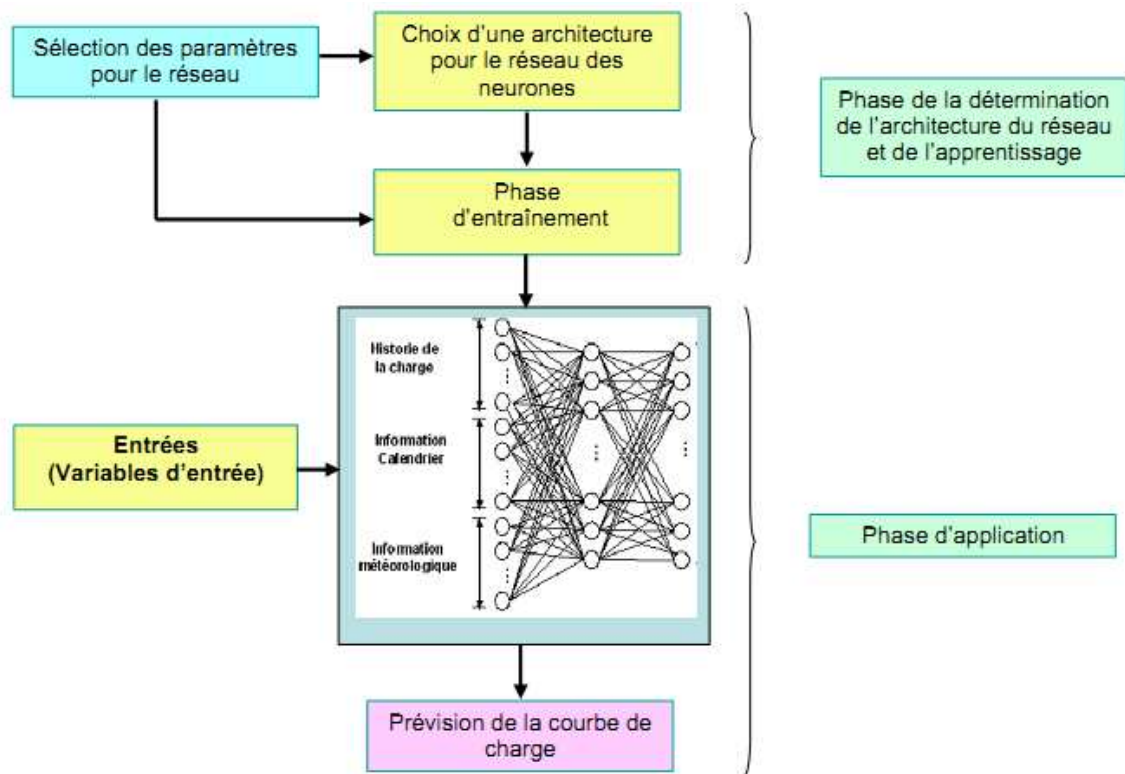


Figure 6.7 – Détermination et application d'un réseau rétro-propagation.

Lorsque les entrées d'un perceptron multi-couche représentent les valeurs antérieures d'une série chronologique et des entrées externes, ceci est équivalent à un modèle *ARX* (Connor *et al.*, 1992). Il a été prouvé que le PMC est un approximateur universel (Funahashi, 1989) de fonctions continues.

Le nombre de couches cachées et les connections entre les nœuds forment la topologie du réseau.

Dans le contexte de la prédiction à court terme de la charge électrique, des études utilisant le plus souvent une seule couche cachée (Reinschmidt, 1995 ; Lee *et al.*, 1996) ont fait l'objet de nombreux travaux scientifiques (Chen *et al.* (1992 ; Mohamad *et al.*, 1996 ; Drezga *et Rahman*, 1998 ; Lu *et Vemuri*, 1993).

D'autres auteurs par contre pensent qu'une seule couche cachée nécessite un nombre important de nœuds pour modéliser la charge. Ainsi, deux couches cachées donnent de meilleurs performances, c'est le cas des travaux effectués par (Hsu *et Yang*, 1991b) et (Kalaitzakis *et al.*, 2002).

La prochaine étape consiste à déterminer le nombre de nœuds dans chaque couche cachée et le nombre de connections. L'approche naturelle se fait en testant différentes topologies et sélectionner la meilleure (par exemple, (Mohammed *et al.*, 1996 ; Drezga *et Rahman*, 1999 ; Lee *et al.*, 1992 ; Gross *et Wagner*, 1996 ; Kiartzis *et al.*, 1995 ; Darbellay *et Slama*, 2000 ; Peng *et al.*, 1992).

Une autre approche dite network pruning (Weigend et al., 1991 ; Hassibi et al., 1992 ; Le Cun et al., 1990) consiste à utiliser un réseau partiellement connecté. Cette approche élimine les connections ayant des poids insignifiants pour en faire un réseau moins compliqué. De plus, si tous les poids associés à un nœud n'existent plus alors ce nœud est enlevé du réseau. Des exemples appliquées à la prédiction de la charge à court terme sont données par Lu (Lu et al., 1993) et Chen (Chen et al., 1992).

Pour la prédiction de la charge électrique, l'algorithme back propagation pour le perceptron multi-couche est le plus populaire (Hipper et al., 2001). Cette approche a été utilisée par exemple par Chiu (Chiu et al., 1997 ; Dash et al., 1995a).

D'autres utilisent les algorithmes génétiques (Global search algorithms) pour l'apprentissage (Mars et al., 1996). C'est le cas de Srinivasan (Srinivasan, 1998) et Yang et Huang (Yang et Huang, 1996). Le principal inconvénient est la lenteur de convergence lorsque l'espace de recherche est grand.

La connaissance traditionnelle de la modélisation des données et le développement récent de la théorie d'apprentissage indiquent clairement qu'après un point critique, le RNA entraîné avec la rétro propagation continuera à faire mieux dans la base d'apprentissage mais la performance de base du test va commencer à se détériorer. Ce phénomène est appelé sur apprentissage (Rasool et Al-kababjie, 2009). L'arrêt prématuré (early stopping) consiste à arrêter l'apprentissage au point de la généralisation maximale.

6.5 Cas d'application : Approche de prédiction par système expert flou.

Dans cette partie, nous allons présenter un exemple de prévision de la charge électrique algérienne en utilisant un système expert basé sur la logique floue avec génération automatique de règles d'inférence. Les règles utilisées sont celles définies par Mamdani (Mamdani, 1977). Ce travail utilise les classes identifiées au chapitre 2, 3 et 4.

6.5.1 Méthodologie de conception du système

Le système projeté est composé des étapes suivantes :

- Préparation des données initiales
- Choix de la base
- Architecture du système projeté
- Apprentissage
- Validation
- Test et performance

6.5.2 Préparation des données initiales

Notre base de données contient des valeurs avec une échéance d'un quart d'heures. Ces données représentent le relevé de la charge électrique par région (région1 : Oran, région2 : Alger, région5 : Hassi Messoud, et région6 : Ain Amenass). Les types de jours sont les journées de la semaine (1 samedi, 2 dimanche,...).

Le rôle du prétraitement est de préparer les données à la phase d'apprentissage, en les normalisant, c'est-à-dire de les rendre dans l'intervalle [0..1], en divisant la totalité des valeurs sur la valeur maximale.

Nous procédons par la suite à l'extraction de données en sélectionnant à partir des données initiales les paramètres suivants : La consommation électrique nationale, le type jour, la date et une entrée additionnelle pour indiquer si c'est le ramadhan. (Valeur binaire).

6.5.3 Choix de la base

Nous divisons la base en 4 sous bases correspondant aux 4 classes utilisées au chapitre 5. Chaque classe correspond à une saison. Nous ne traiterons pas les deux classes restantes (Ramadhan et jours fériés).

6.5.4 Architecture du système projeté

L'architecture du système est la même pour tous les cas, seulement les paramètres, ainsi que les données changent d'un cas à un autre. Dans notre système expert on peut faire sortir autant d'heures qu'on peut, nous nous sommes, pour des raisons de précision, limités à 24 heures de prévision seulement. Le système a comme entrée 24 heures, le type de jour, la date et comme sortie les 24 heures de la date suivante.

On a proposé aussi la récupération automatique de la date à partir de la date du système d'exploitation utilisé, mais il ya toujours possibilité de la mettre à jour par l'utilisateur. Pour chaque date correspond une période et par conséquent, en fonction des entrées, une seule base de règles (parmi celles sauvegardées en phase d'apprentissage) sera sélectionnée.

L'architecture du système aura la forme suivante :

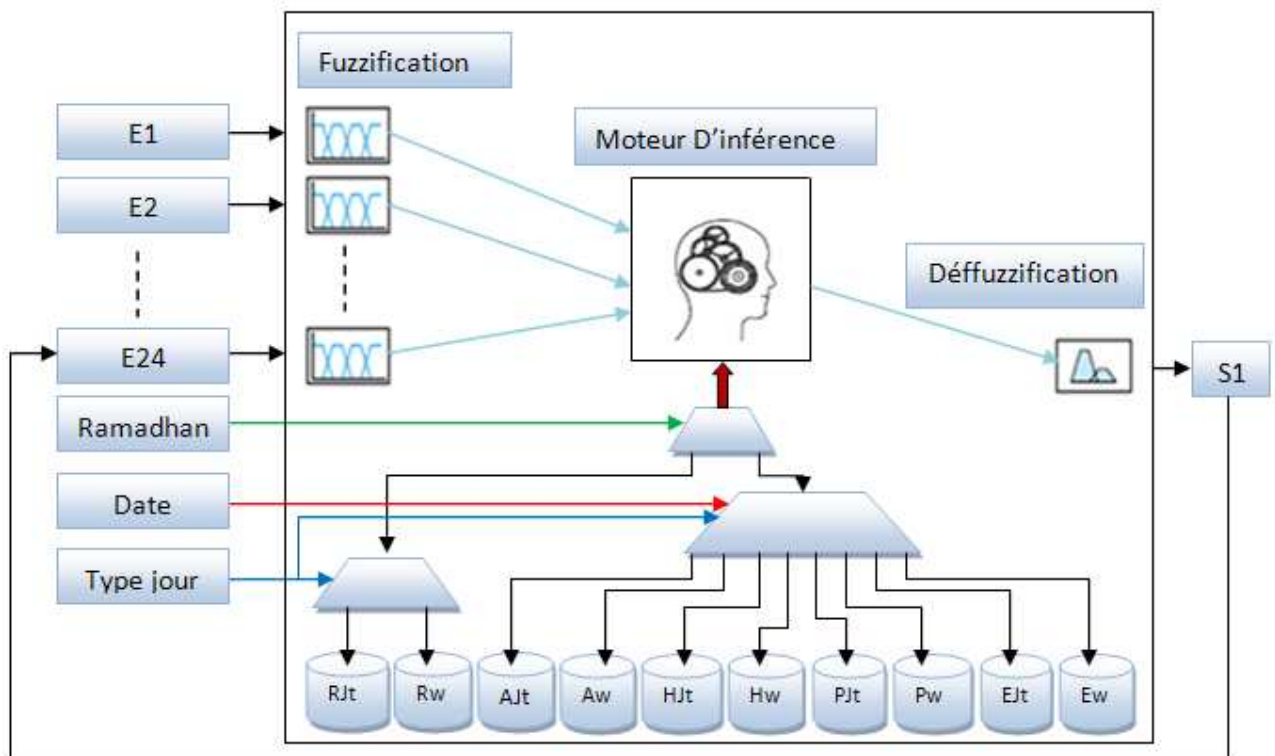


Figure 6.8 - Architecture du système

L'enchaînement des étapes est comme suit :

- Déterminer un ensemble de règles floues
Fuzzifier les entrées en utilisant les fonctions d'appartenance d'entrée
Combiner les entrées fuzzifiées selon les règles floues pour établir une règle forcée.
- Trouver la conséquence de la règle en combinant la règle forcée et la fonction d'appartenance des sorties.
- Défuzzifier la distribution des sorties.

6.5.5 Apprentissage

Dans cette étape nous allons utiliser le pourcentage de données affectées à l'apprentissage pour générer les règles d'inférences. Pour ce fait, on utilise la méthode de génération automatique des règles d'inférences expliquée au paragraphe 6.3.3. Une fois l'apprentissage terminé on passe à la validation et le test.

6.5.6 Validation

Le but principal de la validation est de vérifier si les règles d'inférences sont valides. Ainsi, on introduit la base de validation, la matrice d'entrées /sortie est calculée, On isole ensuite les entrées de la sortie en ignorant les sorties réelles temporairement et on calcule d'après nos règles la sortie dite prédite. On compare les valeurs prédites et réelles en associant l'erreur moyenne absolue l'EMA (Palit et Popovic, 2005). Après plusieurs itérations d'ajustement et des qu'une erreur acceptable est atteinte, les règles sont validés puis sauvegardés.

6.5.7 Test

En utilisant la base de test, on injecte les 24 premières valeurs comme paramètres d'entrée du système, dans un premier pas on aura une valeur prédite. On décale les entrées d'un pas et on injecte la valeur prédite à la place de la dernière entrée (E_{24}) et ainsi de suite jusqu'à 24 pas en avant. On obtient de cette manière les 24 sorties qui correspondent à une journée.

On peut augmenter l'horizon de la prévision mais il y a risque d'agrandir l'erreur de prévision, qui se multiplie d'une itération à une autre. C'est la raison pour laquelle, on limitera l'horizon à 24 pas en avant.

6.5.8 Résultats de calcul

Dans ce chapitre, nous mettons en pratique les différentes étapes de conception.

Le tableau 6.3 résume les paramètres utilisés : σ_a , σ_b , D , d , L , et n .

σ_a	σ_b	D	d	L	n
0.04	0.04	24	1	1	121

Tableau 6.3 Tableau des paramètres d'apprentissage choisis

L'entier n , la valeur de Sigma_a , et la valeur de Sigma_b , sont choisies de tel sorte que deux ensembles flous qui se suivent soient chevauchés. Des recherches ont montré que la qualité de la prévision dépend de l'étendu de ce chevauchement, il ne faut pas qu'il soit trop étroit, ni trop large (Palit et Popovich,2005).

Pour modéliser le comportement de chaque série, on fait un jeu de test sur ces paramètres de modélisation (Sigma_a , Sigma_b , D , d , L , n).

La quantité d'information utilisée lors de la validation et test sont fixés avec les proportions suivantes : Une semaine pour la validation et une journée pour le test.

Séparation en quatre saisons (4 classes)

Dans ce cas, on construit les 4 sous séries temporelles, automne, hiver, printemps, été. Chaque série correspond à un cluster ou classe définie au chapitre 5.

Les résultats obtenus sont :

Pour l'automne : On a utilisé 80% des jours d'automne pour l'apprentissage, une semaine pour la validation, et une journée pour le test. Les résultats obtenus sont :

Pour la validation :

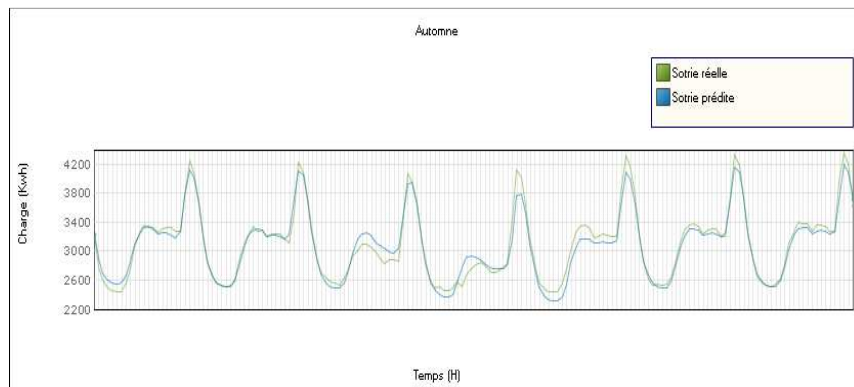


Figure 6.9 - Valeurs réelles et prédites en phase validation pour l'automne.

L'évolution de l'erreur de validation est illustrée par la figure 6.10 :

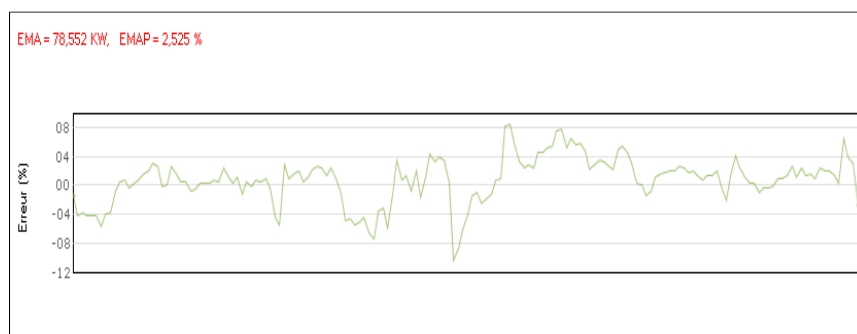


Figure 6.10 - Erreur en pourcentage, concernant la série de l'automne en phase de validation.

Valeurs réelles et prédites en phase test pour l'automne :

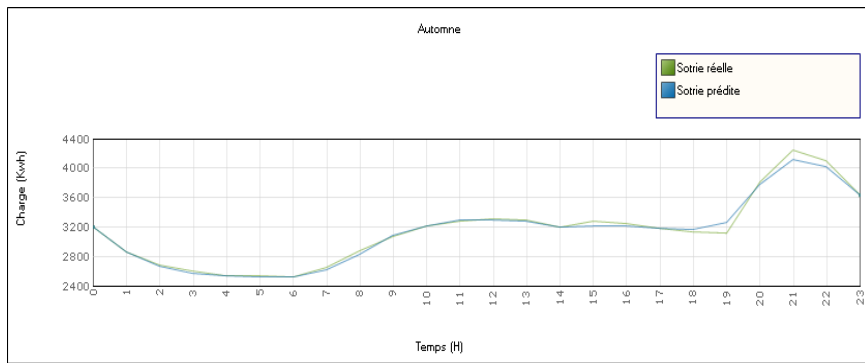


Figure 6.11 - Valeurs réelles et prédites en phase test pour l'automne.

Evolution de l'erreur en pourcentage, concernant la série de l'automne en phase de test.

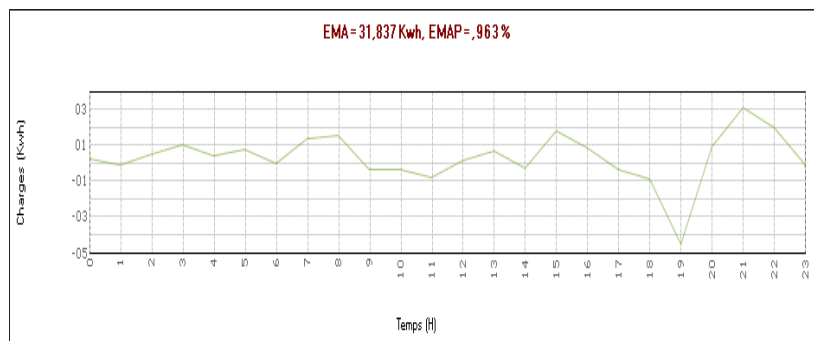


Figure 6.12 - Erreur en pourcentage, concernant la série de l'automne en phase de test.

Pour l'été : On a utilisé 70% des jours d'été pour l'apprentissage, une semaine pour la validation, et une journée pour le test. Voilà les résultats obtenus :

Valeurs réelles et prédites en phase de validation pour l'été.

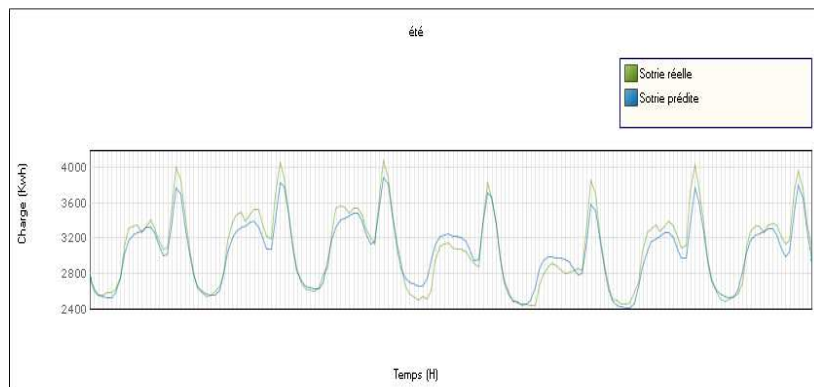


Figure 6.13 - Valeurs réelles et prédites en phase validation pour l'été.

Evolution de l'erreur en pourcentage, concernant la série de l'été en phase de validation.

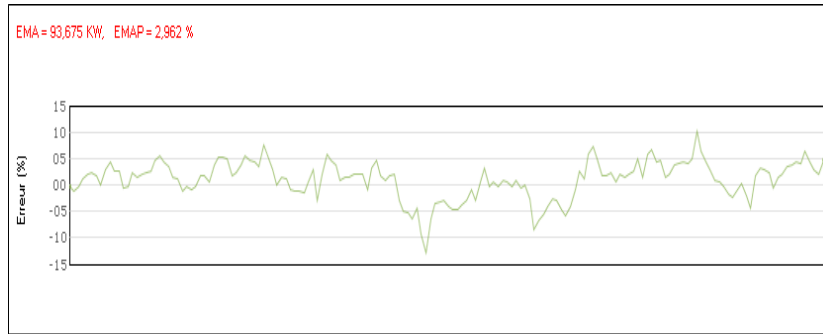


Figure 6.14- Erreur en pourcentage, concernant la série de l'été en phase de validation.

Valeurs réelles et prédites en phase de test pour l'été.

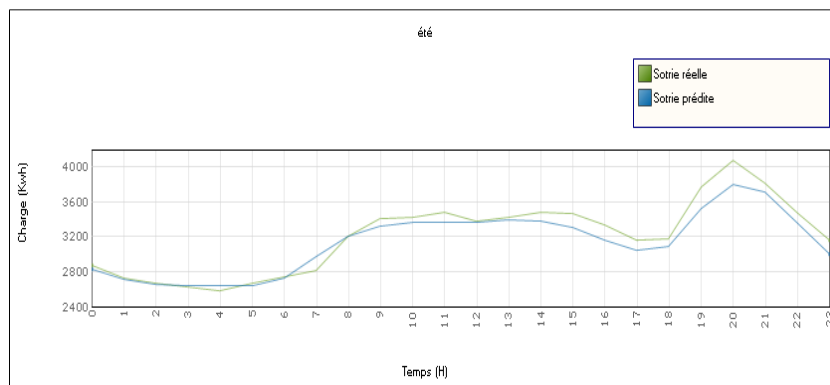


Figure 6.15 - Valeurs réelles et prédites en phase de test pour l'été.

Evolution de l'erreur en pourcentage, concernant la série de l'été en phase de test.

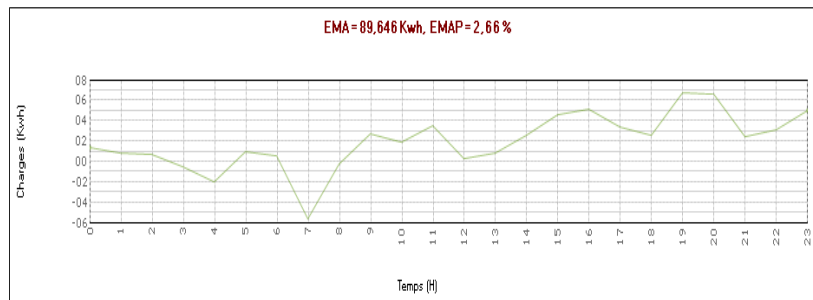


Figure 6.16 - Erreur en pourcentage, concernant la série de l'été en phase de test.

Pour l'hiver : On a utilisé 70% des jours d'hiver pour l'apprentissage, une semaine pour la validation, et une journée pour le test. Voilà les résultats obtenus :

Valeurs réelles et prédites en phase validation pour l'hiver.

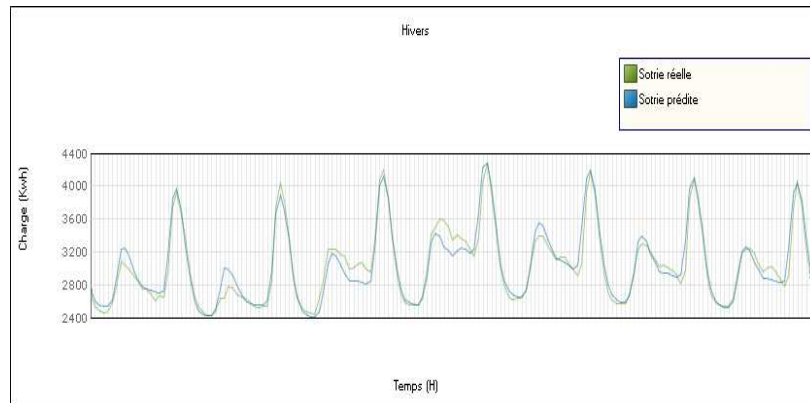


Figure 6.17- Valeurs réelles et prédites en phase validation pour l'hiver.

L'Evolution de l'erreur en pourcentage, concernant la série de l'hiver en phase de validation.

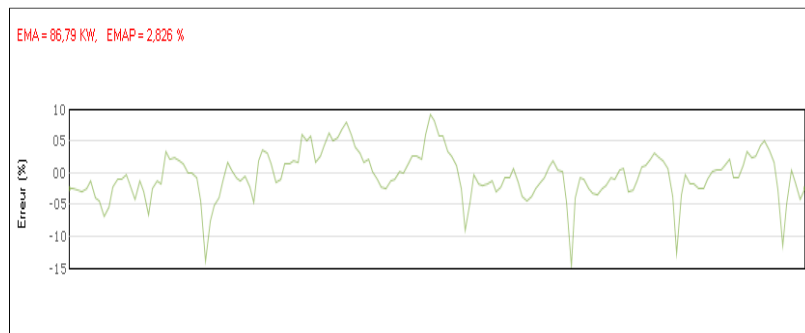


Figure 6.18 - Erreur en pourcentage, concernant la série de l'hiver en phase de validation.

Valeurs réelles et prédites en phase de test pour l'hiver.

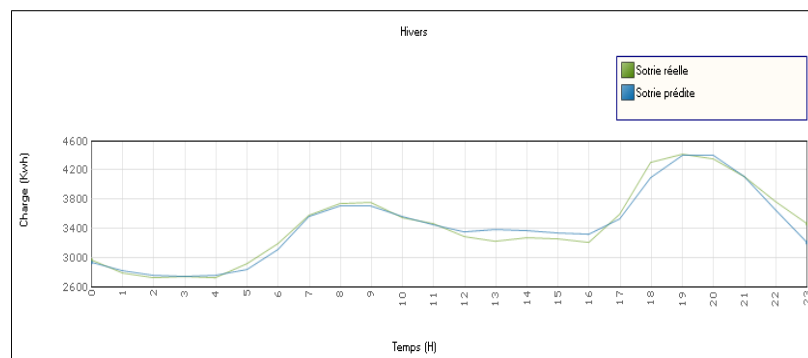


Figure 6.19 - Valeurs réelles et prédites en phase de test pour l'hiver.

L'Evolution de l'erreur Erreur en pourcentage, concernant la série de l'hiver en phase de test.

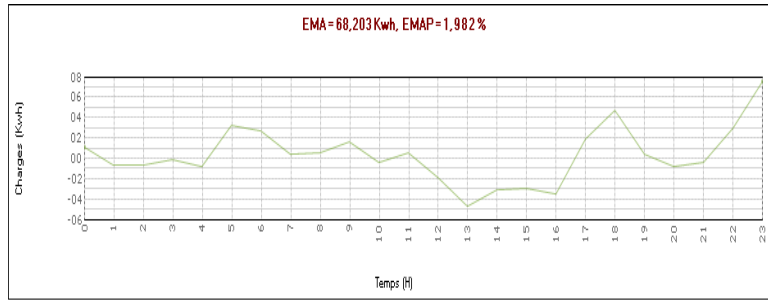


Figure 6.20 - Erreur en pourcentage, concernant la série de l'hiver en phase de test.

Pour le printemps : On a utilisé 85% des jours de printemps pour l'apprentissage, une semaine pour la validation, et une journée pour le test. Voilà les résultats obtenus :

Valeurs réelles et prédites en phase validation pour le printemps.

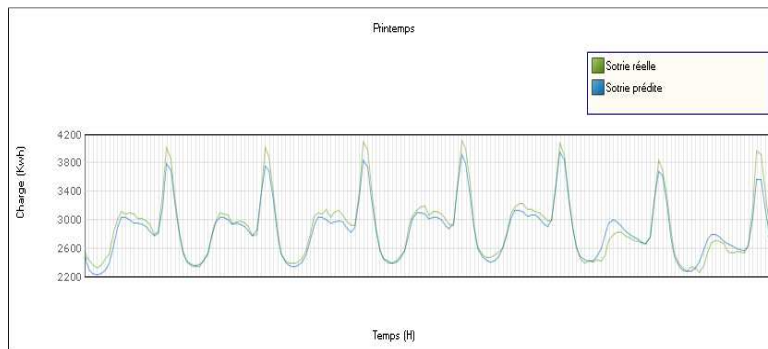


Figure 6.21 - Valeurs réelles et prédites en phase validation pour le printemps.

Evolution de l'erreur en pourcentage, concernant la série du printemps en phase de validation.

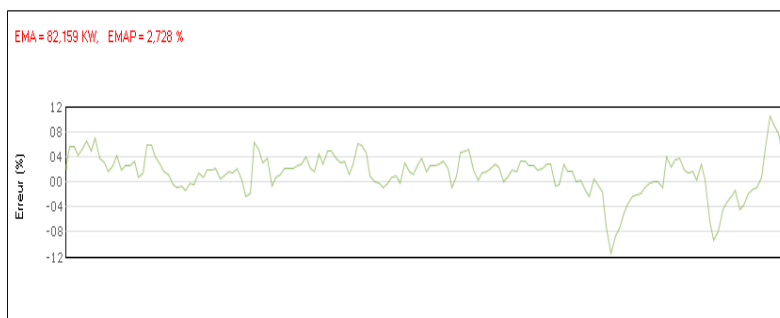


Figure 6.22 - Erreur en pourcentage, concernant la série du printemps en phase de validation.

Valeurs réelles et prédites en phase de test pour le printemps

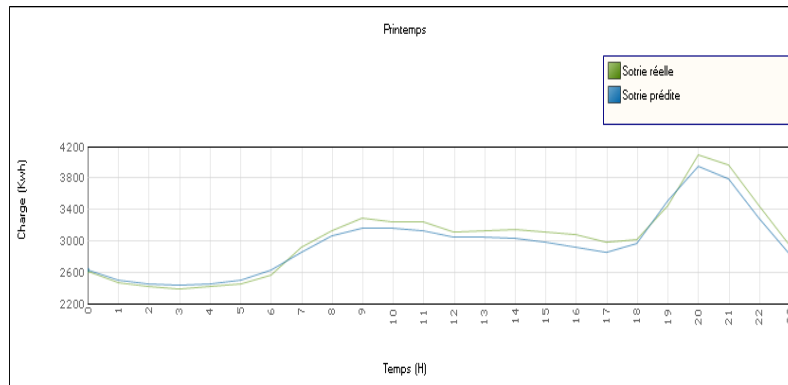


Figure 6.23 - Valeurs réelles et prédites en phase de test pour le printemps

Evolution de l'erreur en pourcentage, concernant la série du printemps en phase de test.

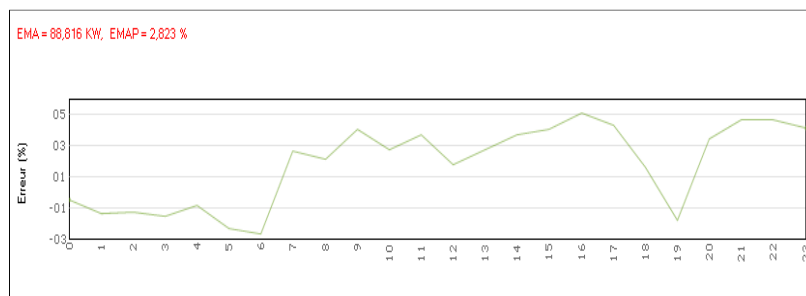


Figure 6.24 - Erreur en pourcentage, concernant la série du printemps en phase de test.

Interprétation :

Dans cette partie, la série totale est divisée en quatre séries, une pour chaque saison. Les tableaux 6.4 et 6.5 résumant les résultats de validation et de test.

Validation :

Type Série	Apprentissage	Validation	EMA(KW)	EMAP(%)
Eté	70 %	168	93.675	2.962 %
Automne	80 %	168	78.552	2.525 %
Hiver	70 %	168	86.79	2.82 %
Printemps	85 %	168	82.159	2.72 %

Tableau 6.4 Tableau des résultats de la phase de validation.

Test :

Type Série	EMA(KW)	EMAP(%)
Eté	89.646	2.66 %
Automne	31.837	0.963 %
Hiver	68.203	1.982 %
Printemps	88.216	2.823 %

Tableau 6.5 Tableau des résultats de la phase de test.

Les résultats obtenus sont très satisfaisants et montrent l'efficacité de cette approche.

6.6 Conclusion

Au cours de ce chapitre consacré à la prédiction de la charge électrique, nous avons introduit les notions de base, les différents types de prédiction : les approches classiques : régression linéaire, Box et Jenkins basé sur les séries chronologiques mais aussi des approches plus élaborés : système expert flou, réseaux de neurones.

Sur la base de la classification élaborée aux chapitres 4 et 5, un exemple d'approche de prédiction de la charge électrique algérienne a été initié. Cet exemple utilise un système expert basé sur la logique floue avec génération automatique de règles d'inférence. Ces règles sont générées avec un moteur d'inférence de type Mamdani. Cette approche a donné des résultats très satisfaisants.

7

Conclusion générale

Les résultats proposés à l'issue de ce travail de recherche apportent une contribution aux méthodes d'identification de types de jours pour la prédiction de la charge électrique algérienne à la fois nationale et régionale.

Avant de nous pencher sur le problème d'identification, nous avons présenté au cours du chapitre 2, l'analyse de la charge électrique algérienne et les notions de séries chronologiques. Le troisième chapitre a exposé l'état de l'art des approches utilisées en classification automatique ainsi que leurs évaluations par les indices de validités.

La principale contribution de ce travail apparait au cours du quatrième, cinquième et sixième chapitre.

Des approches hybrides de classification ont été mises en œuvre pour la charge électrique algérienne (chapitre 4) et régionale (Chapitre 5) pour pallier aux insuffisances des approches classiques.

Le quatrième chapitre étudie l'identification des types de jours dans le contexte de la charge électrique nationale. L'algorithme des K-means combiné aux cartes de Kohonen avec l'indice de validité de Davies-Bouldin a permis de déterminer le nombre de classes égale à 6. Cette approche s'applique dans le cas où les classes sont hyper-sphérique compactes et bien séparées. Cependant, l'étude de la charge électrique globale n'est pas significative par rapport aux spécificités de chaque région. De plus, l'application de SOM est tributaire d'un bon choix des paramètres permettant de définir la carte SOM. Aussi, il importe de faire un bon choix des paramètres de la carte de Kohonen (Benabbas et al., 2007).

Au cours du cinquième chapitre, l'identification est consacrée à la charge électrique régionale en utilisant les cartes de Kohonen puis une approche hybride combinant SOM et K-moyenne d'une part et SOM et Fuzzy C-means d'autre part. Des critères de qualités pour le bon choix de l'architecture SOM et le nombre optimal de classes sont envisagées à la fois pour la classification automatique dure et floue (Benabbas et al., 2009, 2010,2011).

Dans le sixième chapitre, sur la base de cette classification a été élaboré un système expert flou avec génération automatique des règles d'inférence. Ces règles sont générées avec un moteur d'inférence de type Mamdani. Cette approche a donnée des résultats très satisfaisants.

Le secret du choix d'une approche de prédiction appropriée, réside avant tout dans la bonne connaissance des données mises à disposition. Cette connaissance, est en grande partie donnée par l'identification des types de jours quand une prédiction à court terme est projetée. L'analyse de la charge électrique algérienne établis en deux niveaux d'abstraction (SOM-K-

moyenne et SOM-FCM), a donnée d'excellents résultats. En effet, les classes des types de jours de la charge électrique Algérienne, national et régional, ont été définis avec succès, s'en est suivi une meilleure compréhension des phénomènes contenu dans la série temporelles constituant cette dernière. Sur la base de ses connaissances et de la classification obtenu, à savoir : six (06) classes pour la charge nationale, quatre (04) classes pour les villes du nord et trois (03) classes pour les villes du sud; des approches de prédictions peuvent alors être considérés dans le cadre d'un système multi-modèle. Dans notre cas, un exemple donné par un Système Expert- Flou a été implémenté et tester sur la base de la classification obtenu précédemment. Le prédicteur multi-modèles a démontré de bonne aptitudes à prédire la charge électrique surtout en l'absence d'entrées exogènes tel que la température, le PIB, etc.

Plusieurs perspectives de développement peuvent être envisagées comme suite à ce travail :

Dans cette thèse, nous avons traité la segmentation des données de la charge avec l'identification par type de jour, on pourra s'intéresser à une segmentation par heure du jour. les modèles séquentiels développés par Choueiki et Papadakis (Choueiki et al., 1997, Papadakis et al., 1999) et les modèles parallèles traitées par Hill et Gupta (Hill, 1998 et Gupta, 1985) où L'approche dite Multi-timeScale (MTS) (Gupta, 1985, Khotanzad et al., 1996, Murray et al., 2000) qui combine les deux peuvent être des exemples d'utilisation.

Nous pouvons envisager également l'utilisation d'autres algorithmes de classification bien adaptés aux séries chronologiques ainsi que d'autres types de visualisation des classes obtenues, ce domaine est toujours d'actualité.

D'autres types de modélisation non-linéaire peuvent être envisagées. L'approche neuronale avec les Perceptron Multi-Couches (PMC) ou l'approche neuro-fuzzy combinant les deux en se basant sur la segmentation des données. Les recherches actuelles ayant montré l'efficacité de ces méthodes notamment en associant d'autres facteurs pertinents pour la charge électrique tels que la température, l'humidité, les précipitations, la luminosité et le BIP (produit Intérieur Brut).

Pour finir, l'un des enjeux actuels c'est la capacité de traiter de grandes masses de données. Des approches de segmentation doivent être élaborées à cet effet et appliquer le type de prédiction approprié pour des prédictions à moyen et long terme.

Bibliographie

Ajoy K.P., Popovic, P. (2005). Computational Intelligence in Time Series Forecasting Theory and Engineering Applications, Springer, London.

Barakat, E.H., Qayyum, M.A., Hamed, M.N., Al Rashed, S.A. (1990). Short-term peak demand forecasting in fast developing utility with inherit dynamic load characteristics. I. Application of classical time-series methods. II. Improved modelling of system dynamic load characteristics, IEEE Transactions on Power Systems, 5(3), 813-824.

Benabbas, F., Khadir, M.T., Fay D., Boughrira A. (2007). Regional Day Type Identification For Algerian Electricity Load Using Kohonen Map. In Proceedings of The International Conference of Neural Networks, EANN 2007, Thessaloniki, Greece, 29-31 aout.

Benabbas, F., Khadir, M.T., Fay D., Boughrira A. (2008). Cartes de Kohonen Combiné à L'algorithme des K-Means pour l'identification des Types de Jours de la Charge Electrique Algerienne. Proceedings de la Conférence Internationale francophone d'automatique, CIFA 2008, Bucarest, Roumanie, 3-5 septembre.

Benabbas, F., Khadir, M.T., Fay D., Boughrira A. (2008). Kohonen Map Combined to the K-Means Algorithm for the Identification of Day Types of Algerian Electricity Load. Proceedings of The International Conference of Computer Information Systems and Industrial Management Applications, CISIM 2008, Ostrava, Czech Republic, 26-28 juin.

Benabbas, F., Khadir, M.T., Fay D., Boughrira A. (2008). Kohonen Map Combined to the K-Means Algorithm for the Identification of Day Types of Regional Algerian Electricity Load. Proceedings of Second International Conference on Electrical Engineering Design and Technologies, ICEEDT 2008, Hammamet, 8-10 novembre, Tunisia.

Benabbas, F., Khadir, M.T. (2011). Fuzzy c-means clustering and kohonen maps for the indentification of regional electricity load day types. International Journal of Hybrid Intelligent Systems, 8(2), 81-82.

Bensaid, A.M., Hall, L.O Bezdek, J.C., Clarke, L.P., Silbiger M.L. and Arrington J.A., Murtagh, R.F. (1996). Validity-guided (Re)Clustering with applications to image segmentation. IEEE Transactions on Fuzzy Systems, 4, 112-123.

Benlembarek , K., Khadir, M.T., Benabbas F. (2010). A Web Based System for Short-Term Forecasting of Algerian Electricity Load Using Artificial Neural Network, Journal of Automation and Systems Engineering, 4(2), 94-100.

Bezdek, J.C. (1974a). Numerical taxonomy with fuzzy sets. J. Math. Biol., 57-71.

Bezdek, J.C. (1974b). Cluster validity with fuzzy sets. J. Cybernet. 3, 58-78.

Bezdek, J.C. (1981). Pattern recognition with fuzzy objective function algorithms. New York, NY : Plenum Press.

Bezdek, J.C. and Pal, N.R. (1988). Some new indexes of cluster validity, *IEEE Transactions on Systems, Man and Cybernetics – Part B*, 28, 301-315.

Bowerman, B., O'Connell, R.T. (1987). *Time series forecasting unified concepts and computer implementations*, Duxbury press.

Box, G.E.P., Cox, D.R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society*, 26 (B), 211-243.

Box, G.E.P., Jenkins, G.M. (1970). *Time Series Analysis : Forecasting and Control*, San Francisco: Holden Day.

Bretschneider, P., Rauschenbach, T., Wernstedt, J. (1999). Forecast using an adaptive fuzzy classification algorithm for load, In: *Proceedings, the 5th European Control Conference*, Karlsruhe, Germany, August, (not paginated on CD-ROM).

Brockwell, P.J., Davies, R.A. (2009). *Times Series : Theory and Methods*, 2nd edition, Springer.

Calinski, R., Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics* , 3, 1-27 .

Cannon, R.L, Dave J.V, Bezdek J.C. (1986). Efficient implementation of the fuzzy c-means clustering algorithm. *IEEE Trans. Pattern Anal. Machine Intell*, 8, 248-255.

Charytoniuk, W., Chen, M.S., Kotas, P., Van Olinda, P. (1999). Demand forecasting in power distribution systems using non-parametric probability density estimation, *IEEE Transactions on Power Systems*, 14(4), 1200-1206.

Chen, B. J., Yang, B., Dong, J. et Abraham, A. (2004). Load forecasting using support vector machines. *IEEE Transactions on Power Systems*, 19, 4, 1821-1830.

Chen, S.T., Yu, D.C., Moghaddamjo, A.R., (1992). Weather sensitive short-term load forecasting using non-fully connected artificial neural network, *IEEE Transactions on Power Systems*, 7 (3), 1098-1104.

Chi, S.C, Yang,C.C., (2008). A two stages clustering Method combining Ant Colony SOM and K-means, *Journal Of Information Science and engineering*, 24, 1445-1460.

Chiu, C.C., Kao, L.J., Cook, D.F. (1997). Combining a neural network with a rule based expert system approach for short term power load forecasting in Taiwan, *Expert Systems with Applications*, 13(4), 299-305.

Choueiki, M.H., Mount-Campbell, C.A., Ahalt, S.C. (1997). Building a 'quasi optimal' neural network to solve the short-term load forecasting problem, *IEEE Transactions on Power Systems*, 12(4), 1432-1439.

Connor, J., Atlas, L.E., Martin, D.R., (1992). Recurrent networks and NARMA modelling, *Advances in Neural Information Processing Systems*, 4, 301-308.

Dash, P.K., Satpathy, H.P., Liew, A.C., Rahman, S. (1997). A real-time short-term load forecasting system using functional link network, *IEEE Transactions on Power Systems*, 12(2), 675-681.

Dash, P.K., Liew, A.C., Rahman, S. (1995a). Peak load forecasting using a fuzzy neural network, *Electric Power Systems Research*, 32, 19-23.

Duda, R., Hart, P., Stork, D. (2001). *Pattern classification*, 2 edition, New York, NY : John Wiley & Sons.

Darbellay, G.A., Slama, M. (2000). Forecasting the short-term demand for electricity, do neural networks stand a better chance?, *International Journal of Forecasting*, 16, 71-83.

Davies, D., Bouldin, D. (1979). A cluster separation measure, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1, 224-227 .

Di Caprio, U., Genesio, R., Pozzi, S., Vinino, A. (1985). Comparison of ARMA and extended Wiener filtering for load prediction at ENEL. In: Bunn, E.D., Farmer, E.D., eds. *Comparative Models for Electrical Load Forecasting*, NY: John Wiley and sons, 109-119.

Drezga, S., Rahman, S. (1998). Input variable selection for ANN-based short-term load forecasting, *IEEE Transactions on Power Systems*, 13(4), 1238-1244.

Dubes, R. (1993). Cluster analysis and related issue, In *Handbook of Pattern Recognition and Computer Vision*, C. Chen, L. Pau, and P. Wang, Eds., River Edge, NJ : World Science Publishing Company, 3-32 .

Dreyfus, G., Martinez, J.M., Samuelides, M., Gordon, M.B., Badran, F., Thiria, S. and Hérault, L. (2004). *Réseaux De Neurones : Méthodologie et Application*, Eyrolles ISBN : 2-212-11464-8, France.

Dunn, J. (1974b). Well separated clusters and optimal fuzzy partitions, *Journal of Cybernetics*, 4, 95-104.

Elkateb, M.M., Solaiman, K., Al-Turki, Y. (1998). A comparative study of medium-weather-dependant load forecasting using enhanced artificial/fuzzy neural network and statistical techniques, *Neurocomputing*, 23, 3-13.

Eschrich, S., Ke J., Hall L., and Goldof D. (2003). Fast accurate fuzzy clustering through data reduction, *IEEE Trans. Fuzzy syst*, 11(2), 262-270.

Everitt, B., Landau, S., and Leese, M. (2001). *Cluster analysis*, 4 edition, London : Arnold .

Faya, D., D., Ringwood, J.V., Condon, M. et Kellyc, M. (2003). 24-electrical load data – a sequential of portioned time series, *Neurocomputing*, 55, 469-498.

Forgy, E. (1965) Cluster analysis of multivariate data: efficiency vs. interpretability of classifications, *Biometrics* , 21, 768-780 .

- Funahashi, K.I. (1989). On the approximate realization of continuous mapping by neural networks, *Neural Networks*, 2(3), 183-192.
- Gelper, S., Fried, R., Croux, C. (2010). Robust Forecasting with exponential and Holt-Winters smoothing, *Journal of forecasting*, 29, 285-300.
- Gordon, A. (1998). Cluster validation in *Data Science, Classification, and Related Methods*, C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H. Bock, and Y. Bada, Eds., New York, NY : Springer - Verlag , 22-39 .
- Guérif, S. (2006). Réduction de dimension en Apprentissage Numérique non supervisé, Ph.D thesis, Université Paris 13, France.
- Gupta, P.C. (1985). Adaptive short-term forecasting of hourly loads using weather information. In: Bunn, E.D., Farmer, E.D., eds. *Comparative Models for Electrical Load Forecasting*, NY: John Wiley and sons, 43-56.
- Hajek, P., Olej, V. (2007). Municipal Creditworthiness modelling by Clustering Methods. In *Proceedings of the International Conference of Neural Networks*, Thessaloniki, Greece.
- Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002). Cluster validity methods : Part I & II, *SIGMOD Record*, 31 (2 & 3).
- Hann, H., Meyer-Nieberg, Pickl, S. (2009). Electric load forecasting methods : Tools for decision making, *European Journal of operational Research*, 199, 902-907.
- Harvey, A.C. (1981). *Time Series Analysis*, Hertfordshire UK: Philip Allan.
- Hassibi, B., Stork, D.G., Wolff, G.J. (1992). Optimal brain surgeon and general network pruning, in: *Proceedings, IEEE International Conference on Neural Networks*, 1, San Francisco: CA, IEEE, 293-299.
- Haykin, S. (1999). *Neural networks: A comprehensive foundation*, 2 edition, Upper Saddle River, NJ: Prentice Hall.
- Himberg, A. (2000). SOM Based Cluster Visualization and Its Application for False Coloring, *Proceedings of International Joint Conference on Neural Networks (IJCNN)*, 3, 587-592.
- Hippert, S.H., Pedriera, C.E., Souza, R.C. (2001). Neural networks for short-term load forecasting: a review and evaluation, *IEEE Transactions on Power Systems*, 16 (1), 44-55.
- Ho, K.L., Hsu, Y.Y., Liang, C.C., Lai, T.S. (1990). Short term load forecasting of Taiwan power system using a knowledge-based expert system, *IEEE Transactions on Power Systems*, 5 (4), 1214-1221.
- Hoppner, F., Klawonn, F., and Kruse, R. (1999). *Fuzzy cluster analysis : Methods for classification, data analysis and image recognition*. New York, NY, Wiley.

Hsu, Y.Y., Yang, C.C. (1991a). Design of artificial neural networks for short-term load forecasting Part I: Self-organizing feature maps for day type identification, IEE Proceedings-C, 138(5), 407-413.

Hubele, N.F., Cheng, C.S., 1990, Identification of seasonal short-term forecasting models using statistical decision functions, IEEE Transactions on Power Systems, 5(1), 40-45.

Infield, D.G., Hill, D.C. (1998). Optimal smoothing for trend removal in short term load forecasting, IEEE Transactions on Power Systems, 13 (3), 1115-1120.

Jain, K., Dubes, R.C. (1988) Algorithms for Data Clustering, Prentice Hall, Englewoods Cliffs, new Jersey.

Jolliffe, I.T. (2002). Principal Component Analysis, Springer, New York.

Kalaitzakis, K. Stavrakakis, G.S. Anagnostakis, E.M. (2002). Short-term load forecasting based on artificial neural networks parallel implementation, Electric Power Systems Research, 63, 185-196.

Khadir, M. T., Khedairia, S., Benabbas, F. (2010). Kohonen Maps Combined to K-means in a Two Level Strategy for Time Series Clustering Application to Meteorological and Electricity Load data. Self-Organizing Maps, In Tech, 63-80.

Khadir, M. T., Benabbas, F. (2011). Kohonen Maps Combined to Fuzzy C-means, a Two Level Clustering Approach. Application to Electricity Load Data, Self Organizing Maps Applications and Novel Algorithm Design, InTech, 541-558.

Khotanzad, A., Davis, M.H., Abaye, A., Maratulum, D.J. (1996). An artificial neural network hourly temperature forecaster with applications in load forecasting, IEEE Transactions on Power Systems, 11(2), 870-876.

Khurshid, A., Bogdan, V. (2004). Learning to Visualise High-dimensional Data, Journal of International Conference on Information Visualisation, (IV'04), 2004, 507-512.

Kodogiannis, V.S., Anagnostakis, E.M. (1999). A study of advanced learning algorithms for short-term load forecasting, Engineering Applications of Artificial Intelligence, 12, 159-173.

Kohonen, T. (1990). The Self-Organising Map, Proceedings IEEE, 78 (9).

Kohonen, T. (1989). Self-organization and associative memory, 3rd edition, Berlin : Springer Verlag.

Kohonen, T. (2001). Self-organizing maps, 3 edition, Berlin, Hiedelberg : Springer .

Kosko, B. (1997). Fuzzy Engineering, N.J.: Prentice Hall.

Le Cun, Y., Denker, J.S., Solla, S.A. (1990). Optimal brain damage, In: proceedings, Advances in Neural Information Processing Systems, 2, San Mateo: CA: Morgan-Kaufmann, 598-605.

Lee, K.Y., Cha, Y.T., Park, J.H. (1992). Short-term load forecasting using an artificial neural network, *IEEE Transactions on Power Systems*, 7(1), 124 - 132.

Ljung, L. (1987). *Systems Identification – Theory for the user*, Prentice Hall.

Lu, C.N., Wu, H.T., Vemuri, S. (1993). Neural Network based short term load forecasting, *IEEE Transactions on Power Systems*, 8(1), 336-342.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations, In *Proceedings of the Fifth Berkeley Symposium*.

Mamdani, E.H. (1977). Application of fuzzy logic to approximate reasoning using linguistic systems, *Fuzzy Sets and Systems* 26: 1182-1191.

Mars, P., Chen, J.R., Nambiar, R. (1996). *Learning Algorithms: Theory and Applications in Signal Processing, Control and Communications*, N.Y.: CRC Press.

Mastorocotas P.A., Theocharis, J.B., Bakirtzis, A.G. (1999). Fuzzy modeling for short term load forecasting using the orthogonal least squares method, *IEEE Transactions on Power Systems*, 14 (1), 29-35.

Mbamalu, G.A.N., El-Hawary, M.E. (1993). Load forecasting via sub-optimal seasonal autoregressive models and iteratively reweighted least squares method, *IEEE Transactions on Power Systems*, 8 (1), 343-348.

Mogham, I., Rahman, S. (1989). Analysis and evaluation of five short-term load forecasting techniques, *IEEE Transactions on Power Systems*, 4(4), 1484-1491.

Mohamad, E.A., Mansour, M.M., El-Debeiky, Mohamad, K.G., Rao, N.D., Ramakrishna, G. (1996). Results of Egyptian unified grid hourly load forecasting using an artificial neural network with expert system interface, *Electric Power Systems Research*, 39, 171-177.

Muller, H., Petrisch, G. (1998). Energy and load forecasting by fuzzy-neural networks. In: Jurgen, H., Zimmermann, H.J., eds., *Proceedings, European Congress on Intelligent Techniques and Soft Computing*, Aachen, Germany, September. Aachen: Elite foundation, 1925-1929.

Muller, H., Schatzel, F. (1999). Daily load curve clustering and prediction by neural model tool box for power systems with non-stochastic load components, In: *Proceedings, the 5th European Control Conference*, Karlsruhe, Germany, August, (not paginated on CD-ROM)

Murray, F.T. (1996). *Forecasting methodologies for electricity supply systems*, Dublin: Ph.D. thesis, School of Electronic Engineering, Dublin City University, Ireland.

Murray, F.T., Ringwood, J.V., Austin, P.C. (2000). Integration of multi-time-scale models in time series forecasting electricity consumption, *International Journal of Systems Science*, 31 (10), 1249-1260.

Pal N.R., Bezdek J.C. (1995). On cluster validity for fuzzy c-means model. *IEEE Trans. Fuzzy Systems* 3, 370-379.

Palit, A.K. and Popovic, D. (1999). Fuzzy logic based automatic rule generation and forecasting of time series, *Proc. of FUZZ-IEEE*, 1: 360-365.

Palit, A.K. and Popovic, D. (2005). *Computational Intelligence in Time Series Forecasting, Theory and Engineering Applications*, Springer Verlag, London.

Papadakis, S.E., Theocharis, J.B., Bakirtzis, A.G. (1999). Fuzzy short-term load forecasting models based on load-curve prototype fuzzy clustering, In *Proceedings, the 5th European Control Conference*, Karlsruhe, Germany, August, (not paginated on CD-ROM)

Pedrycz, Vasilakos, Karnouskos. (2003) *IEEE Trans. on Syst. Man and Cybern.*, special issue on computational intelligence in telecommunication networks and internet service. Pt.-I, 33(3): 294-426; Pt.-II, 33(4): 429-501; Pt.-III, 34(1):1-96.

Pelikan, E., Matejka, P., Slama, M., Vinkler, K. (1996). Interactive forecasting of the electric load using Kohonen self-organizing feature maps. In: *Proceedings, International Neural Network Society 1996 Annual Meeting*, Mahwah, NJ, USA. NJ: Lawrence Erlbaum Assoc, 443-446.

Peng, T.M., Hubele, N.F., Karadg, G.G. (1992). Advancement in the application of neural networks for short-term load forecasting, *IEEE Transactions in Power Systems*, 7(1), 250–257.

Rajurkar, K.P.; Newill, R.E. (1985). Multiple series modelling and forecasting of short-term load demand by data dependent systems. In: *Proceedings of the International Conference on Cybernetics and Society*, New York, USA. NY: IEEE, 448-52.

Reinschmidt, K.F. (1995). Artificial neural networks for short term load forecasting, In: *Proceedings Annual Meeting of the American Power Conference*, Chicago, Illinois, 1144-1149.

Sinha, N.K. (1991). *Linear Systems*, NY, USA: John Wiley and sons.

Srinivasan, D., Tan, S. S., Chang, C. S., Chan, E. K. (1999). Parallel neural network-fuzzy expert system for short-term load forecasting: system implementation and performance evaluation, *IEEE Transactions on Power Systems*, 14(3), 1100-1106.

Szilágyi, S., Szilágyi, L., IclĆnzan, D., Dávid, L., Frigy, A. and Benyó Z. (2009). Intensity in homogeneity correction and segmentation of magnetic resonance images using a multi-stage fuzzy clustering approach, *International Journal on non-standard computing and artificial intelligence*, 19(5), 513-528.

Theodoridis, S., Koutroumbas, K. (2006). *Pattern recognition*, 3, San Diego, CA : Academic Press .

Tipping M.E. and Bishop C.M. (1999). Mixtures of probabilistic Principal Component Analysis, *Neural Computation* 11(1999), 443–48.

Van Laerhoven, K. (2001). Combining the Self-Organizing Map and K-Means clustering for on-line classification of sensor data. *Artificial Neural Networks - ICANN 2001, International Conference*, 2130, 464-469.

Vesanto, J., Alhoniemi, E. (2000). Clustering of the Self-Organizing Map, *IEEE Transactions On Neural Networks*, 11(3).

Vesanto, J. (1999). SOM-Based Data Visualization Methods, *Intelligent Data Analysis*, 3 (2), 111-126.

Weigend, A.S., Rumelhart, D.E., Huberman, B.A. (1991). Generalisation by weight elimination with application to forecasting, *Advances in Neural Information Processing Systems*, 3, 875-882.

Wu, S., & Chow, T. W. S. (2004). Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density, *Pattern Recognition*, 37(1), 175-188.

Xie, X., Beni, G. (1991). A validity measure for fuzzy clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8): 841-847.

Xu, R., Wunsch D.C. (2009). *Clustering*, IEEE Computational Intelligence Society, John Wiley & Sons.

Yang, H.T., Huang, C.M., Huang, C.L. (1996). Identification of ARMAX model for short-term load forecasting: an evolutionary programming approach, *IEEE Transactions on Power Systems*, 11(1), 403-408.

Yu, J., Cheng Q. Huang H. (2004). Analysis of the weighting exponent in the FCM and its application to parameter selection. *IEEE Transactions on Fuzzy Systems*, 13(1), 164- 176.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8, 338-353.