

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOHAMED BOUDIAF - M'SILA

FACULTE : TECHNOLOGIE
DEPARTEMENT D'ELECTRONIQUE
N° : 2018/ INST 11 / 87



DOMAINE : SCIENCES ET TECHNOLOGIES
FILIERE : ÉLECTRONIQUE
OPTION : INSTRUMENTATION

**Mémoire présenté pour l'obtention
Du diplôme de Master Académique**

Par: DAOUD youcef

SOUYEH khiredine

Intitulé

**SYSTEME DE PREDICTION DE LA VITESSE DU VENT BASE
SUR LA SELECTION DE CARACTERISTIQUE ET LES
RESEAUX DE NEURONES ARTIFICIELS**

Soutenu devant le jury composé de:

KOUDA souhile	Université M'sila	Président
MEZAACHE Hatem	Université M'sila	Rapporteur
BENDIB toufik	Université M'sila	Examineur

Année universitaire : 2017 /2018

Remerciements

Tout d'abord je remercie le bon dieu, le clément de m'avoir donné la foi et le courage de terminer ce projet.

Nous tiens à adresser mes plus vifs remerciements à Mon encadreur **Mezaache Hatem** qui m'a aidé à terminer ce travail.

Aussi étendre mes remerciements à Mme **Nahed ZEMOURI** que je lui souhaite Bonne chance dans ses études, pour son aide, son soutien et ses conseils qui m'ont fourni une grande motivation à accomplir ce travail.

Mon vif remerciements vont aussi à mes parents qui nous ont donné la lumière et guidés et à tous mes professeurs qui m'ont orientés, assistés et aidés à concevoir ce mémoire ainsi qu'à tous ceux qui, durant ma vie scolaire et universitaire m'ont apporté leurs précieux conseils et m'ont permis d'être ce que je suis aujourd'hui.

Nous e tiens à exprimer toute ma gratitude à tous ceux qui ont contribué, de près ou de loin, à la concrétisation de ce travail.

Mes remerciements vont aussi à tous les enseignants du département d'Electronique qui ont contribué à ma formation.

Mes sincères remerciements à tous les membres du jury qui ont Acceptés de juger ce travail.

En fin, je tiens à exprimer ma reconnaissance à tous.

DAOUD YOUSSEF

SOUHEB KHIREDDINE

Dédicaces

Aux êtres les plus chers à mon cœur, ma mère et mon père, qui m'ont beaucoup soutenu durant mon cursus d'étude et à qui je serai éternellement reconnaissant, qui m'ont permis de devenir Ce que je suis aujourd'hui.

je dédie ce modeste travail à mon père et à ma mère

À toute ma famille

À mes amies

À tous les enseignants qui m'ont aidé

Enfin à tous ceux et celles qui m'ont encouragé et soutenu.

Sommaire

Sommaire

Liste de figure

Liste de tableau

Liste des symboles et abréviations

Introduction générale.....	1
----------------------------	---

CHAPITRE I : Energie renouvelable

Introduction.....	3
1. Définition de l'énergie éolienne.....	4
1.1 Descriptif d'une éolienne.....	6
1.2 Principaux composants d'une éolienne.....	8
1.3. Le principe de fonctionnement d'une éolienne.....	10
1.4. Evolution de la taille et de la puissance des éoliennes.....	11
2. L'énergie éolienne en quelques chiffre.....	12
2.1. Production mondiale d'énergie éolienne.....	12
2.2. Les plus grands producteurs d'énergie éolienne dans le monde	13
2.3. Capacité mondiale de production d'énergie éolienne dans le monde.....	13
2.4. Différents types d'éoliennes.....	14
2.5. Les raisons pour choisir une éolienne plus petite.....	15
2.5.1. Les éoliennes a axes verticaux.....	15
2.5.2. Les éoliennes à axe horizontal.....	16
3. Avantage et inconvénients de l'énergie éolienne.....	17
3.1. Les avantages.....	17
3.2. Les inconvénients.....	19
Conclusion.....	20

CHAPITRE II : Sélection de caractéristiques

Introduction.....	21
1. Réduction basée sur une sélection de caractéristiques.....	22
1.1 . Définition de la sélection.....	22
1.2. Pertinence d'une caractéristique.....	23
1.3.Caractéristiques générales des méthodes de sélection.....	23

1.3.1. Initialisation et procédures de recherche.....	25
1.3.2. Procédures d'évaluation.....	25
1.4. Critère d'arrêt.....	29
1.5. Revue de quelques méthodes de sélection.....	29
1.5.1. SFS et SBS.....	30
1.5.2. LDA (Analyse Discriminante Linéaire) :	30
1.5.3. MRMR (Max-relevance, Min-Redundancy).....	30
1.5.4. Les algorithmes génétiques.....	31
2. Réduction basée sur une transformation de données.....	31
2.1. Méthodes linéaires.....	32
2.1.1. Analyse en Composantes Principales.....	32
2.1.2. Analyse Linéaire Discriminante.....	34
2.1.3. Positionnement Multi-Dimensionnel.....	34
2.2. Méthodes non-linéaires.....	34
Conclusion.....	35

CHAPITRE III : Réseaux des neurones Artificiels

Introduction.....	36
1. Réseaux de neurones artificiels.....	36
1.1. Le neurone.....	36
1.2. Les bases biologiques.....	36
1.3. Définition les réseaux de neurones artificiels.....	37
1.4. Structure des réseaux de neurones artificiels.....	38
1.5. Utilité de réseaux de neurones artificiels.....	38
2. Le neurone formel.....	38
2.1. Définition et représentation.....	38
2.2. Poids des connections.....	41
2.3. Transformation de l'information.....	41
3. Fonction de Transfer.....	41
4. Quelques modèles des réseaux de neurones artificiels.....	42
4.1. Les réseaux multicouches.....	42
4.2. les réseaux ELM (Extreme Learning Machines).....	43
4.2.1. Modèle mathématique.....	44
4.2.2. Algorithme.....	45

4.2.3. Caractéristiques principales.....	45
4.3. Réseaux de neurones à base radiale.....	45
4.3.1. Présentation des réseaux RBF.....	45
4.3.2. Architecture générale d'un réseau RBF.....	46
4.3.3. Fonction à base radiale.....	47
4.3.4. Phase d'apprentissage.....	48
4.4. Avantages et inconvénients des réseaux RBF.....	51
5. Apprentissage.....	52
5.1. Rétro-propagation de l'erreur.....	52
5.2. Algorithmes d'optimisation.....	53
6. L'organisation des neurones en réseaux.....	54
Conclusion.....	54

CHAPITRE IV : Résultats et discussion

Introduction.....	55
1.Série temporelle.....	55
1.1. Utilisation.....	55
2. Présentation de la base de données.....	55
2.1. Caractéristiques des site.....	55
3.Application de la technique MRMR avec les réseaux de neurones type RBF et ELM.....	56
3.1. Etape de test.....	57
4. Résultats de simulation des deux sites en utilisant le logiciel MATLAB.....	57
4.1. Prédiction par la méthode de ELM.....	58
4.2. Prédiction par la méthode de RBF.....	62
Conclusion.....	66
CONCLUSION GENERAL.....	67
REFERENCES BIBLIOGRAPHIQUES	

LISTE DE FIGURE

CHAPITRE I : Energie renouvelable

Figure I-1: Estimée de la production et l'investissement dans les (ER)	3
Figure I-2 : Conversion de l'énergie cinétique du vent.....	5
Figure I-3 : Principaux organes du système de conversion éolien.....	5
Figure I-4 : Conversion de l'énergie cinétique du vent	6
Figure I-5 : Schéma d'une éolienne de type aérogénérateur.....	7
Figure I-6 : Tour d'une turbine.....	8
Figure I-7: Rotor d'une turbine.....	9
Figure I-8 : Constitution d'une nacelle.....	9
Figure I.9 : Puissance nominale moyenne des éoliennes – Allemagne.....	12
Figure I-10 : Production d'énergie éolienne dans le monde en 2013.....	14
Figure I-11 : Différents types d'éoliennes.....	16
Figure I-12 : Différents types d'éoliennes.....	17

CHAPITRE II : Sélection de caractéristiques

Figure II.1 Procédure générale d'un algorithme de sélection de caractéristiques.....	22
Figure II.2 La procédure du modèle "filtre"	26
Figure II.3 La procédure du modèle "wrapper".....	28
Figure II.4. ACP sur des données linéaires.....	32
Figure II.5. ACP sur des données non-linéaires.....	33

CHAPITRE III : Réseaux des neurones Artificiels

Figure III.1: Un neurone biologique et ses principaux composants.....	37
Figure III.2: Modèle du neurone formel.....	39
Figure III.3: Fonction de transfert : (a) du neurone «seuil» ; (b) du neurone «linéaire ´ », et (c) du neurone «sigmoïde»	42
Figure III.4 Schéma d'un réseau multicouches.....	43
Figure III.5: Illustration de la structure du réseau de neurones ELM.....	44
Figure III.6: Architecture d'un Réseau de Neurone RBF.....	47
Figure III.7: Quelques fonctions radiales.....	48

LISTE DE FIGURE

CHAPITRE IV : Résultats et discussion

Figure IV.1 : La courbe réelle avec les séries temporelles des deux sites (base de données)	56
Figure IV.2: Courbe comparaison valeurs réels et prédites de ELM pour 50 neurones site Colorado.....	58
Figure IV.3: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 50 neurones site Colorado.....	58
Figure IV.4: Courbe comparaison valeurs réels et prédites de ELM pour 100 neurones site Colorado.....	59
Figure IV.5: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 100 neurones site Colorado.....	59
Figure IV.6: Courbe comparaison valeurs réels et prédites de ELM pour 500 neurones site Colorado.....	59
Figure IV.7: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 500 neurones site Colorado.....	59
Figure IV.8: Courbe comparaison valeurs réels et prédites de ELM pour 50 neurones pour le site Connecticut.....	60
Figure IV.9: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 50 neurones site Connecticut.....	60
Figure IV.10: Courbe comparaison valeurs réels et prédites de ELM pour 100 neurones pour le site Connecticut.....	61
Figure IV.11: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 100 neurones site Connecticut.....	61
Figure IV.12: Courbe comparaison valeurs réels et prédites de ELM pour 500 neurones pour le site Connecticut.....	61
Figure IV.13: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 500 neurones site Connecticut.....	61
Figure IV.14 : Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Colorado.....	62
Figure IV.15 : Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 10 neurones pour le site Colorado.....	62
Figure IV.16: Courbe comparaison valeurs réels et prédites de RBF pour 50 neurones pour le site Colorado.....	63

LISTE DE FIGURE

Figure IV.17: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 50 neurones pour le site Colorado.....	63
Figure IV.18: Courbe comparaison valeurs réels et prédites de RBF pour 100 neurones pour le site Colorado.....	63
Figure IV.19: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 100 neurones pour le site Colorado.....	63
Figure IV.20: Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Connecticut.....	64
Figure IV.21: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 10 neurones pour le site Connecticut.....	64
Figure IV.22: Courbe comparaison valeurs réels et prédites de RBF pour 50 neurones pour le site Connecticut.....	65
Figure IV.23: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 50 neurones pour le site Connecticut.....	65
Figure IV.24: Courbe comparaison valeurs réels et prédites de ELM pour 100 neurones pour le site Connecticut.....	65
Figure IV.25: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 100 neurones pour le site Connecticut.....	65

LISTE DE TABLAEU

Liste des tableaux

CHAPITRE I : Energie renouvelable

Tableau I-1 : classification des turbines éoliennes.....	14
--	----

CHAPITRE II : Sélection de caractéristiques

Tableau II.1 : Opération pour trouver les composantes principales.....	33
--	----

CHAPITRE III : Réseaux des neurones Artificiels

Tableau III.1: analogie entre le neurone biologie et le neurone forme.....	39
Tableau III.2: Fonctions de transfert $a = f(n)$	41
Tableau III.3: Algorithme 1 Rétro-propagation élémentaire du gradient.....	53

CHAPITRE IV : Résultats et discussion

Tableau IV.1: Caractéristiques géographiques du site Colorado	56
Tableau IV.2 : Caractéristiques géographiques du site Connecticut.....	56
Tableau IV.3: valeurs réels et prédites de ELM et réduction MRMR avec site Colorado.....	58
Tableau IV.4: valeurs réels et prédites de ELM et réduction MRMR avec Site Connecticut.....	60
Tableau IV.5: valeurs réels et prédites de RBF et réduction MRMR avec Site Colorado.....	62
Tableau IV.6: valeurs réels et prédites de RBF et réduction MRMR avec site Connecticut.....	64

Notations

Les notations définies ci-dessous sont celles que nous utilisons le plus dans ce manuscrit :

E	Energie cinétique
m	La masse
v	La vitesse
P_{m t}	La puissance de l'air en mouvement
P_m	La puissance de l'air
A₂, A₁	Les sections transversales avant et après le convertisseur
F	La force exercée par le vent
v'	La vitesse du flux
C_p	Coefficient de puissance
 · , · 	Norme et valeur absolue
f	Fonction d'activation
σ	Fonction d'activation
W₁	La matrice des pondérations d'entrée-à-couche cachée
W₂	La matrice des pondérations de la couche de sortie
β	Le poids de sortie
K(x,y)	La fonction de noyau
W_{ij}	Le noyau de chaleur
I :	le nombre d'exemple soumis au réseau.
n :	le nombre de neurone.
Y:	les outputs attendus
M:	le nombre de centrioles.
N:	le nombre d'exemples dans la base d'apprentissage.

Abréviations

RNA	Réseaux Neuronaux Artificiels
ACP	L'Analyse en Composantes Principale
ACI	Analyse en Composantes Indépendantes
PMC	Perceptron Multi Couches
MLP	Multilayer perceptron
RBF	Le réseau de neurone fonctions à base radial
ELM	Extreme Learning Machines
MAPE	L'erreur Absolue Moyenne en Pourcentage (The Mean AbsolutePercentage Error).
RMSE	L'erreur Quadratique Moyenne (The Root Mean Squared Error)
R^2	Le Coefficient de Détermination
MRMR	Max-relevance, Min-Redundancy
LDA	Analyse Discriminante Linéaire
SFS	Sequential Forward Selection
SBS	Sequential Backward Selection où (sélection séquentielle arrière).
SNR	Signal-to-Noise Ratio coefficient
ÉR	énergies renouvelables
SCE	Système de Conversion Eolien
MDS	Multi Dimensional Scaling

Introduction Générale

Introduction Générale

À cause du réchauffement climatique et de la quantité limitée des stocks des énergies fossiles dans le monde, il est urgent de trouver des solutions alternatives à ces sources d'énergies. L'augmentation de la pénétration des énergies renouvelables peut contribuer à cela. Pour que les énergies renouvelables soient une alternative crédible à l'énergie fossile, il faut d'une part que leurs coûts de production soient concurrentiels et d'autre part que leur intégration aux grands réseaux à grande échelle, ainsi qu'aux réseaux décentralisés soient viables, c'est-à-dire que les gestionnaires des réseaux soient capables d'équilibrer l'offre et la demande de l'énergie. Toutefois, les périodes de production et de consommation d'énergie ne sont pas toujours les mêmes. Il devient important de trouver des systèmes de gestion et de stockage d'énergie. La connaissance de la prévision des sources d'énergies renouvelables est plus pertinente, à cause de leurs intermittences. Cette prévision devrait aider les systèmes de gestion des réseaux électriques à augmenter l'intégration des sources d'énergies renouvelables dans les réseaux.

Ce travail consiste à développer un système de prédiction des sources d'énergie renouvelable dans le but de l'intégrer à un système de gestion intelligente d'un micro réseau électrique afin de favoriser la pénétration des énergies renouvelables dans les grands réseaux électriques et ceux isolés.

La prédictibilité des données en détail est analysée afin d'évaluer la pertinence de garder les composantes qui ont un faible potentiel de prédictibilité. Cette démarche a pour but de déterminer les composants des données qui ont un apport négligeable ou négatif à la précision de la production pour permettre d'économiser des ressources de calcul.

L'application et l'intégration de la prédiction des sources d'énergies renouvelables aux réseaux électriques intelligents (Smart Grid) permettront d'optimiser la gestion des réseaux intégrant les sources d'énergies renouvelables.

Organisation du manuscrit

Dans ce mémoire, nous proposons la méthode de MRMR et les réseaux de neurones artificiels pour la prédiction de la vitesse du vent en se basant sur des modèles dites les séries temporelles.

Ce présent document est constitué de la présente introduction, de quatre chapitres et d'une conclusion générale.

Introduction Générale

Il est organisé comme suit :

- Le premier chapitre est consacré pour une présentation générale sur les énergies renouvelables et plus précisément l'énergie éolienne où on a donné une description sur l'éolienne, ses composants et le principe de son fonctionnement.
- L'objectif du deuxième chapitre est de présenter la stratégie de la réduction de la dimensionnalité qui est basée sur la sélection de caractéristiques.
- Par la suite le chapitre trois est aussi consacré pour la présentation et la description des réseaux de neurones artificiels.
- Le quatrième chapitre est le fruit de notre travail, où on donne des rappels théoriques sur les séries temporelles ainsi la méthode utilisée pour prédire la vitesse du vent, en utilisant le logiciel Matlab dans notre simulation et à la fin nous terminerons notre travail par une conclusion générale.

CHAPITRE I : Energie renouvelable

Sommaire

Introduction.....	3
1. Définition de l'énergie éolienne.....	4
1.1 descriptif d'une éolienne.....	6
1.2 principaux composants d'une éolienne.....	8
1.3. Le principe de fonctionnement d'une éolienne.....	10
1.4. Evolution de la taille et de la puissance des éoliennes.....	11
2. L'énergie éolienne en quelques chiffre.....	12
2.1. Production mondiale d'énergie éolienne.....	12
2.2. Les plus grands producteurs d'énergie éolienne dans le monde	13
2.3. La capacité mondiale de production d'énergie éolienne dans le monde.....	13
2.4. Les différents types d'éoliennes.....	14
2.5. Les raisons pour choisir une éolienne plus petite.....	15
2.5.1. Les éoliennes a axes verticaux.....	15
2.5.2. Les éoliennes à axe horizontal.....	16
3. Avantage et inconvénients de l'énergie éolienne.....	17
3.1. Les avantages.....	17
3.2. Les inconvénients.....	19
Conclusion.....	20

Introduction

Depuis quelques années, on remarque un intérêt de plus en plus grandissant pour les énergies renouvelables (ÉR) dans le monde. Les gouvernements adoptent des politiques pour favoriser la diminution du coût de nombreuses technologies d'ÉR. L'augmentation des prix des énergies fossiles, la hausse de la demande d'énergie et d'autres facteurs (réchauffement climatique, développement durable etc.) ont favorisé l'émergence des ÉR [1].

De la fin 2006 à la fin 2011, on constate une croissance très rapide des énergies renouvelables dont le photovoltaïque avec une capacité opérationnelle en augmentation moyenne de 58% par an et l'énergie éolienne de 26 % par an. ce qui se traduit par une très forte croissance des investissements dans le domaine des énergies renouvelables. de 2004 à 2011 l'investissement dans le monde est passé de 39 Milliards US à 257 Milliards US [2]. Cet investissement a été motivé principalement par la croissance des gaz à effets de serre dans l'atmosphère et aussi par l'épuisement, dans un avenir proche des stocks d' énergies fossiles dans le monde [3] .

Cette évolution est représentée dans figure I-1.

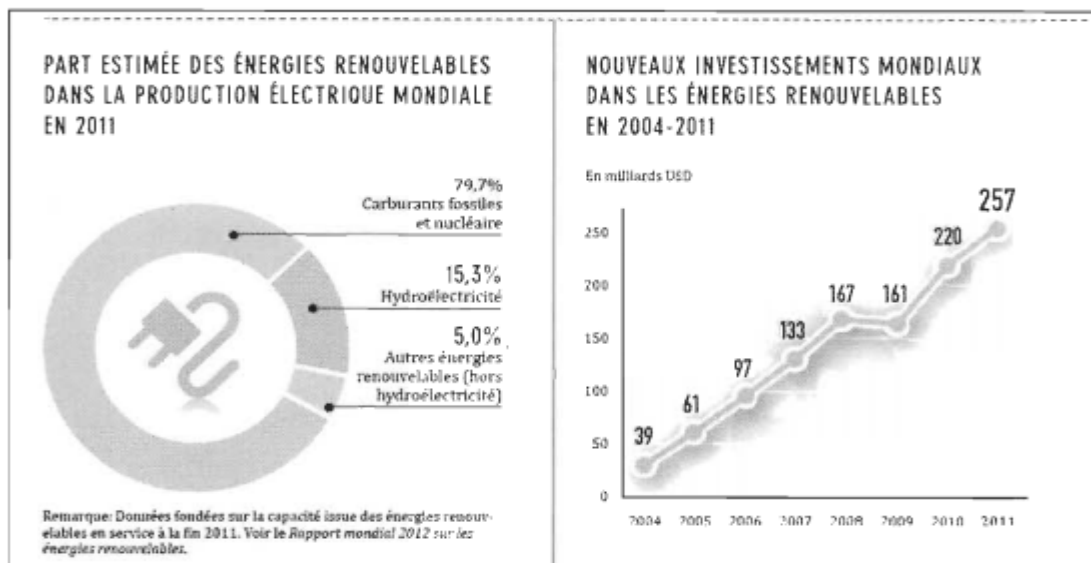


Figure I-1: Estimée de la production et l'investissement dans les ÉR [2]

Étant donné que les énergies renouvelables, tels que l'éolienne et le solaire, sont des sources intermittentes dont la production est influencée par les conditions météorologiques, leurs intégrations aux réseaux électriques et à grande échelle est un défi considérable à cause de leur caractère intermittent.

Pour que les énergies renouvelables soient une alternative crédible à l'énergie fossile il faut que leurs intégrations aux grands réseaux et à ceux décentralisés prennent une place importante dans la fourniture de l'électricité dans le monde. le défi principal des gestionnaires des réseaux électriques est d'être capable d'équilibrer l'offre et la demande de l'énergie du réseau. L'intermittence des sources comme l'éolienne et le solaire rend la gestion de cet équilibre plus ardu, lorsqu'ils représentent une part important dans les réseaux. d'où la pertinence de la prédiction des sources d'énergies renouvelables (l'éolienne et le solaire) et de la consommation de la charge. il est aussi important dans la gestion de ces réseaux de considérer l'aspect du stockage de l'énergie à cause de la dynamique de charge et de décharge des moyens de stockage.

À cause de l'intermittent des sources d'énergie renouvelable, l'application et l'intégration des techniques de prédiction aux réseaux intelligents peut être un outil très puissant pour optimiser la gestion des réseaux intégrant les sources d'énergies renouvelables intermittents.

L'énergie éolienne est une forme indirecte de l'énergie solaire. l'absorption du rayonnement solaire dans l'atmosphère engendre des différences de température et de pression qui mettent les masses d'air en mouvement , et créent le vent.

Depuis Plusieurs dizaines d'années , l'énergie éolienne connaît une croissance considérable dans le monde. Cependant , la production de cette énergie est dépendante de l'intensité et la vitesse du vent, fortement volatile, et est donc caractérisée par un haut degré d'incertitude.

C'est pour cela que la production de l'énergie éolienne est importante et qu'il est intéressant de pouvoir trouver des solutions afin de produire toujours d'énergie de manière plus propre et plus durable. cette méconnaissance de la production future d'énergie à tendance à rendre la gestion du réseau électrique plus difficile .

1. Définition de l'énergie éolienne

Une éolienne est une machine qui transforme l'énergie cinétique du vent (déplacement d'une masse d'air) en énergie mécanique ou électrique [1].

Les éoliennes sont conçues de manière à produire un maximum de puissance pour des vents de force moyennes fréquemment rencontrées.. Si le vent devient plus violent, la machine subit des contraintes plus importantes. Elle est alors freinée grâce à un système de régulation électronique qui lui permet de rester à la puissance maximale (atteinte dès 50 km/h) tout en limitant les efforts sur la structure. Au delà d'un certain seuil (90 km/h, soit 25 m/s), la

régulation ne suffit plus. La machine est alors stoppée afin de lui éviter de subir des charges trop importantes [4].

Cette énergie offre deux grands avantages, puisqu'elle est totalement propre et renouvelable. lors de son exploitation, elle n'entraîne aucun rejet (pas d'effet de serre ou de pluies acides) et aucun déchet. Mais le principal inconvénient de cette source d'énergie renouvelable est son manque de flexibilité et son inconstance. le vent ne souffle pas forcément quand on en a besoin ! en moyenne, une éolienne tourne à sa puissance nominale 1/5 du temps sur une année. il convient alors de diviser par 5 la puissance installée pour obtenir la puissance réelle moyenne .



Figure I-2 : Conversion de l'énergie cinétique du vent

Le SCE (Système de Conversion Eolien) est constitué d'un générateur électrique, entrainé par une turbine éolienne à travers le multiplicateur, d'un système de commande, d'un convertisseur statique, d'un transformateur et enfin d'un réseau électrique

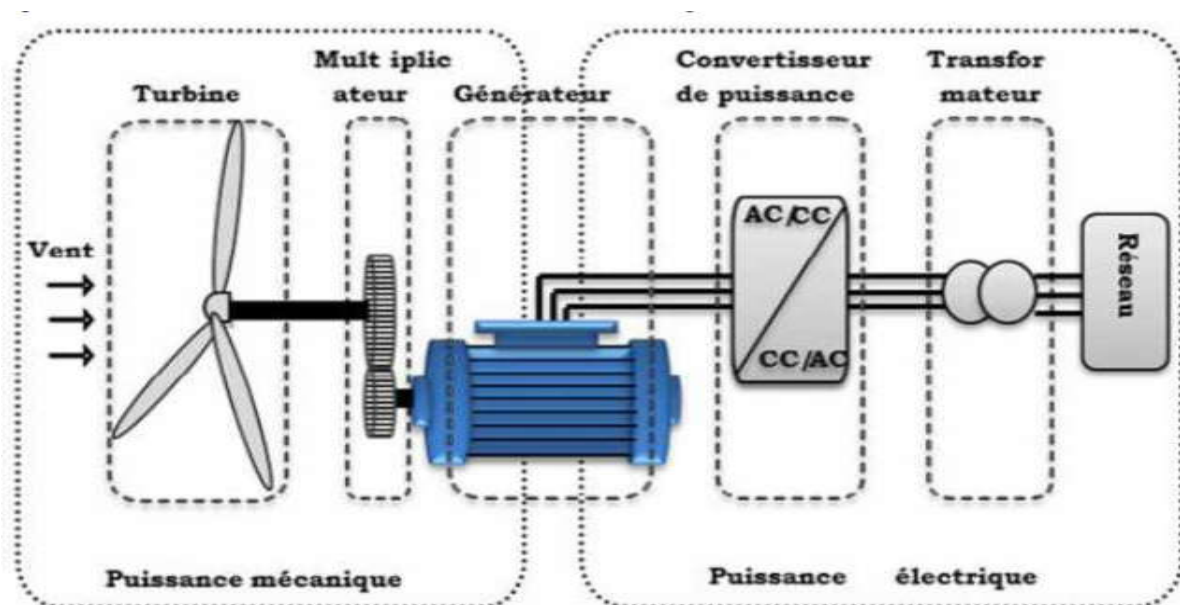


Figure I-3 : Principaux organes du système de conversion éolien

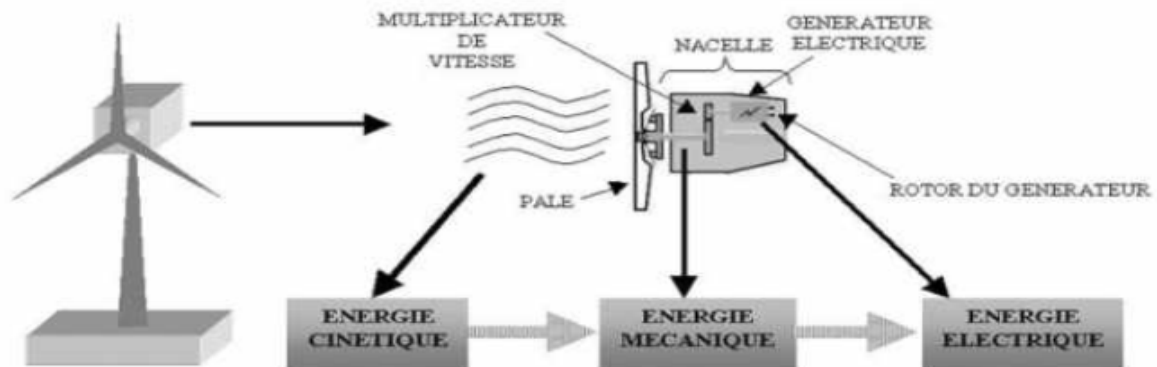


Figure I-4 : Conversion de l'énergie cinétique du vent

L'énergie éolienne est une énergie "renouvelable" non dégradée, géographiquement disponible, et surtout en corrélation saisonnière (l'énergie électrique est largement plus demandée en hiver et c'est souvent à cette période que la moyenne des vitesses des vents est la plus élevée).

De plus, c'est une énergie qui ne produit aucun rejet atmosphérique ni déchet radioactif. Elle est toutefois aléatoire dans le temps et son captage reste assez complexe, nécessitant des mâts et des pales de grandes dimensions (jusqu'à 60m pour des éoliennes de plusieurs mégawatts) dans des zones géographiquement dégagées pour éviter les phénomènes de turbulences.

L'énergie éolienne fait partie des nouveaux moyens de production d'électricité décentralisée proposant une alternative viable à l'énergie nucléaire sans pour autant prétendre la remplacer (l'ordre de grandeur de la quantité d'énergie produite étant largement plus faible). Les installations peuvent être réalisées sur terre mais également de plus en plus en mer où la présence du vent est plus régulière. de plus, les éoliennes sont ainsi moins visibles et occasionnent moins de nuisances sonores.

1.1. Descriptif d'une éolienne

Une éolienne permet de transformer l'énergie cinétique du vent en énergie électrique. L'hélice d'une éolienne entre en rotation par la force du vent et permet ainsi la production d'énergie mécanique ou électrique .

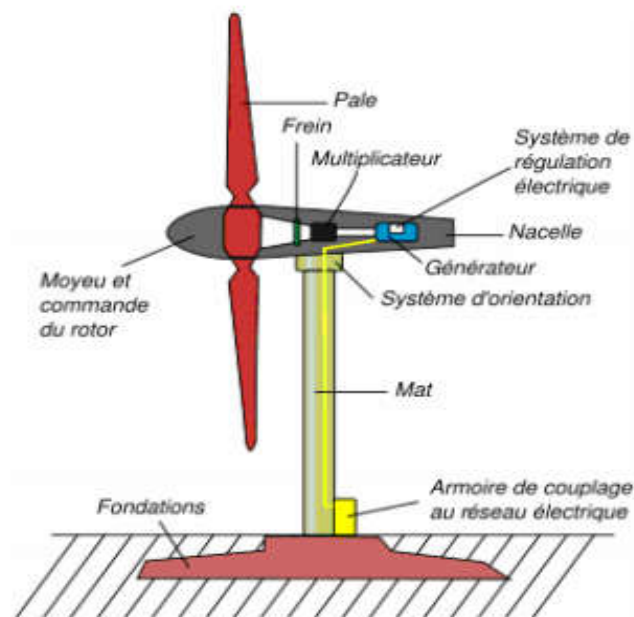


Figure I-5 : Schéma d'une éolienne de type aérogénérateur

Pour convertir l'énergie éolienne en énergie électrique le générateur va utiliser le phénomène d'induction.

En effet ce dernier est composé de deux parties, une partie mobile le rotor et une partie fixe, le stator, permettant de créer un champ magnétique et de générer un courant électrique. L'éolienne est également équipée d'une girouette permettant l'orientation des pales en fonction de la direction du vent. elle doit être également fixée solidement au sol. s'agissant du mât, il doit être dimensionné précisément en fonction de la machine, des fondations ,plusieurs systèmes existent : haubané, haubané basculant, treillis, autoporteur, pour qu'une éolienne produise ,quelques notions importantes diapositives plusieurs éléments doivent être pris avant installation.

L'éolienne doit être adaptée au site et à ses besoins, il faut vérifier le vent, on considère qu'un site est exploitable lorsque l'on dispose d'une vitesse minimale de vent de 4m/s, à une hauteur du sol de 10 m .La donnée 'vent' est essentielle à l'installation d'une éolienne, cependant le coût d'une étude précise, à l'aide d'un anémomètre peut se révéler élevé. il faut que cela reste cohérent avec le prix global de votre projet. l'observation du site (la présence de moulins, d'éoliennes altitude, vents dominants, végétation...) et l'intuition peuvent se révéler de bons indicateurs du potentiel éolien dans un premier temps.

Ce potentiel conditionnera le choix d'une machine de puissance nominale adaptée. un autre critère important reste l'évaluation de ses besoins en électricité au regard de ces

consommations journalières et annuelles. cette étape permettra également de réfléchir plus largement à maîtriser ses consommations électriques.

Pratiquement, c'est beaucoup plus complexe : les éoliennes doivent résister à des tempêtes aussi violentes que celles que nous connaissons en Bretagne. Les principaux avantages de l'énergie éolienne sont l'autonomie en électricité, la possibilité de produire de jour comme de nuit et un impact environnemental réduit par des précautions simples vis à vis de la population et du paysage. L'énergie produite peut être soit stockée dans des batteries, soit injectée aux normes dans le réseau. Cette dernière solution étant économiquement et techniquement beaucoup plus pertinente [6].

1.2. Principaux composants d'une éolienne

Il existe plusieurs configurations possibles d'aérogénérateurs qui peuvent avoir des différences importantes. Néanmoins, une éolienne "classique" est généralement constituée de:

- ❖ le mât, généralement en métal, supporte l'ensemble des équipements permettant de produire l'électricité (nacelle + rotor). Il est fixé sur une fondation implantée dans le sol, une lourde semelle en béton qui assure l'ancrage et la stabilité de l'éolienne. Le mât des éoliennes atteint aujourd'hui 80 m de haut pour les plus puissantes (exceptionnellement jusqu'à 100 m). Les éoliennes sont-elles si haut perchées c'est parce que le vent souffle plus fort à quelques dizaines de mètres de hauteur, où il n'est pas perturbé par l'effet des obstacles : relief, arbres, maisons. Et la puissance fournie par une éolienne est proportionnelle au cube de la vitesse du vent [7].



Figure I-6 : Tour d'une turbine.

La tour doit être suffisamment solide pour supporter non seulement la nacelle et le rotor, mais aussi les charges puissantes provoquées par le vent : d'une part la puissance exercée par le vent directement sur la tour, d'autre part la puissance transmise par le rotor .

- ❖ un rotor, composé de plusieurs pales (en général 3) et du nez de l'éolienne. Les pales sont aujourd'hui faites de matériaux composites à la fois légers et assurant une rigidité et une résistance suffisantes : polyester renforcé de fibre de verre et/ou fibre de carbone. Leur longueur atteinte actuellement entre 30 et 55 mètres, soit un diamètre du rotor compris entre 60 et 110 mètres. la puissance d'une éolienne est proportionnelle à la surface balayée par ses pales (un cercle), donc au carré de son diamètre rotor[8].



Figure I-7: Rotor d'une turbine

Un rotor balaye un disque circulaire au cours d'une rotation et peut donc récolter l'énergie des molécules d'air traversant ce disque. où r est le rayon du disque circulaire, d est le diamètre et π est ($\pi = 3,1415\dots$).

Le rotor est relié à la nacelle par le moyeu, elle transforme l'énergie cinétique du vent en énergie mécanique .

- ❖ une nacelle montée au sommet du mât et abritant les composants mécaniques et pneumatiques et certains composants électriques et électroniques nécessaires au fonctionnement de la machine . le transport de l'électricité produite dans la nacelle jusqu'au sol est assuré par des câbles électriques descendant à l'intérieur du mât de l'éolienne.

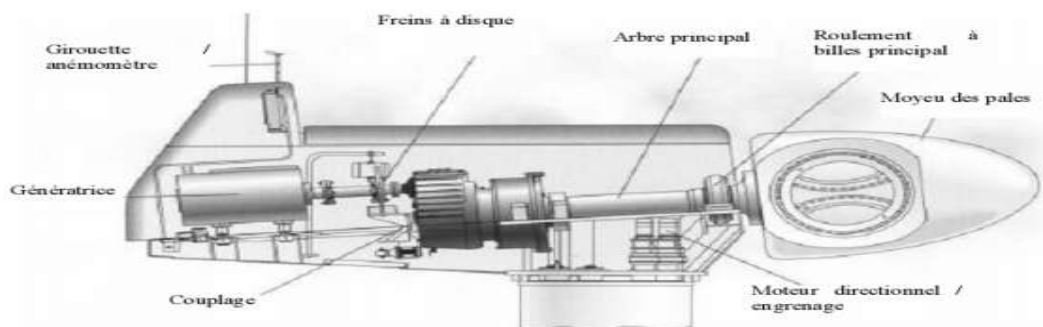


Figure I-8 : Constitution d'une nacelle

- ❖ Le multiplicateur de vitesse : il sert à élever la vitesse de rotation entre l'arbre primaire et l'arbre secondaire qui entraîne la génératrice électrique.
- ❖ L'arbre secondaire comporte généralement un frein mécanique qui permet d'immobiliser le rotor au cours des opérations de maintenance et d'éviter l'emballement de la machine.
- ❖ La génératrice : c'est elle qui convertit l'énergie mécanique en énergie électrique.
- ❖ Un contrôleur électronique chargé de surveiller le fonctionnement de l'éolienne. il s'agit en fait d'un ordinateur qui peut gérer le démarrage de la machine lorsque la vitesse du vent est suffisante (de l'ordre de 5 m/s), gérer le pas des pales, le freinage de la machine, l'orientation de l'ensemble « rotor plus nacelle » face au vent de manière à maximiser la récupération d'énergie. Pour mener à bien ces différentes tâches, le contrôleur utilise les données fournies par un anémomètre (vitesse du vent) et une girouette (direction du vent), habituellement situés à l'arrière de la nacelle.

Enfin, le contrôleur assure également la gestion des différentes pannes éventuelles pouvant survenir.

- ❖ Divers dispositifs de refroidissement (génératrice, multiplicateur) par ventilateurs, radiateurs d'eau ou d'huile.

Grâce à un système de supervision et contrôle d'une éolienne peut être arrêtée automatiquement et très rapidement en cas de nécessité. la sécurité du fonctionnement des éoliennes est ainsi assurée en continu.

- ❖ dans le cas des éoliennes produisant de l'électricité, un poste de livraison situé à proximité du parc éolien permet de relier ce parc au réseau électrique pour y injecter l'intégralité de l'énergie produite [9].

1.3. Principe de fonctionnement d'une éolienne

Sous l'effet du vent, le rotor tourne. Dans la nacelle, l'arbre principal entraîne un alternateur qui produit l'électricité. La vitesse de rotation du rotor (de 12 à 15 tours/minute) doit être augmentée par un multiplicateur de vitesse jusqu'à environ 1500 tours/minute, vitesse nécessaire au bon fonctionnement de l'alternateur. des convertisseurs électroniques de puissance ajustent la fréquence du courant produit par l'éolienne à celle du réseau électrique auquel elle est raccordée (50 Hz en Europe), tout en permettant au rotor de l'éolienne de tourner à vitesse variable en fonction du vent.

La tension de l'électricité produite par l'alternateur, de l'ordre de 600 à 1000 volts, est ensuite élevée à travers un transformateur de puissance, situé dans la nacelle ou à l'intérieur

du mât, jusqu'à un niveau de 20 ou 30 KV. ce niveau de tension permet de véhiculer l'électricité produite par chacune des éoliennes d'une centrale éolienne jusqu'au point de raccordement au réseau électrique public (en France, le réseau EDF). la tension de l'électricité produite par la centrale peut alors être de nouveau transformée, en fonction du niveau de tension de raccordement de la centrale au réseau public. Pour les centrales éoliennes de 10 à 15 MW de capacité, le niveau de tension de raccordement est, en France, généralement de 20 KV. Pour les centrales de capacité plus importante, le niveau de tension de raccordement peut aller de 60 à 90 KV, voire même 225 KV.

Pour pouvoir démarrer, une éolienne a besoin d'une vitesse de vent minimale, de l'ordre de 10 à 15 km/h. Et au-delà de 90 km/h, les turbines s'arrêtent de tourner. Tout d'abord, la fréquence d'occurrence des vents d'une vitesse supérieure à 90 km/h est généralement faible (inférieure à 1 %), et si les éoliennes fonctionnaient dans ces conditions, elles subiraient des efforts importants qui entraîneraient une usure prématurée de leurs équipements.

Compte tenu du faible gain relatif sur la production que représente un fonctionnement par vent fort, les ingénieurs préfèrent, dans ces conditions, stopper les machines et attendre le retour de vents plus modérés et plus réguliers. Si les éoliennes ne fonctionnent pas au-delà d'une vitesse de vent de 90 km/h, leurs fondations n'en sont pas moins conçues pour résister à des vents beaucoup plus importants... la puissance d'une éolienne classique est de 1 à 1,5 MW, mais les éoliennes de la nouvelle génération atteignent 2 à 3 MW et des modèles de 5 MW sont d'ores et déjà testés par les constructeurs [10].

1.4. Evolution de la taille et de la puissance des éoliennes

Les premières éoliennes, dans les années 1980, étaient plutôt petites (15 à 20 m de diamètre pour une puissance de 50 à 100 kW) mais la puissance des éoliennes en exploitation commerciale a augmenté régulièrement. La figure 1.2 représente la puissance moyenne des éoliennes installées en Allemagne de 2000 à 2012. Au cours de cette même période, la puissance nominale moyenne des éoliennes a plus que doublé, passant de 1 114 kW à 2 419 kW.

En même temps que la puissance nominale des éoliennes augmentait, les moyeux s'élevaient. en règle générale, la hauteur du moyeu est égale au diamètre du rotor, mais de nombreux fabricants proposent des tours plus élevées en option pour permettre une production plus importante. plusieurs éoliennes d'un diamètre d'environ 70 m ont un moyeu à 100 m de hauteur.

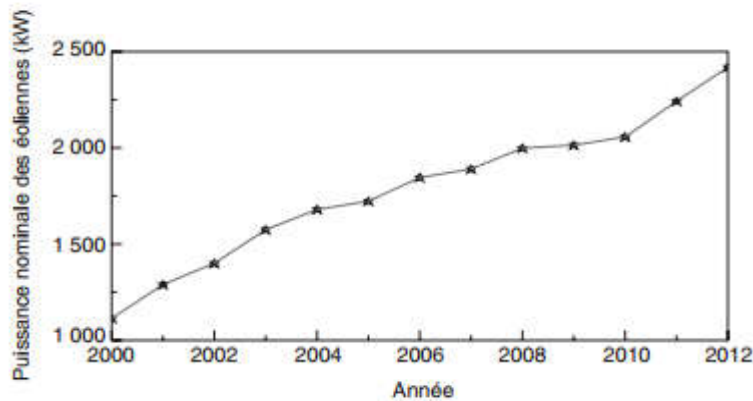


Figure I.9 : Puissance nominale moyenne des éoliennes – Allemagne

Une autre façon d'augmenter le rendement énergétique est d'augmenter le taux d'utilisation du générateur. Mais étant donné que les vitesses de vent capables de générer la puissance de sortie maximale ne sont disponibles que quelques heures par an, il n'y a aucun intérêt économique à installer des générateurs à grande puissance. Pour tenir compte des pressions économiques qu'ils subissent, la plupart des fournisseurs d'éoliennes ont fixé à environ 400 W/m² les puissances nominales maximales des rotors.

Une éolienne de 40 m de diamètre commercialisée au début des années 1990 avait donc une puissance d'environ 500 kW. Pour une vitesse moyenne de vent de 7,5 m/s et avec une distribution normale des vitesses de vent, elle pouvait atteindre sa puissance maximale en sortie pendant environ 8 % de l'année. La puissance nominale des éoliennes a cependant augmenté, tendant vers 600 W/m² pour les éoliennes les plus puissantes, car la concurrence est de plus en plus forte.

Les éoliennes de forte puissance ne conviennent pas aux sites à vitesses de vent faibles. Pour ce genre de sites, les fabricants proposent des éoliennes à grand rotor et puissance nominale modeste. Ainsi, le fabricant d'éoliennes Vestas propose une machine de 2 MW avec rotor de 110 m de diamètre pour les sites à vitesses de vent faibles, une machine de même puissance mais avec rotor de 80 m de diamètre pour les sites à vitesses de vent élevées, et diverses autres options.

2. L'énergie éolienne en quelques chiffres

2.1. Production mondiale d'énergie éolienne

La capacité éolienne installée dans le monde est de 318 GW fin 2013, et devrait augmenter de plus de 400% et atteindre 1.107 GW en 2030. la puissance éolienne installée dans le monde devrait être multipliée par 3 pour atteindre 910 GW en 2025 et le marché mondial de l'éolien devrait plus que doubler au cours de cette période, passant de 21 à 47 milliards €[11] .

2.2. Les plus grands producteurs d'énergie éolienne dans le monde

L'entreprise chinoise Xinjiang Glodwin Science and Technology a été à la tête des producteurs mondiaux d'énergie éolienne en 2015. Le groupe chinois a installé 7.8 Gigawatts de turbines éoliennes, toutes en Chine, un bond lié à l'essor du secteur impulsé par le gouvernement du pays. Le danois Vestas occupe à la deuxième place, avec 7.3 GW installés, 2.5 qu'en 2014. La compagnie a élargie son marché, avec des projets dans un total de 32 pays.

General Electric est descendu à la troisième position reste la compagnie du secteur la plus puissante des Etats-Unis, qui compte 62% des installations mondiales faites par des groupes américains. GE a installé 5.9 gigawatts, 700 mégawatts plus que l'année précédent. Le quatrième poste du top 10 est partagé entre l'allemande Siemens et l'espagnole Gamesa, car les deux compagnies ont installé 3.1GW en 2015.

Une autre entreprise allemande, Enercon, occupe la sixième position, et reste le leader du pays avec 48% des turbines éoliennes installés. l'étude a été élaborée par Bloomberg New Energy Finance, qui souligne que c'est la première fois qu'un constructeur chinois occupe ce rang de leader sur ce marché. En 2014, Goldwind était quatrième avec 4,5 GW installés.

La totalité de ses commandes ont été enregistrées sur le marché chinois, alors que la Chine a représenté en 2015 la moitié du marché mondial. quatre fabricants chinois apparaissent d'ailleurs dans la queue du top 10 : Guodian, Ming Yang, En vision et CSIC.

2.3. Capacité mondiale de production d'énergie éolienne dans le monde

Les chiffres les plus actuelles concernant la capacité mondiale de production d'énergie éolienne datent de 2013 C'est 318.137 MW , En 2015, la capacité mondiale augmenta de 12,5% et de 10,3% en Europe pour atteindre 318.137 MW fin 2013 (soit +35.467 MW en 2013) et l'équivalent de 227 réacteurs nucléaires EPR ! (source GWEC -conseil Mondial pour l'Energie Eolienne).

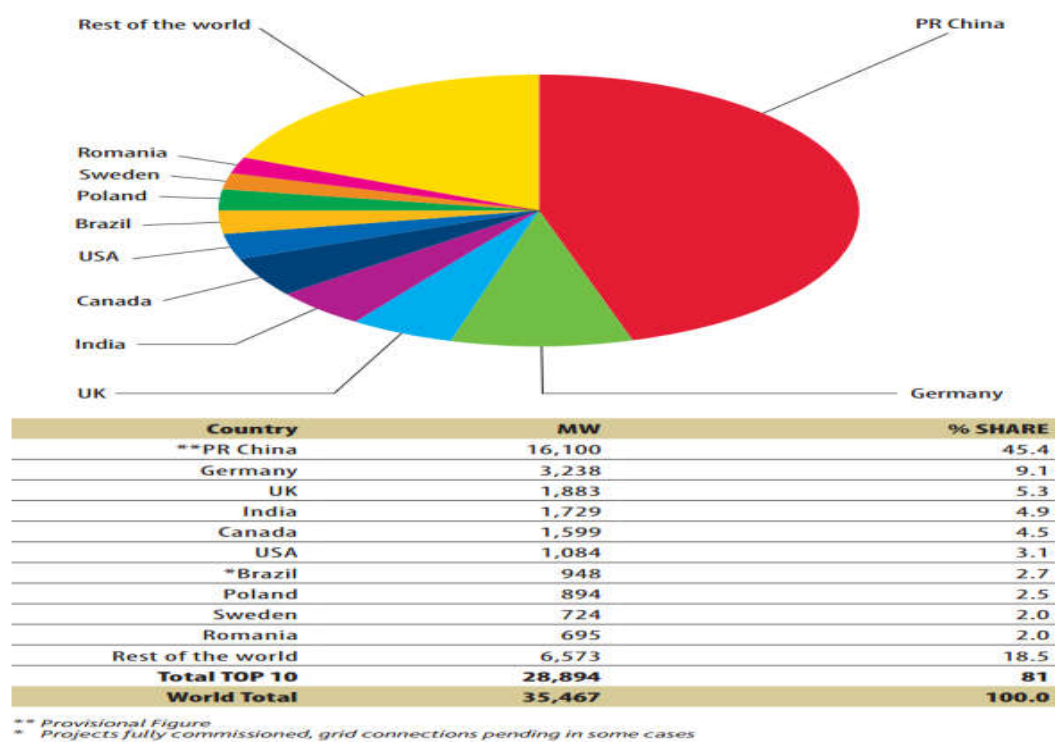


Figure I-10 : Production d'énergie éolienne dans le monde en 2013

2.4. Différents types d'éoliennes

Les solutions techniques permettant de recueillir l'énergie du vent sont très variées. le tableau I-2 présente une classification des turbines éoliennes.

ECHELLE	DIMÈTRE DE L'HÉLICE	PUISSANCE DELIVRÉE
Petite	Moins de 12 m	Moins de 40 KW
Moyenne	12 m à 45 m	De 40 KW à 1 MW
Grande	46 m et plus	1 MW et plus

Tableau I-1 : classification des turbines éoliennes .

Les raisons pour choisir une grande éolienne :

1. Le principe des économies d'échelle vaut évidemment également pour les éoliennes. Ainsi, une grande éolienne produit normalement de l'électricité à un moindre coût qu'une petite. la raison pour cela est que les coûts de fondations, de construction, de raccordement au réseau et d'autres composants de l'éolienne (le système contrôle commande, p.ex.) sont plus ou moins les mêmes, quelque soit la taille de l'éolienne.

2. Les grandes éoliennes sont particulièrement appropriées à l'installation en mer. le coût

des fondations n'augmente pas proportionnellement avec la taille de l'éolienne, et les coûts d'entretien sont dans une large mesure indépendants de la taille.

3. Dans les zones où il est difficile de trouver des sites pour plus qu'une seule éolienne, une grande éolienne avec une tour haute tire mieux partie de la ressource éolienne qu'une petite.

2.5. Les raisons pour choisir une éolienne plus petite

1. Il arrive que le réseau électrique local soit trop faible pour supporter la production électrique d'une grande éolienne. c'est souvent le cas dans les parties les plus extérieures du réseau où la densité de la population et les besoins en électricité est très basse.

2. La production d'électricité est moins fluctuante dans un parc éolien composé de plusieurs petites éoliennes, étant donné que les variations du vent sont aléatoires, ayant donc tendance à s'annuler. et en plus, comme déjà mentionné, le choix d'éoliennes plutôt petites peut se révéler avantageux dans un réseau électrique faible.

3. Les coûts liés à l'usage de très grandes grues et à la construction de chemins suffisamment robustes pour supporter le transport des composants de l'éolienne constituent un autre facteur qui, dans certains endroits, rend plus économique le choix de petites éoliennes.

4. Avec plusieurs éoliennes d'une moindre puissance, on assure la répartition du risque en cas de défaillance temporaire d'une éolienne (p.ex. par suite d'une foudre).

Des considérations esthétiques du paysage peuvent parfois dicter le choix d'éoliennes plus petites. cependant, il faut savoir, que la vitesse de rotation d'un grand rotor est en général beaucoup moins rapide que celles d'un petit, ce qui a pour résultat qu'une seule grande éolienne attire souvent moins l'attention que plusieurs petites.

On distingue aussi le "grand éolien", qui concerne les machines de plus de 250 KW, de l'éolien de moyenne puissance (entre 36 KW et 250 KW) et du petit éolien (inférieur à 36 KW) [1].

Les éoliennes se divisent en deux grandes familles : celles à axe vertical et celles à axe horizontal:

2.5.1. Les éoliennes à axes verticaux

Les éoliennes à axe vertical ont été les premières structures développées pour produire de l'électricité paradoxalement en contradiction avec le traditionnel moulin à vent à axe horizontal. Elles possèdent l'avantage d'avoir les organes de commande et le générateur au niveau du sol donc facilement accessibles et il s'agit d'une turbine à axe vertical de forme cylindrique qui peut facilement être installée sur le toit d'une maison moderne et dont les

avantages sont : faible impact visuel, pratiquement pas de bruit et très grande tolérance aux vents forts [12].

Il existe des systèmes grâce auxquels les ailes se décalent plus ou moins pour augmenter l'étendue des vitesses d'action. Si la vitesse du vent est basse, les ailes sont complètement déployées, si la vitesse est trop forte, les ailes sont complètement fermées et l'éolienne forme un cylindre.

Même si quelques grands projets industriels ont été réalisés, les éoliennes à axe vertical restent toutefois marginales et peu utilisées voire actuellement abandonnées.

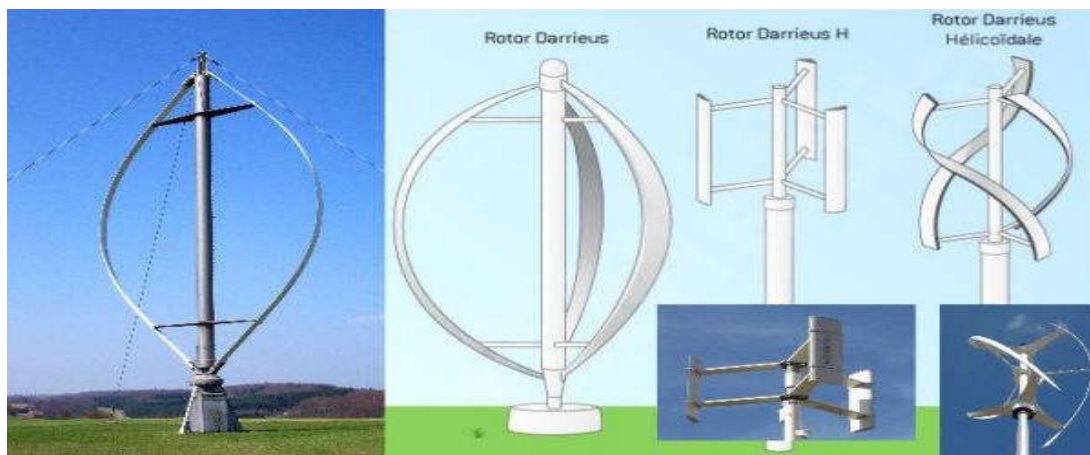


Figure I-11 : Différents types d'éoliennes.

2.5.2. Les éoliennes à axe horizontal

Les éoliennes à axe horizontal sont basées sur la technologie ancestrale des moulins à vent. Elles sont constituées de plusieurs pales profilées aérodynamiquement à la manière des ailes d'avion. Dans ce cas, la portance n'est pas utilisée pour maintenir un avion en vol mais pour générer un couple moteur entraînant la rotation.

Le nombre de pales utilisé pour la production d'électricité varie classiquement entre 1 et 3, le rotor tripale étant le plus utilisé car il constitue un compromis entre le coefficient de puissance, le coût et la vitesse de rotation du capteur éolien, ce type d'éolienne a pris le dessus sur celles à axe vertical car elles représentent un coût moins important, elles sont moins exposées aux contraintes mécaniques et la position du récepteur à plusieurs dizaines de mètres du sol privilégie l'efficacité.

Notons cependant que certains travaux, défendent la viabilité du rotor vertical en réalisant des études multicritères. Les concepts abordés dans la suite de cette étude se limiteront uniquement au cas des éoliennes à axe horizontal [12].

Aujourd'hui, pratiquement les seules éoliennes commerciales sont à axe horizontales. Les plus grandes éoliennes mesurent jusqu'à 180m en bout de pale avec un moyeu à 120m pour une puissance de 6 MW.

Les éoliennes à axe vertical ont été prometteuses dans les années 80 et au début des années 90, mais leur faible rendement aérodynamique ainsi que les fluctuations élevées de la puissance électrique générée les ont écartées du marché .[13]

Dans ce contexte, quelques notions principales au sujet de la technologie éolienne ont été données comme les méthodes de description du gisement éolien, des exemples d'architectures, les génératrices utilisées, etc. Nos travaux portent sur un système éolien de petite puissance dédié à un site perturbé comme le milieu urbain, des exemples de réalisations pratiques et industrielles de cette catégorie sont proposés.

La turbine à axe vertical de type Savonius qui semble être particulièrement adaptée est alors étudiée du point de vue de son rendement en collaboration avec l'Institut de Mécanique des Fluides de Toulouse donnant des indices sur les différents perfectionnements du rendement possibles.



Figure I-12 : Différents types d'éoliennes.

3. Avantage et inconvénients de l'énergie éolienne

La croissance de l'énergie éolienne est évidemment liée aux avantages de l'utilisation de ce type d'énergie. Cette source d'énergie a également des désavantages qu'il faut étudier, afin que ceux-ci ne deviennent pas un frein à son développement.

3.1. Les avantages

❖ L'énergie éolienne, propre, fiable, économique, et écologique, c'est une énergie qui

respecte l'environnement .

❖ Bien que ne pouvoir envisager de remplacer totalement les sources traditionnelles d'énergie, l'énergie éolienne peut toutefois proposer une alternative intéressante et renouvelable.

Elle s'inscrit parfaitement dans l'effort global de réductions des émissions de CO₂, etc..

- ❖ L'énergie éolienne est une énergie renouvelable propre, gratuit, et inépuisable .
- ❖ Chaque mégawatheure d'électricité produit par l'énergie éolienne aide à réduire de 0,8 à 0,9 tonne les émissions de CO₂ rejetées chaque année par la production d'électricité d'origine thermique .
- ❖ Parmi toutes les sources de production d'électricité, celle d'origine éolienne subit de très loin le plus fort taux de croissance .
- ❖ L'énergie éolienne n'est pas non plus une énergie à risque comme l'énergie nucléaire et ne produit pas de déchets toxiques ou radioactifs .
- ❖ L'exploitation de l'énergie éolienne n'est pas un procédé continu puisque les éoliennes en fonctionnement peuvent facilement être arrêtées, contrairement aux procédés continus de la plupart des centrales thermiques et des centrales nucléaires. Ceux-ci fournissent de l'énergie même lorsque que l'on n'en a pas besoin, entraînant ainsi d'importantes pertes et par conséquent un mauvais rendement énergétique [14].
- ❖ Les parcs éoliens se démontent très facilement et ne laissent pas de trace .
- ❖ C'est une source d'énergie locale qui répond aux besoins locaux en énergie. Ainsi les pertes en lignes dues aux longs transports d'énergie sont moindres. Cette source d'énergie peut de plus stimuler l'économie locale, notamment dans les zones rurales.
- ❖ La durée de vie des éoliennes modernes est maintenant de 20 à 25 ans, ce qui est comparable à de nombreuses autres technologies de production d'énergie conventionnelles
- ❖ C'est l'énergie la moins chère entre les énergies renouvelables , selon l'article le coût de l'éolienne à diminuer presque 90% depuis le début des années 80. Le coût de l'énergie éolienne continue de diminuer grâce aux percées technologiques, à l'accroissement du niveau de production et à l'utilisation de grandes turbines .
- ❖ Cette source d'énergie est également très intéressante pour les pays en voie de développement.

Elle répond au besoin urgent d'énergie qu'ont ces pays pour se développer. L'installation d'un parc ou d'une turbine éolienne est relativement simple. Le coût d'investissement

nécessaire est faible par rapport à des énergies plus traditionnelles, ce type d'énergie est facilement intégré dans un système électrique existant déjà [13].

- ❖ L'énergie éolienne se révèle une excellente ressource d'appoint d'autres énergies, notamment durant les pics de consommation, en hiver par exemple.

3.2. Les inconvénients

Même s'ils ne sont pas nombreux, l'éolien a quelques désavantages :

- ❖ L'impact visuel : ca reste néanmoins un thème subjectif . des images de synthèse sont élaborées pour montrer l'impact visuel. dans la plus grande majorité des cas, les enquêtes réalisées montrent une réelle acceptation des populations voisines ou visitant.
- ❖ Les bruits mécaniques ou aérodynamiques ont été réduits par l'utilisation de nouveaux profils , extrémités de pale, mécanismes de transmission etc. et ne sont plus une gêne, même proche des machines (50-60 dB équivalent à une conversation) [15].

Une distance d'environ huit fois le diamètre permet de ne plus distinguer aucun bruit lié à cette activité (< 40 dB). De plus, il faut souligner que le bruit naturel du vent, pour des vitesses supérieures à 8 m/s, a tendance à masquer le bruit rayonné par l'éolienne .

- ❖ Les éoliennes peuvent nuire à la migration des oiseaux en étant un obstacle mortel. en effet, les pales en rotation sont difficilement visibles par mauvais temps ou la nuit. Les oiseaux peuvent alors entrer en collision avec celles-ci. Plus le parc éolien est dense plus ce risque est grand. Des lumières sur les pales peuvent réduire ce danger.

Cependant, aucune étude sérieuse ne semble actuellement avoir démontré la réalité du danger pour les oiseaux.

- ❖ La source d'énergie éolienne étant stochastique , la puissance électrique produite par les aérogénérateurs n'est pas constante. la qualité de la puissance produite n'est donc pas toujours très bonne. Jusqu'à présent, le pourcentage de ce type d'énergie dans le réseau était faible, mais avec le développement de l'éolien, notamment dans les régions à fort potentiel de vent, ce pourcentage n'est plus négligeable. ainsi, l'influence de la qualité de la puissance produite par les aérogénérateurs augmente et par suit, les contraintes des gérants du réseau électrique sont de plus en plus strictes.
- ❖ Les systèmes éoliens coûtent généralement plus cher à l'achat que les systèmes utilisant des sources d'énergie classiques, comme les groupes électrogènes à essence, mais à long terme, ils constituent une source d'énergie économique et ils demandent peu d'entretien.

- ❖ Selon l'article de , il a fallu plusieurs décennies pour réaliser des éoliennes silencieuses, esthétiques et résistantes aux conditions météorologiques très capricieuses. [13].

Conclusion

Dans ce chapitre une brève description sur les systèmes éoliens a été présentée. comme beaucoup d'énergies renouvelables, le gisement éolien mondial est considérable puisqu'il est très supérieur à la consommation énergétique actuelle de l'humanité. cependant le développement de son exploitation dépendra, non pas des difficultés technologiques surmontées actuellement côté pratique, mais de données économiques et politiques favorisant ou non les diverses formes d'énergies exploitables.

Et on peut dire aussi que l'éolienne est une source de production d'énergie qui représente dans certains cas l'une des meilleurs solutions adaptée. et ne consomme aucun combustible et ne participe pas à l'effet de serre.

Vue l'importance de l'énergie éolienne on s'intéressera dans les chapitres suivants à la conception et la simulation d'un modèle d'une éolienne avec une génératrice synchrone à aimant permanent.

CHAPITRE II : Sélection de caractéristiques

Sommaire

Introduction.....	21
1. Réduction basée sur une sélection de caractéristiques.....	22
1.1 . Définition de la sélection.....	22
1.2.La pertinence d'une caractéristique.....	23
1.3.Caractéristiques générales des méthodes de sélection.....	23
1.3.1. Initialisation et procédures de recherche.....	25
1.3.2. Procédures d'évaluation.....	25
1.4. Critère d'arrêt.....	29
1.5. Revue de quelques méthodes de sélection.....	29
1.5.1. SFS et SBS.....	30
1.5.2. LDA (Analyse Discriminante Linéaire) :	30
1.5. 3. MRMR (Max-relevance, Min-Redundancy).....	30
1.5.4. Les algorithmes génétiques.....	31
2. Réduction basée sur une transformation de données.....	31
2.1. Méthodes linéaires.....	32
2.1.1. Analyse en Composantes Principales.....	32
2.1.2. Analyse Linéaire Discriminante.....	34
2.1.3. Positionnement Multi-Dimensionnel.....	34
2.2. Méthodes non-linéaires.....	34
Conclusion.....	35

Introduction

Dans de nombreux domaines (vision par ordinateur, reconnaissance de formes,...etc.), la résolution des problèmes se base sur le traitement de données extraites à partir des données acquises dans le monde réel, et structurées sous forme de vecteurs. la qualité du système de traitement dépend directement du bon choix du contenu de ces vecteurs. mais dans de nombreux cas, la résolution pratique du problème devient presque impossible à cause de la dimensionnalité trop importante de ces vecteurs. par conséquent, il est souvent utile, et parfois nécessaire, de réduire celle-ci à une taille plus compatible avec les méthodes de résolution, même si cette réduction peut conduire à une légère perte d'informations. parfois, la résolution de phénomènes complexes avec des descripteurs de grande taille pourrait être gérée en utilisant peu de caractéristiques extraites des données initiales, il suffit qu'elles représentent les variables pertinentes pour le problème à résoudre.

Une méthode de réduction de la dimensionnalité est souvent définie comme un processus de prétraitement de données qui permet de supprimer les informations redondantes et bruitées. Avec l'accroissement de la quantité de données, mises à disposition, la redondance et le bruit dans les informations sont toujours présents. cette multiplication de données n'est pas sans introduire de bruit qui vient complexifier la résolution du problème.

Les méthodes de réduction de la dimensionnalité sont généralement classées en deux catégories :

- L'extraction de caractéristiques qui permet de créer de nouveaux ensembles de caractéristiques, en utilisant une combinaison des caractéristiques de l'espace de départ ou plus généralement une transformation effectuant une réduction du nombre de dimensions.
- La sélection de caractéristiques qui regroupe les algorithmes permettant de sélectionner un sous-ensemble de caractéristiques parmi un ensemble de départ, en utilisant divers critères et différentes méthodes.

L'approche par sélection permet de mieux appréhender la modélisation d'un problème et de limiter les mesures qui permettent la résolution du problème, par contre l'extraction de nouvelles caractéristiques conserve une vision globale des observations et ne permet pas d'économiser des mesures pour décrire le phénomène observé. dans ce mémoire, nous nous intéressons aux techniques de sélection de caractéristiques.

Comme mentionné précédemment, la sélection de caractéristiques est une technique permettant de choisir les caractéristiques, variables ou mesures les plus intéressantes, pertinentes, adaptées à un système de résolution d'un problème particulier. la difficulté des

problèmes à résoudre et la masse des données disponibles conduisent à la complexification des systèmes. une phase de sélection constitue alors un module important qui est intégré au système complexe.

1. Réduction basée sur une sélection de caractéristiques

Après avoir précisé l'objectif de cette approche, défini la pertinence d'une caractéristique et les différentes étapes qui interviennent classiquement dans les systèmes reposant sur un tel principe, nous présenterons quelques méthodes développées dans la littérature et que nous avons choisies en fonction de leur représentativité dans le domaine.

1.1. Définition de la sélection

La sélection de caractéristiques est généralement définie comme un processus de recherche permettant de trouver un sous-ensemble "pertinent" de caractéristiques parmi celles de l'ensemble de départ. la notion de pertinence d'un sous-ensemble de caractéristiques dépend toujours des objectifs et des critères du système. en général, le problème de sélection de caractéristiques peut être défini par :

Soit $F = \{f_1, f_2, \dots, f_N\}$ un ensemble de caractéristiques de taille N , ou N représente le nombre total de caractéristiques étudiées. soit E_v une fonction qui permet d'évaluer un sous-ensemble de caractéristiques. nous supposons que la plus grande valeur de E_v soit obtenue pour le meilleur sous-ensemble de caractéristiques. l'objectif de la sélection est de trouver un sous-ensemble F' ($F' \subseteq F$) de taille N' ($N' \leq N$) tel que :

$$E_v(F') = \max_{Z \subseteq F} E_v(Z) \tag{II.1}$$

Ou $|Z| = N'$ et N' est, soit un nombre prédéfini par l'utilisateur ou soit contrôlé par une des méthodes de génération de sous-ensembles que nous décrirons dans la section II.1 Une procédure générale proposée par (Dash et Liu [1997]) pour une méthode de sélection de caractéristiques est illustrée par la figure II.1 [16].

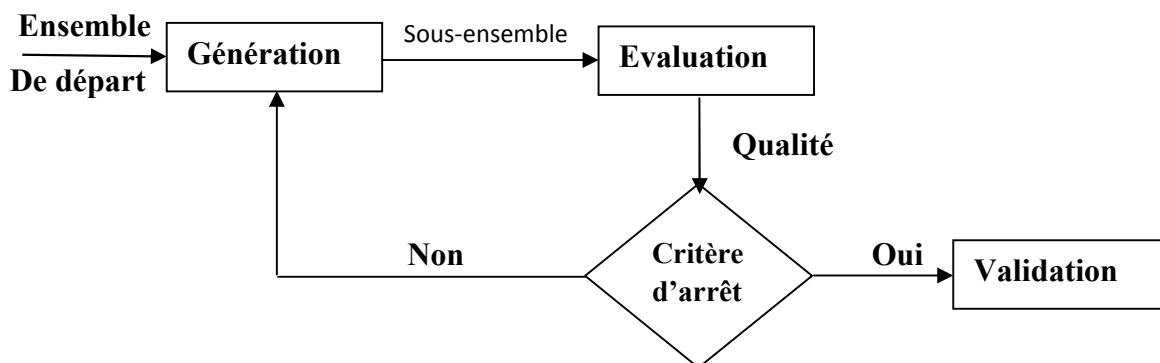


Figure II.1: Procédure générale d'un algorithme de sélection de caractéristiques[16]

Il existe trois types de stratégies de sélection de caractéristiques :

Dans la première stratégie, la taille du sous-ensemble à sélectionner (N' par exemple) est prédéfinie et l'algorithme de sélection cherche à trouver le meilleur sous-ensemble de cette taille.

La deuxième stratégie consiste à sélectionner le plus petit sous-ensemble dont la performance est plus grande ou égale à un seuil prédéfini.

La troisième stratégie cherche à trouver un compromis entre l'amélioration de la performance (l'erreur de classification par exemple) et la réduction de la taille du sous-ensemble.

Le but est de sélectionner le sous-ensemble qui optimise les deux objectifs en même temps.

1.2. Pertinence d'une caractéristique

La performance d'un algorithme d'apprentissage dépend fortement des caractéristiques utilisées dans la tâche d'apprentissage. la présence de caractéristiques redondantes ou non pertinentes peut réduire cette performance. dans la littérature, il existe plusieurs définitions de la pertinence d'une caractéristique, la plus connue est celle de (John et al. [1994], John [1997]). Selon cette définition, une caractéristique est classée comme étant très pertinente, peu pertinente et non pertinente[17].

Très pertinente : Une caractéristique f_i est dite très pertinente si son absence entraîne une détérioration significative de la performance du système de classification utilisé.

Peu pertinente : Une caractéristique f_i est dite peu pertinente si elle n'est pas "très pertinente" et s'il existe un sous-ensemble V tel que la performance de $V \cup \{f_i\}$ soit significativement meilleure que la performance de V .

Non pertinente : Les caractéristiques qui ne sont ni "peu pertinentes" ni "très pertinentes" représentent les caractéristiques non pertinentes. ces caractéristiques seront en général supprimées de l'ensemble de caractéristiques de départ.

1.3. Caractéristiques générales des méthodes de sélection

Une méthode de sélection qui cherche à maximiser la fonction (E_v) de l'équation (II.1) passe généralement par quatre étapes (Liu et Yu [2005]) :

Les deux premières consistent à initialiser le point de départ à partir duquel la recherche va commencer et à définir une procédure de recherche ou une procédure de génération de sous-ensemble de caractéristiques. une fois la stratégie de recherche définie, et les sous-ensembles générés, une méthode d'évaluation est définie dans la troisième étape.

Les étapes deux et trois se répètent jusqu' à un critère d'arrêt. ce test d'arrêt représente la quatrième étape de la méthode. les quatre étapes sont détaillées dans les sections suivantes.

1.3.1. Initialisation et procédures de recherche

La première question que l'on peut se poser avant d'appliquer la procédure de recherche est : "Sur quel point de l'espace de caractéristiques la recherche peut-elle commencer ?" Pour répondre à cette question, il est nécessaire de définir un point de départ (ou direction de recherche). par exemple, une recherche peut commencer par une ensemble vide de caractéristiques, et continuer par l'ajout successif, à chaque itération, d'une ou plusieurs caractéristiques. inversement, la recherche peut commencer avec l'ensemble de toutes les caractéristiques et continuer par la suppression séquentielle, à chaque itération, de la caractéristique la moins pertinente.

Une autre façon est de commencer la recherche par un sous-ensemble quelconque de caractéristiques. une fois que le point de départ est bien choisi, une procédure de recherche (également connue sous le nom "organisation de la recherche"), servant à générer des sous-ensembles de caractéristiques doit être définie. en général, les stratégies de recherche peuvent être classées en trois catégories : exhaustive, heuristique et aléatoire.

a) Génération exhaustive

Dans cette approche, une recherche exhaustive sur tous les sous-ensembles de caractéristiques est effectuée afin de sélectionner le "meilleur" sous-ensemble de caractéristiques. cette stratégie de recherche garantit de trouver le sous-ensemble optimal. le problème majeur de cette approche est que le nombre de combinaisons croit exponentiellement en fonction du nombre de caractéristiques. pour un ensemble de N caractéristiques, et quand N devient grand, les 2^N combinaisons possibles rendent la recherche exhaustive impossible (problème NP-complet, Blum et Rivest [1993]).

b) Génération heuristique

Dans cette catégorie, une approche heuristique pour guider la recherche est utilisée. les algorithmes qui utilisent cette approche sont généralement des algorithmes itératifs dont chaque itération permet de sélectionner ou de rejeter une ou plusieurs caractéristiques.

Les avantages de ces algorithmes sont leur simplicité et leur rapidité. en revanche, ils ne permettent pas de parcourir totalement l'espace de recherche. dans la littérature, les trois sous-catégories les plus connues de cette approche sont :

Forward : cette approche est également appelée ascendante, son principe est de commencer avec un ensemble de caractéristiques vide et à chaque itération une ou plusieurs caractéristiques seront ajoutées.

Backward : cette approche procède à l'inverse de "Forward". L'ensemble de départ représente l'ensemble total des caractéristiques et à chaque itération, une ou plusieurs caractéristiques seront supprimées. cette approche est aussi qualifiée de descendante.

Stepwise : cette approche est un mélange des deux précédentes et consiste à ajouter ou supprimer des caractéristiques au sous-ensemble courant.

c) génération aléatoire

Pour un ensemble de données et une initialisation particulière, une stratégie de recherche heuristique retourne toujours le même sous-ensemble, ce qui la rend très sensible au changement de l'ensemble de données.

La procédure de recherche aléatoire (appelée aussi stochastique ou non-déterministe) consiste à générer aléatoirement un nombre fini de sous-ensembles de caractéristiques afin de sélectionner le meilleur. en outre, les stratégies de recherche aléatoires convergent en général rapidement vers une solution "semi-optimale", ce qui est préférable pour éviter le phénomène de sur-apprentissage.

1.3.2. Procédures d'évaluation

Les méthodes utilisées pour évaluer un sous-ensemble de caractéristiques dans les algorithmes de sélection peuvent être classées en trois catégories principales: "**Filter**", "**Wrapper**" et "**Embedded**".

a - Filter

Le modèle "Filter" a été le premier utilisé pour la sélection de caractéristiques. dans celui-ci, le critère d'évaluation utilisé évalue la pertinence d'une caractéristique selon des mesures qui reposent sur les propriétés des données d'apprentissage. cette méthode est considérée, davantage comme une étape de prétraitement (filtrage) avant la phase d'apprentissage.

En d'autres termes, l'évaluation se fait généralement indépendamment d'un classificateur (John et al. [1994]). les méthodes qui se basent sur ce modèle pour l'évaluation des caractéristiques, utilisent souvent une approche heuristique comme stratégie de recherche. La procédure du modèle "filtre" est illustrée par la figure (II.2) [18].

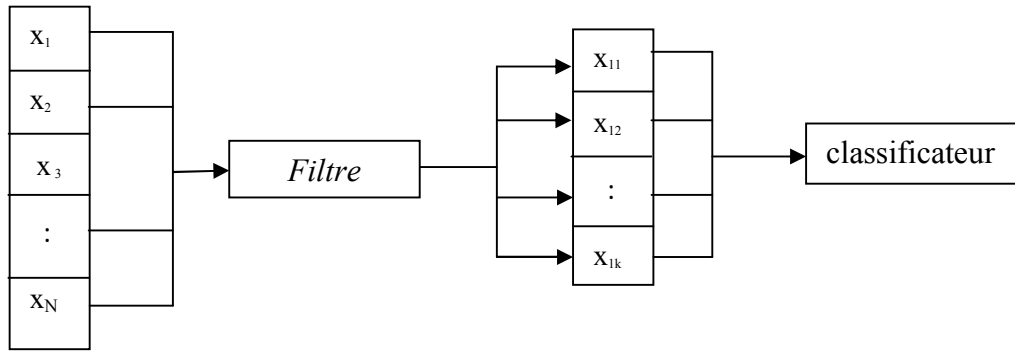


Figure II.2: La procédure du modèle "filtre"[18].

Les caractéristiques sont généralement évaluées par des mesures calculées pour chacune des caractéristiques.

Soit $X = \{x_k | x_k = (x_{k1}, x_{k2}, \dots, x_{kn}) ; k = 1, 2, \dots, m\}$ un ensemble de m exemples d'apprentissage dans un espace de représentation comportant n caractéristiques.

Soit $Y = \{y_k, k = 1, 2, \dots, m\}$ ou y_k représente l'étiquette de la classe de l'exemple x_k .

Si $x_i = (x_{1i}, x_{2i}, \dots, x_{mi})$

représente la $i^{\text{ème}}$ caractéristique ($i = 1, 2, \dots, n$) alors le but d'une méthode d'évaluation "filtre" est de calculer un score pour évaluer le degré de pertinence de chacune des caractéristiques (x_i). ci-dessous, nous présentons quelques mesures utilisées dans la littérature comme score ou critère d'évaluation (Guyon et Elisseeff [2003]) :

a.1- Le critère de corrélation : ce score est utilisé dans le cas d'une classification binaire $y_k \in \{-1, 1\}$. Il est estimé comme suit :

$$C(i) = \frac{\sum_{k=1}^m (x_{ki} - \mu_i)(y_k - \mu_y)}{\sqrt{\sum_{k=1}^m (x_{ki} - \mu_i)^2 \sum_{k=1}^m (y_k - \mu_y)^2}} \tag{II.2}$$

où μ_i et μ_y représentent respectivement les valeurs moyennes de la $i^{\text{ème}}$ caractéristiques et des étiquettes de l'ensemble d'apprentissage, $\|.\|$ est la norme euclidienne usuelle.

Cette fonction calcule le cosinus de l'angle entre chacune des caractéristiques et le vecteur des étiquettes. en d'autres termes, et pour une caractéristique donnée, une grande valeur absolue de cette mesure indique sa forte corrélation linéaire avec le vecteur des étiquettes (Y).

a.2- Le critère de Fisher : permet de mesurer le degré de séparabilité des classes à l'aide d'une caractéristique donnée (Duda et al. [2000], Furey et al. [2000]). Il est défini par :

$$f(i) = \frac{\sum_{c=1}^C n_c (\mu_c^i - \mu^i)^2}{\sum_{c=1}^C n_c (\sigma_c^i)^2} \tag{II.3}$$

où n_c , μ_c^i et σ_c^i représentent respectivement l'effectif, la moyenne et l'écart type de la $i^{\text{ème}}$ caractéristique au sein de la classe c . μ^i est la moyenne globale de la $i^{\text{ème}}$ caractéristique.

On pourrait dire que la mesure est liée à la variance interclasse de la caractéristique.

a.3- L'information mutuelle est une mesure de dépendance entre les distributions de deux populations (Fraser et Swinney [1986]). Soient X et Y deux variables aléatoires dont les instances sont respectivement les valeurs de la $i^{\text{ème}}$ caractéristique et les étiquettes des classes. l'information mutuelle $I(i)$ est définie comme la divergence de Kullback-Leibler (KL)(Cover et Thomas [1991]) entre la probabilité $P(x^i, y)$ et le produit des probabilités ($P(x^i)P(y)$) [19]. l'information mutuelle est estimée empiriquement par :

$$I(i) = \sum_{x_i} \sum_y P(X = x_i, Y = y) \log \frac{P(X=x_i, Y=y)}{P(X=x_i)P(Y=y)} \quad (II.4)$$

où les probabilités $P(x_i)$, $P(y)$ et $P(x_i, y)$ sont estimées par les fréquences des différentes valeurs possibles.

SNR(Signal-to-Noise Ratio coefficient) est un score qui mesure le pouvoir de discrimination d'une caractéristique entre deux classes. d'une manière similaire au critère de fisher, cette méthode classe les caractéristiques en calculant le rapport de la valeur absolue de la différence des moyennes des classes et de la moyenne des écart-types des classes.

La formule de SNR pour une caractéristique et pour un problème à deux classes est calculée par :

$$SNR(i) = \frac{2 \times |\mu_{C_{i1}} - \mu_{C_{i2}}|}{(\sigma_{C_{i1}} + \sigma_{C_{i2}})} \quad (II.5)$$

D'autres critères d'évaluation sont proposés dans (Golub et al. [1999], Tusher et al. [2001], Hastie et al. [2001]).

Le principal avantage des méthodes de filtrage est leur efficacité calculatoire et leur robustesse face au sur-apprentissage. Malheureusement, ces méthodes ne tiennent pas compte des interactions entre caractéristiques et tendent à sélectionner des caractéristiques comportant de l'information redondante plutôt que complémentaire (Guyon et Elisseeff [2003]). De plus, ces méthodes ne tiennent absolument pas compte de la performance des méthodes de classification qui suivent la sélection (Kohavi et John [1997])

b. Wrapper

Le principal inconvénient des approches "*filter*" est le fait qu'elles ignorent l'influence des caractéristiques sélectionnées sur la performance du classificateur à utiliser par la suite. Pour résoudre ce problème, Kohavi et John ont introduit le concept "*wrapper*" pour la sélection de caractéristiques (Kohavi et John [1997]). Les méthodes "*wrapper*", appelées aussi

méthodes enveloppantes, évaluent un sous-ensemble de caractéristiques par sa performance de classification en utilisant un algorithme d'apprentissage. La procédure du modèle "wrapper" est illustrée par la figure II.3.

L'évaluation se fait à l'aide d'un classificateur qui estime la pertinence d'un sous-ensemble

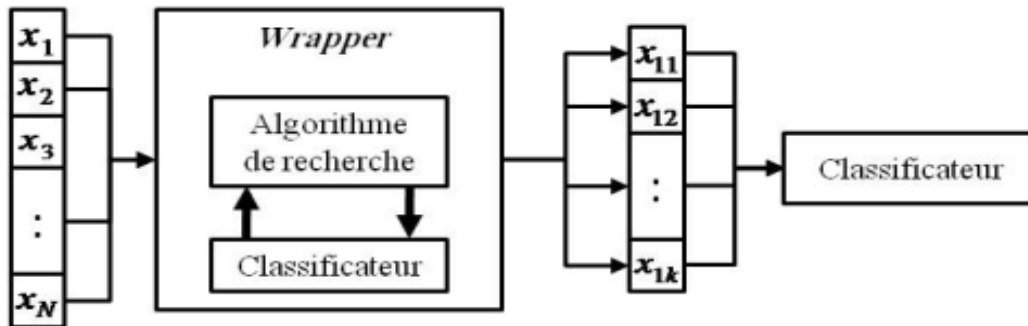


Figure II.3: La procédure du modèle "wrapper"

donné de caractéristiques. Les sous-ensembles de caractéristiques sélectionnés par cette méthode sont bien adaptés à l'algorithme de classification utilisé, mais ils ne sont pas forcément valides si on change le classificateur. la complexité de l'algorithme d'apprentissage rend les méthodes "wrapper" très coûteuses en temps de calcul.

En général, pour diminuer le temps de calcul et pour éviter les problèmes de sur-apprentissage, le mécanisme de validation croisée est fréquemment utilisé. une meilleure performance des méthodes "wrapper" par rapport à certaines méthodes de filtrage a été démontré par (Kohavi et John [1997]).

Le problème de la complexité de cette technique rend impossible l'utilisation d'une stratégie de recherche exhaustive (problème NP-complet). par conséquent, des méthodes de recherche heuristiques ou aléatoires peuvent être utilisées. la recherche devient néanmoins, de plus en plus irréalisable avec l'augmentation de la taille de l'ensemble initial de caractéristiques.

Les méthodes "wrapper" sont généralement considérées comme étant meilleures que celles de filtrage selon (Li et Guo [2008], Huang et al. [2008]). elles sont capables de sélectionner des sous-ensembles de caractéristiques de petite taille qui sont performants pour le classificateur utilisé mais il existe deux inconvénients principaux qui limitent ces méthodes:

- 1) La complexité et le temps de calcul nécessaire pour la sélection.

Le principal inconvénient de l'approche "wrapper" est le temps nécessaire pour la sélection des caractéristiques, il est nettement plus long que celui des approches de filtrage et d'autres approches de sélection de caractéristiques. l'utilisation d'un classificateur pour

évaluer les sous-ensembles ainsi que les techniques d'évaluation (validation croisée par exemple) rendent les méthodes "Wrapper" très coûteuses en terme de temps de calcul.

2) La dépendance des caractéristiques pertinentes sélectionnées par rapport au classificateur utilisé.

La deuxième limitation de l'approche "wrapper" est que l'évaluation des caractéristiques se fait par un seul classificateur lors de la sélection. chaque classificateur à ses spécificités et ses hypothèses. ainsi le sous-ensemble sélectionné dépend toujours du classificateur utilisé.

c . Embedded

A la différence des méthodes "wrapper" et "filter", les méthodes "embedded" (appelées aussi méthodes intégrées) incorporent la sélection de variables lors du processus d'apprentissage. un tel mécanisme intégré pour la sélection des caractéristiques peut être trouvé, par exemple, dans les algorithmes de type SVM, AdaBoost, ou dans les arbres de décisions .

Dans les méthodes de sélection de type "wrapper", la base d'apprentissage est divisée en deux parties : une base d'apprentissage et une base de validation pour valider le sous-ensemble de caractéristiques sélectionné. en revanche, les méthodes intégrées peuvent se servir de tous les exemples d'apprentissage pour établir le système. cela constitue un avantage qui peut améliorer les résultats. un autre avantage de ces méthodes est leur plus grande rapidité par rapport aux approches "Wrapper" parce qu'elles évitent que le classificateur recommence de zéro pour chaque sous-ensemble de caractéristiques.

1.4. Critère D'arrêt

Certains critères doivent être définis pour arrêter le processus de recherche sur les sous-ensembles de caractéristiques. pour les méthodes de filtrage, le critère d'arrêt couramment utilisé est basé sur l'ordre des caractéristiques, rangées selon certains scores de pertinence (généralement des mesures statistiques). une fois les caractéristiques ordonnées, celles qui ont les scores les plus élevés seront choisies et utilisées par un classificateur.

Pour les méthodes de type "Wrapper", le processus de recherche peut s'arrêter lorsque il n'y a plus d'amélioration de précision. en d'autres termes, lorsqu'il n'y a plus la possibilité de trouver un sous-ensemble meilleur que le sous-ensemble actuel. un critère d'arrêt pour les méthodes enveloppantes est de continuer à rechercher jusqu'au moment où la précision dépasse un certain seuil défini par l'utilisateur.

1.5. Revue de quelques méthodes de sélection

Dans cette section, nous offrons quelques moyens de sélectionner les caractéristiques de la littérature, Nous avons choisi d'introduire des techniques basées sur différentes techniques parmi lesquelles on a étudié la méthode MRMR.

Recherche prédéfinie ainsi que diverses techniques d'évaluation.

1.5.1. SFS et SBS

SFS (Sequential Forward Selection) ou (sélection séquentielle croissante) est la première méthode proposée pour la sélection de caractéristiques. cette méthode à été proposée en 1963 par Marill et Green (Marill et Green [1963]). une approche heuristique de recherche est utilisée dans cette méthode, en commençant par un ensemble vide de caractéristiques.

A chaque itération, la meilleure caractéristique parmi celles qui restent sera sélectionnée. le processus de sélection continue jusqu'à un critère d'arrêt en 1971, Whitney (Whitney [1971]) a proposé une méthode similaire au SFS appelée SBS (Sequential Backward Selection) où (sélection séquentielle arrière). a la différence de la méthode SFS, cette méthode commence par l'ensemble de toutes les caractéristiques et à chaque itération, Bien que les deux méthodes SFS et SBS semblent similaires, Ahan et Bankert (Aha et Bankert [1995]) ont montré que la méthode SBS est plus performante parce qu'elle prend en considération l'interaction d'une caractéristique avec un ensemble de caractéristiques plus large, contrairement au SFS qui ne prend en considération que l'interaction de cette caractéristique avec le sous-ensemble déjà sélectionné. par ailleurs, l'évaluation des sous-ensembles de grande taille avec la méthode SBS pose un problème au niveau de temps de calcul[20].

1.5.2. LDA (Analyse Discriminante Linéaire) :

La LDA est une méthode d'analyse numérique qui permet de chercher la combinaison linéaire des variables qui représentent au mieux les données, elle est très utilisée dans le domaine de la reconnaissance des formes à savoir la reconnaissance de visage. elle permet de maximiser l'éparpillement inter-classes (the between-class scatter) et de réduire l'éparpillement intraclasses (the withinclass scatter) [21].

1.5.3 MRMR (Max-relevance, Min-Redundancy)

MRMR (Max-relevance, Min-Redundancy) est une méthode de filtrage pour la sélection de caractéristiques proposée par Peng et al. en 2005 (Peng et al. [2005]). Cette méthode est basée sur des mesures statistiques classiques comme l'information mutuelle, la corrélation etc. l'idée de base est de profiter de ces mesures pour essayer de minimiser la redondance (MR) entre les caractéristiques et de maximiser la pertinence (MR). les auteurs

proposent deux variantes de leur méthode. une pour des données discrètes et l'autre pour des données continues. pour les données discrètes, les auteurs utilisent l'information mutuelle pour calculer les deux facteurs **MR** et **MR**. le calcul de la redondance et de la pertinence d'une caractéristique est donné par l'équation II.6.

$$\mathbf{Redondance}(i) = \frac{1}{|F|^2} \sum_{i,j \in F} I(i,j), \mathbf{Pertinence}(i) = \frac{1}{|F|^2} \sum_{i,j \in F} I(i,Y) \quad (\text{II.6})$$

où F et $|F|$ représentent, respectivement, l'ensemble des caractéristiques et sa taille. $I(i,J)$ est l'information mutuelle entre la $i^{\text{ème}}$ et la $j^{\text{ème}}$ caractéristique et finalement $I(i,Y)$ est l'information mutuelle entre la $i^{\text{ème}}$ caractéristique et l'ensemble des étiquettes de classes (Y). le score d'une caractéristique est la combinaison de ces deux facteurs tel que :

$$\mathbf{score}(i) = \frac{\mathbf{Pertinence}(i)}{\mathbf{Redondance}(i)} \text{ ou } \mathbf{score}(i) = \mathbf{Pertinence}(i) - \mathbf{Redondance}(i) \quad (\text{II.7})$$

Pour les données continues, les auteurs ont remplacé l'information mutuelle par d'autres mesures. Pour la redondance ils ont utilisé la mesure de corrélation, par contre, la mesure F-statistique est utilisée pour calculer la pertinence.

Après cette évaluation individuelle des caractéristiques, une technique de recherche avant séquentielle est utilisée avec un classificateur pour sélectionner le sous-ensemble final de caractéristiques. en d'autres termes, un classificateur est utilisé pour évaluer les sous-ensembles en commençant par la caractéristique qui a le meilleur score, les deux meilleures, etc., jusqu'à trouver le sous-ensemble qui minimise l'erreur de classification.

1.5.4. Les algorithmes génétiques

Les algorithmes génétiques ont été utilisés dans le domaine de la sélection de caractéristiques afin d'accélérer la recherche et d'éviter les optima locaux. de nombreuses études rapportées dans la littérature ont montré que les méthodes qui utilisent les AGs comme technique de recherche ont donné de meilleurs résultats que les résultats obtenus par les autres méthodes de sélection (Jain et Zongker [1997], Kuncheva et Jain [1999], Ishibuchi et Nakashima [2000]). le chapitre 4 sera consacré à la présentation d'une description détaillée des algorithmes génétiques ainsi que des méthodes de sélection qui utilisent ces techniques[22].

2. Réduction basée sur une transformation de données

La réduction de la dimensionnalité par une transformation de données (appelée aussi extraction de caractéristiques) ne se fait pas par une sélection de certaines caractéristiques, mais par une construction de nouvelles caractéristiques obtenues en combinant les caractéristiques initiales. Une transformation de données risque de faire perdre la sémantique

de l'ensemble initial de caractéristiques et donc l'utilisation de cette famille de méthodes n'est applicable que dans le cas où la sémantique n'intervient plus dans les étapes qui suivent la réduction. Les sections suivantes décrivent brièvement plusieurs techniques de réduction connues. Elles sont généralement groupées en deux catégories : les méthodes linéaires et les méthodes non linéaires.

2.1. Méthodes linéaires

Nous rappelons brièvement les principes de quelques méthodes classiques d'analyse de données, elles sont le fondement de plusieurs méthodes non linéaires plus récentes.

2.1.1. Analyse en composantes principales

L'Analyse en Composantes principales (ACP) fait partie du groupe des méthodes descriptives multidimensionnelles appelées méthodes factorielles. L'ACP est une technique qui permet de trouver des espaces de dimensions plus petites dans lesquels il est possible d'observer au mieux les individus. Sa démarche essentielle consiste à transformer les variables quantitatives initiales, plus ou moins corrélées entre elles, en des variables quantitatives, non corrélées, combinaisons linéaires des variables initiales et appelées composantes principales.

Les composantes principales sont donc de nouvelles variables indépendantes, combinaisons linéaires des variables initiales, possédant une variance maximale. Globalement l'ACP consiste à rechercher la direction suivant laquelle le nuage de points des observations s'étire au maximum. A cette direction correspond la première composante principale. La seconde composante principale est déterminée de telle sorte qu'elle soit la plus indépendante possible de la première, elle est donc perpendiculaire à celle-ci. Ces deux composantes forment le premier plan principal. Cette opération est répétée de manière à trouver toutes les composantes principales expliquant le maximum de variance.

La figure II.4 montre à gauche, un exemple de données 3D qui se trouvent dans un plan 2D et à droite les deux premières composantes principales sur ces données.

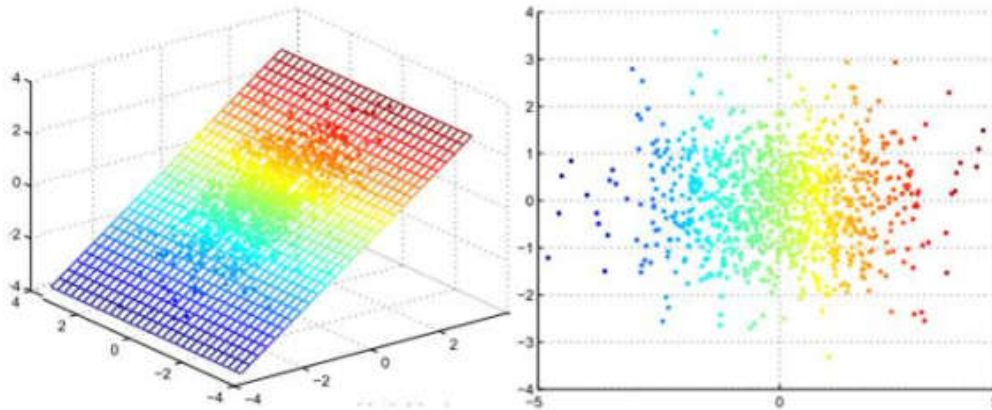


Figure II.4: ACP sur des données linéaires

Supposons que nous ayons un ensemble de données $X = \{x_1, x_2, \dots, x_M\}$ composé de M observations où chaque observation $x_i = (x_{i1}, x_{i2}, \dots, x_{iN})$ est composée de N caractéristiques. X est associé à une matrice de données A de taille $N \times M$ où chaque colonne représente une caractéristique. En pratique, le calcul de l'ACP pour la matrice X revient à réaliser les opérations ci-dessous afin de trouver les composantes principales :

1	Calculer le vecteur $\mu = (\mu_1, \mu_2, \dots, \mu_m)^T$ qui représente le vecteur moyen où μ_i est la moyenne de la $i^{\text{ème}}$ composante des données.
2	Calculer la matrice χ en soustrayant le vecteur moyen à toutes les colonnes de A dans le but d'obtenir des données centrées.
3	Calculer la matrice S (de taille $N \times N$) de covariance de χ avec ($S = \chi \cdot \chi^T$).
4	Calculer la matrice U (de taille $N \times N$) qui est composée des coordonnées des vecteurs propres \vec{u}_j de S triés par ordre décroissant des modules des valeurs propres λ_j (la première colonne de U est le vecteur propre qui correspond à la plus grande valeur propre)
5	Garder les R premières colonnes de U pour former la matrice $\tilde{U} : N \times R$ qui représente les R premières composantes principales.

Tableau II.1 : Opération pour trouver les composantes principales

L'ACP étant une méthode de réduction de dimension, il est important de savoir qu'elle ne peut pas retenir la totalité de l'information contenue dans le nuage de points initial.

Enfin, l'ACP prend uniquement en compte les dépendances linéaires entre les variables et ne peut donc pas fournir une projection pertinente pour une distribution non-linéaire de points. La figure II.5 montre à gauche, un exemple de données non-linéaires (non réparties

dans un plan) et à droite le résultat de leur projection dans un plan génère par les deux premières composantes principales calculées sur ces données.

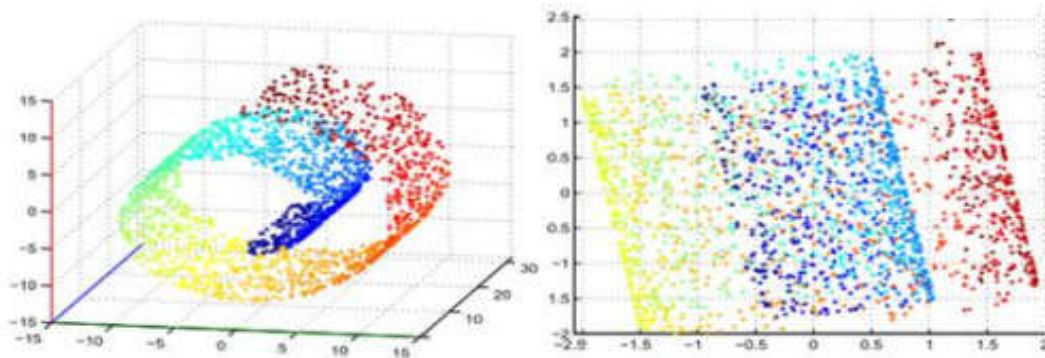


Figure II.5: ACP sur des données non-linéaires

2.1.2. Analyse linéaire discriminante

L'analyse linéaire discriminante, appelée aussi analyse discriminante linéaire de Fisher, est une méthode de réduction du nombre de dimensions proposée par Fisher en 1936 (Fisher [1936]). cette méthode s'applique lorsque les classes des individus sont connues. l'idée de Fisher a été de créer une méthode pour choisir entre les combinaisons linéaires des variables celles qui maximisent l'homogénéité de chaque classe. en d'autres termes, cette méthode consiste à chercher un espace vectoriel de faible dimension qui maximise la variance interclasse (pour une description complète de la méthode voir (Lebart et Morineau [1995])).

2.1.3. Positionnement multi-dimensionnel

La méthode de positionnement multidimensionnel (Messick et Abelson [1956]), MDS (Multi Dimensional Scaling) en anglais, permet de construire une représentation en faible dimension des points de l'espace. son objectif est de construire, à partir d'une matrice de distances ou des mesures de similarité calculées sur chaque paire de points, une représentation euclidienne des individus dans un espace de dimension réduite qui préserve "au mieux" ces distances. supposons que nous ayons un ensemble de données $X = \{x_1, x_2, \dots, x_M\}$ composé de M observations où chaque observation $x_i = (x_{i1}, x_{i2}, \dots, x_{iN})$ est composée de N caractéristiques. Soit d une matrice symétrique de taille $M \times M$ où chaque élément d_{ij} représente la distance entre x_i et x_j et $d_{ii} = 0$. l'idée de MDS est de trouver une configuration de points $y_i, i = 1 \dots M$ dans un espace de dimension plus réduite qui conserverait les distances

entre les points initiaux x_i . Autrement dit, il cherche les points y_i dans un espace de dimension $q < N$ tels que $d(y_i, y_j) \approx d_{ij} = d(x_i, x_j)$. ceci revient à optimiser le critère suivant :

$$J = \sum_{i=1}^N = \sum_{j=1}^N (d_{ij} - d(y_i, y_j))^2 \quad (\text{II.8})$$

Comme l'ACP, l'algorithme de Positionnement Multi-Dimensionnel n'est valable que sur une distribution linéaire de données.

2.2. Méthodes non-linéaires

Les méthodes non linéaires ont pour objectif d'optimiser les représentations afin qu'elles reflètent au mieux la topologie initiale des données vecteur x_i associé à chaque couple (x_i, x_j) , où x_j appartient au voisinage de x_i .

Les w_{ij} permettront par la suite de reconstruire la topologie de voisinage de chaque vecteur x_i dans l'espace de projection.

Conclusion

Dans ce chapitre, nous avons abordé le domaine de la réduction de dimensionnalité. Dans un premier temps, une revue du domaine de la sélection de caractéristiques a été présentée. Après avoir exposé les différents composants nécessaires à un algorithme de sélection de caractéristiques, les alternatives possibles pour leurs mises en œuvre ont été présentées.

Un certain nombre d'algorithmes a alors été décrit en fonction de la combinaison de composants utilisée cela en présentant les avantages et les inconvénients des techniques de sélection en général mais aussi de ceux des algorithmes décrits.

Les méthodes existantes présentent des limitations au niveau de la complexité très élevée des approches "Wrapper" ainsi qu'en raison de la dépendance des caractéristiques pertinentes sélectionnées par rapport au classificateur utilisé. par ailleurs, les approches "Filter" présentent plusieurs limitations concernant la redondance ainsi que les interactions entre les caractéristiques.

CHAPITRE III : Réseaux des neurones

Artificiels

Sommaire

Introduction.....	36
1. Les réseaux de neurones artificiels.....	36
1.1. Le neurone.....	36
1.2. Les bases biologiques.....	36
1.3. Définition les réseaux de neurones artificiels.....	37
1.4. Structure des réseaux de neurones artificiels.....	38
1.5. Utilité de réseaux de neurones artificiels.....	38
2. Le neurone formel.....	38
2.1. Définition et représentation.....	38
2.2. Poids des connexions.....	41
2.3. Transformation de l'information.....	41
3. Fonction de Transfer.....	41
4. Quelques modèles des réseaux de neurones artificiels.....	42
4.1. Les réseaux multicouches.....	42
4.2. les réseaux ELM (Extreme Learning Machines).....	43
4.2.1. Modèle mathématique.....	44
4.2.2. Algorithme.....	45
4.2.3. Les caractéristiques principales.....	45
4.3. les réseaux de neurones à base radiale.....	45
4.3.1. Présentation des réseaux RBF.....	45
4.3.2. Architecture générale d'un réseau RBF.....	46
4.3.3. La fonction à base radiale.....	47
4.3.4. La phase d'apprentissage.....	48
4.4. Avantages et inconvénients des réseaux RBF.....	51
5. Apprentissage.....	52
5.1. rétro-propagation de l'erreur.....	52
5.2. Algorithmes d'optimisation.....	53
6. L'organisation des neurones en réseaux.....	54
Conclusion.....	54

Introduction

Partant d'un souhait initial de mieux comprendre le fonctionnement du cerveau, les réseaux de neurones artificiels sont une classe de modèles ayant eu un impact important dans le domaine de l'analyse et le traitement des données. l'inspiration pour les réseaux de neurones provient cependant de la volonté de créer des systèmes artificiels sophistiqués, voir intelligents, capables d'effectuer des opérations semblables à celle que le cerveau humain effectue de manière routinière.

Les réseaux de neurones intéressent actuellement différents domaines tels que : la médecine, l'électronique, l'informatique, l'automatique, la robotique, la classification, le contrôle, le traitement des signaux, le traitement des images, etc [A. Coulon, 2006].

1. Réseaux de neurones artificiels

Les réseaux de neurones sont une abstraction biologique destinée à modéliser des comportements naturels.

1.1. Le neurone

Le corps cellulaire du neurone se ramifie pour former ce que l'on nomme les dendrites. Celles-ci sont parfois si nombreuses que l'on parle alors de chevelure dendritique ou d'arborisation dendritique. c'est par les dendrites que l'information est acheminée de l'extérieur vers le soma (le corps du neurone). l'information traitée par le neurone chemine ensuite le long de l'axone (unique) pour être transmise aux autres neurones.

La transmission entre deux neurones n'est pas directe. en fait, il existe un espace intercellulaire de quelques dizaines d'Angströms (10^{-9} m) entre l'axone du neurone afférent et les dendrites (on dit une dendrite) du neurone efférent. la jonction entre deux neurones est appelée la synapse.

1.2. Les bases biologiques

Les cellules nerveuses, appelées neurones sont des éléments de base du système nerveux central qui se compose d'environ 10^{12} neurones (mille milliards), le neurone est une cellule composée de trois parties :

a- Un corps cellulaire ou en anglais « **Cell body** » qui contient le noyau et se charge d'effectuer les transformations biochimiques essentielles à la synthèse des éléments assurant la vie du neurone.

b- Les dendrites qui sont des ramifications du corps cellulaire. elles permettent au neurone de capter les signaux lui provenant de l'extérieur.

c- L'**axone** généralement plus long que les dendrites, il se ramifie à son extrémité où il communique avec les autres neurones. il sert de moyen de transport pour les signaux émis par le neurone.

Les connexions entre neurones sont réalisées au niveau des **synapses**, lieu de proximité d'axone émetteur et dendrites réceptrices

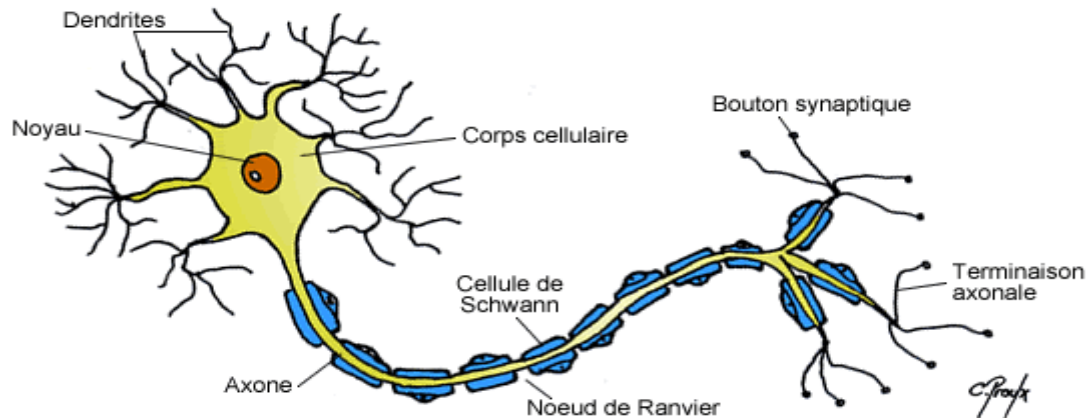


Figure III.1: Un neurone biologique et ses principaux composants.

1.3. Définition les réseaux de neurones artificiels

Un réseau de neurone est un assemblage de neurone formel associé en couches fonctionnant en parallèle. dans un réseau, chaque sous-groupe fait un traitement indépendant des autres et transmet le résultat de son analyse au sous-groupe suivant. l'information donnée au réseau va donc se propager couche par couche, de la couche d'entrée à la couche de sortie, en passant soit par aucune, une ou plusieurs couches intermédiaires (couches cachées).

Chaque neurone dans une couche est connecté à tous les neurones de la couche suivante et de la couche précédente sauf les neurones de la couche d'entrées et les couches de sorties. Contrairement à chacune des fonctions d'activation f , la fonction g qui transforme les valeurs d'entrée en valeurs de sortie à l'échelle du réseau ne peut pas être explicitée facilement. elle est en effet beaucoup plus compliquée puisqu'elle est constituée de la superposition de toutes les fonctions f de chaque neurone.

Un réseau de neurones peut donc être représenté par les poids w des différents neurones. Ces poids peuvent varier au cours du temps, en fonction des entrées présentées E . ce qui représente un vrai problème c'est comment modifier la valeurs de ces poids (la phase d'apprentissage)[23].

1.4. Structure des réseaux de neurones artificiels

Mc Culloch et Pitts en 1943 ont implémenté un système de réseaux neuronaux artificiels (RNA), qui est analogue aux neurones biologiques fondée sur une structure complexe. le système des RNA est considéré comme un arrangement d'éléments de structure identique appelés neurones interconnectées par analogie avec cellules du système nerveux humain.

Il est composé également d'une succession de couches connectées de manière à ce que chaque neurone tienne son entrée de la sortie du neurone précédant. chaque neurone dans ce cas fonctionne indépendamment par rapport aux autres afin que l'ensemble forme un système compact. l'information est emmagasinée de façon répartie dans le réseau sous forme de coefficients synaptiques. le neurone formel calcule régulièrement un résultat qu'il transmet ensuite aux neurones suivant, chaque calcul est associé à un poids qui définit la force de la connexion (Hagan et al. 1996).

1.5. Utilité de réseaux de neurones artificiels

Les réseaux de neurones, en tant que système capable d'apprendre, mettent en œuvre le principe de l'induction, c'est-à-dire l'apprentissage par l'expérience. par confrontation avec des situations ponctuelles, ils infèrent un système de décision intégré dont le caractère générique est fonction du nombre de cas d'apprentissages rencontrés et de leur complexité par rapport à la complexité du problème à résoudre. par opposition, les systèmes symboliques capables d'apprentissage, s'ils implémentent également l'induction, le font sur base de la logique algorithmique, par complexification d'un ensemble de règles déductives.

Grâce à leur capacité de classification, régression et de généralisation, les réseaux de neurones sont généralement utilisés dans des problèmes de nature statistique, tels que la classification automatique de codes postaux ou la prise de décision concernant un achat boursier en fonction de l'évolution des cours.

Les réseaux de neurones sont réellement utilisés, par exemple :

❖ En météorologie, pour la classification de conditions atmosphériques et la prévision statistique du temps approximation d'une fonction inconnue.

2. Le neurone formel

❖ 2.1. Définition et représentation

Le neurone formel (Figure 2), appelé aussi unité, nœud, automate ou élément de processus est un processeur élémentaire. il a été conçu pour la première fois par McCulloch et Pitts [24].

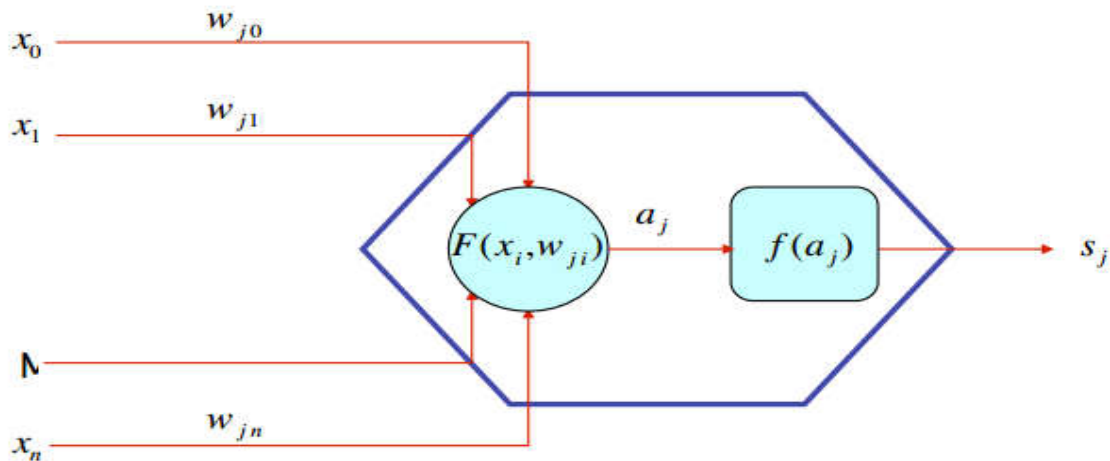


Figure III.2: Modèle du neurone formel.

Ce neurone formel est un élément binaire à seuil. Autrement dit ses entrées et sa sortie sont booléennes. il reçoit de ses voisins des entrées binaires (X_j). après une sommation pondérée de ces entrées ($\sum W_{ij} X_j$) suivie d'une transformation par une fonction de transfert, le neurone formel envoie des sorties (Y_i) binaires aux neurones voisins.

L'activation du neurone formel dépend d'un certain seuil (θ), propre au neurone. si la somme de ses entrées dépasse ce seuil, il se met dans son état actif (+1) sinon, il se met dans son état inactif (-1). malgré l'évolution des neurones formels, le modèle de McCulloch et Pitts constitue, à quelques variantes près, la base des modèles RNA conçus jusqu'à aujourd'hui.

Cette évolution concerne non seulement l'état de neurone qui peut être binaire ou continu mais aussi la fonction de transfert.

Dans le modèle du neurone de McCulloch et Pitts on trouve :

$$X = \sum_{i=1}^n x_i w_i \tag{III.1}$$

$$Y = \begin{cases} +1 & \text{si } X \geq \theta \\ -1 & \text{si } X < \theta \end{cases} \tag{III.2}$$

Neurone Biologique	Neurone Formel
Synapses	Poids de connexion
Axones	Signal de sortie
Dendrite	Signal d'entrée
Corps cellulaire	Fonction d'activation

Tableau III.1: analogie entre le neurone biologique et le neurone formel

2.2. Poids des connections

Les entrées X_j au neurone i sont des sorties des neurones j du même réseau, ou d'un autre réseau, ou encore des entrées extérieures. les connections traduisent l'influence de l'environnement extérieur au neurone et permettent de calculer son entrée globale. à chacune de ces entrées j est associé un poids W_{ij} qui détermine l'intensité de cette influence. l'entrée d'un neurone est fonction des signaux provenant des autres neurones et des intensités des connections. la fonction liant ces quantités est la somme des entrées pondérées de leurs poids correspondants. la fonction d'entrée totale peut donc s'écrire comme suit:

$$S_i = \sum_j W_{ij} X_j \quad (\text{III.3})$$

2.3. Transformation de l'information

L'état interne que peut prendre un neurone peut être défini par une fonction d'activation, ou fonction de transfert. il existe plusieurs formes pour la fonction de transfert : linéaires, binaires, probabilistes, à saturation ou continues. les fonctions binaires à seuil telles que la fonction de Heaviside (ou signe) et les fonctions sigmoïdes, dont plusieurs formes existent, sont des exemples de ces fonctions. la figure 3 représente quelques exemples de fonctions de transfert courantes.

3. Fonction de transfère

Jusqu'à présent, nous n'avons pas spécifié la nature de la fonction d'activation de notre modèle. il se trouve que plusieurs possibilités existent. différentes fonctions de transfert pouvant être utilisées comme fonction d'activation du neurone sont énumérées au tableau III.2. les trois les plus utilisées sont les fonctions «seuil» (en anglais «hard limite»), «linéaire» et «sigmoïde». comme son nom l'indique, la fonction seuil applique un seuil sur son entrée. Plus précisément, une entée négative ne passe pas le seuil, la fonction retourne alors la valeur 0 (on peut interpréter ,ce 0 comme signifiant *faux*), alors qu'une entrée positive ou nulle dépasse le seuil, et la fonction










Nom de la fonction	Relation d'entrée/sortie	Icône	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 $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ 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	$a = \frac{1}{1 + \exp^{-n}}$		logsig
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
compétitive	$a = 1$ si n maximum $a = 0$ autrement		compet

Tableau III.2 : Fonctions de transfert $a = f(n)$.

retourne 1 (vrai). utilisée dans le contexte d'un neurone, cette fonction est illustrée à la figure III.3.a. on remarque alors que le biais b dans l'expression de $a = \text{hardlim}(\mathbf{w}^T \mathbf{p} - b)$ (équation III.3) détermine l'emplacement du seuil sur l'axe $\mathbf{w}^T \mathbf{p}$, ou la fonction passe de 0 à 1. Nous verrons plus loin que cette fonction permet de prendre des décisions binaires. la fonction linéaire est très simple, elle affecte directement son entrée à sa sortie :

$$a = n. \tag{III.4}$$

Appliquée dans le contexte d'un neurone, cette fonction est illustrée à la figure III.3. dans ce cas ,la sortie du neurone correspond a son niveau d'activation dont le passage à zéro se produit lorsque

$$\mathbf{w}^T \mathbf{p} = b \tag{III.5}$$

La fonction de transfert sigmoïde est quant a elle illustrée à la figure . son équation est donnée par :

$$a = \frac{1}{1 + \exp^{-n}} \tag{III.6}$$

Elle ressemble soit a la fonction seuil, soit à la fonction linéaire, selon que l'on est loin ou près de b , respectivement. la fonction seuil est très non-linéaire car il y a une discontinuité lorsque $\mathbf{w}^T \mathbf{p} = b$.

De son côté, la fonction linéaire est tout à fait linéaire. elle ne comporte aucun changement de pente. la sigmoïde est un compromis intéressant entre les deux précédentes. notons finalement, que la fonction «tangente hyperbolique» est une version symétrique de la sigmoïde.

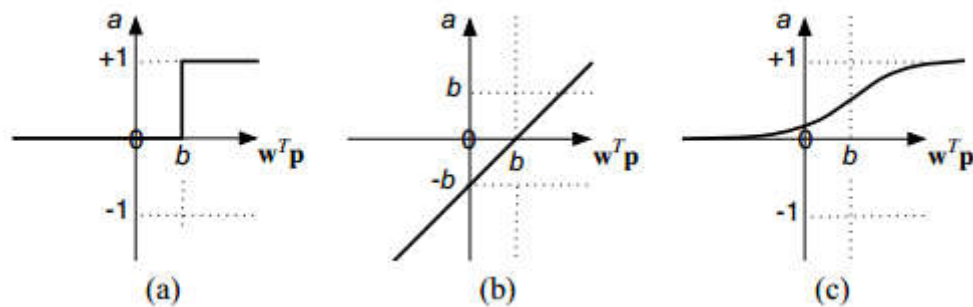


Figure III.3: Fonction de transfert : (a) du neurone «seuil» ; (b) du neurone «linéaire », et (c) du neurone «sigmoïde».

4. Quelques Modèles des Réseaux de Neurones Artificiels :

Il existe plusieurs modèles de réseaux de neurones artificiels dans ce paragraphe en va cité quelques un à titre d'exemple.

- ❖ Les réseaux de Kohonen
- ❖ Les réseaux de Hopfield
- ❖ Les réseaux multicouches
- ❖ Les réseaux ELM (Extrême Learning Machines).
- ❖ Les réseaux RBF (Radial Basis Function).

Dans ce qui suit nous allons donner une petite description sur deux types de réseaux de neurones artificiels : réseaux Extrême Learning Machines (ELM) et Radial Basis Function (RBF) .

4.1. Les réseaux multicouches

Apparus en 1986, les réseaux multicouches sont les plus employés .ces modèles sont fortement inspirés des études biologiques du système visuel. constitués par plusieurs couches, ces réseaux présentent une solution aux limites posées par le perceptron. à la différence des autres réseaux, ils contiennent des couches internes, dites couches cachées.

Les neurones d'une même couche ne sont pas reliés entre eux .

Les réseaux multicouches sont composés:

- ❖ d'une couche d'entrée dont les neurones reçoivent l'information présentée au réseau,
- ❖ d'une couche de sortie qui fournit les résultats des traitements effectués par le réseau,

❖ d'un nombre variable de couches intermédiaires (couches cachées).

Les neurones de chaque couche sont complètement connectés à ceux de la couche suivante. L'information se transmet de la couche d'entrée vers la couche de sortie. on parle d'une propagation en avant, d'où le nom approprié en anglais "feed forward neural networks" (réseaux de neurones nourris vers l'avant).

Un autre type de réseaux multicouches est le réseau récurrent "récurrent neural networks". les connexions récurrentes ramènent l'information en arrière par rapport au sens de la propagation défini dans un réseau multicouches. ce type de réseau est constitué :

❖ d'une couche d'entrée

❖ d'une couche cachée constituée par des neurones similaires aux neurones de Kohonen, influencés par les entrées.

❖ d'une couche de sortie influencée par les réponses désirées.

Les réseaux récurrents sont très peu utilisés . quelques applications sont citées dans la littérature [25].

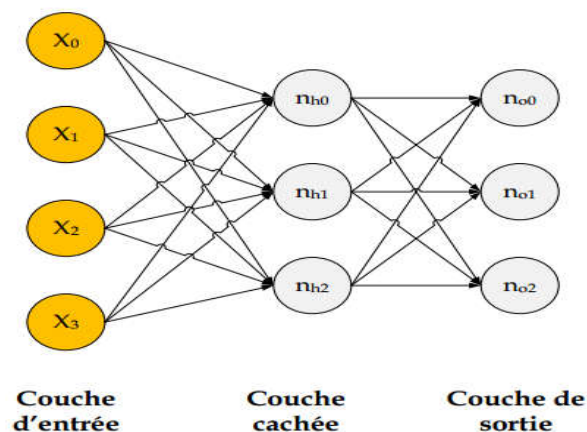


Figure III.4: Schéma d'un réseau multicouches.

4.2. Les réseaux ELM (extrême learning machines)

L'apprentissage automatique ou apprentissage statistique (Machine Learning en Anglais), champ d'étude de l'intelligence artificielle, concerne la conception, l'analyse, le développement et l'implémentation de méthodes permettant à une machine (au sens large) d'évoluer par un processus systématique, et ainsi de remplir des tâches difficiles ou impossibles à remplir par des moyens algorithmiques plus classiques.

En apprentissage automatique, le terme Extreme Learning Machine (ELM) fait référence à un type de réseau de neurones. sa spécificité est de n'avoir qu'une seule couche de nœuds cachés, où les poids des entrées de connexion de nœuds cachés sont répartis au hasard

et jamais mis à jour. ces poids entre les nœuds cachés d'entrée et les sorties sont appris en une seule étape, ce qui revient essentiellement à l'apprentissage d'un modèle linéaire. le nom "Extreme Learning Machine" a été donné à ces modèles par Guang-Bin Huang, mais le principe était déjà connu.

Ces modèles peuvent produire une bonne performance de généralisation et avoir un processus d'apprentissage beaucoup plus rapide que les réseaux formés en utilisant la rétro propagation du gradient [26].

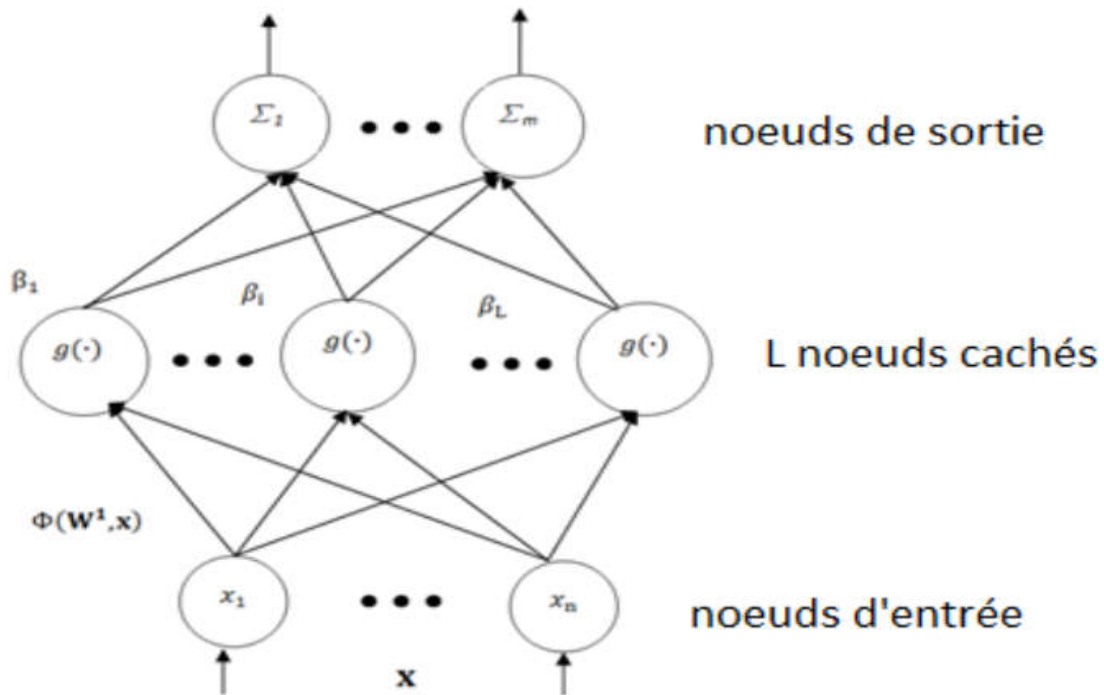


Figure III.5: Illustration de la structure du réseau de neurones ELM [26].

4.2.1. Modèle Mathématique

$\sum_{i=1}^L \beta_i G(a_i, b_i, x_i) = t_j, j = 1, \dots, N$ est équivalent à $H\beta = T$, ou

$$H(a_1, \dots, a_L, b_1, \dots, b_L, x_1, \dots, x_N)$$

$$= \begin{bmatrix} G(a_1, b_1, x_1) & \dots & G(a_L, b_L, x_1) \\ \vdots & \dots & \vdots \\ G(a_1, b_1, x_N) & \dots & G(a_L, b_L, x_N) \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad \text{et} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \tag{III.7}$$

\mathbf{H} est appelé la matrice de sortie de la couche cachée du réseau neuronal; la i th colonne de \mathbf{H} est la sortie du i th nœud caché par rapport aux entrées .

4.2.2. Algorithme

Le plus simple algorithme d'apprentissage ELM a un modèle de la forme

$$\hat{\mathbf{Y}} = \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{x}) \quad (\text{III.8})$$

Où \mathbf{W} est la matrice des pondérations d'entrée-à-couche cachée, σ est une fonction d'activation, et \mathbf{W} est la matrice des pondérations de la couche de sortie.

Étant donné un ensemble de formation

$$\mathfrak{N} = \{ (\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, \dots, N \}, \quad (\text{III.9})$$

onction d'activation g , et le nombre de nœuds cachés L ,

- Attribuer des vecteurs de poids d'entrée aléatoire ou centres et nœud caché partialité ou facteur d'impact $\mathbf{b}_i, i = 1, \dots, L$.
- Calculer la matrice de sortie de la couche cachée \mathbf{H}
- Calculer le poids de sortie $\boldsymbol{\beta}: \boldsymbol{\beta} = \mathbf{H}^{-1} \mathbf{T}$

Avec \mathbf{H}^{-1} est l'inverse de Moore-Penrose généralisé de la matrice de sortie masquée de la couche \mathbf{H} [27].

4.2.3. Les caractéristiques principales

- La vitesse d'apprentissage d'ELM est extrêmement rapide d'où vient le mot extrême
- Contrairement aux algorithmes classiques traditionnels d'apprentissage basé sur les gradients qui travaillent uniquement pour les fonctions d'activation différentiables. l'ELM tend à atteindre les solutions simples avec une méthode plus simple.
- L'algorithme d'apprentissage ELM semble beaucoup plus simple que beaucoup algorithmes d'apprentissage: les réseaux neuronaux et machines à vecteurs de support[28].

4.3. Les réseaux de neurones a base radiale

Les réseaux de neurones à base radiale sont une classe particulière des réseaux de neurones multicouches. [F.Yang, M.Paindavoine, 2005]

4.3.1. Présentation des réseaux RBF

L'idée générale des réseaux RBF dérive de la théorie d'approximation des fonctions, ces réseaux sont une architecture Feedforward puissante. ce type de réseaux a été introduit

pour la première fois par Hardy, et la théorie correspondante a été développée par Powell, ensuite, ces réseaux ont pris le terme de réseaux de neurones grâce à Broomhead et Lowe. Sans oublier les œuvres de MOODY et DARKEN (1989) d'une part, et de POGGIO et GIROSI (1990) d'autre part. La raison de son application vient du fait que le réseau utilise des fonctions gaussiennes standard qui sont à symétrie radiale. son apprentissage est basé sur l'algorithme K-means et l'algorithme des moindres carrés.

Les réseaux de neurones RBFs, sont principalement utilisés pour résoudre des problèmes d'approximation de fonctions dans des espaces de grandes dimensions. Ils sont lus adaptés, en raison d'apprentissage local.

Ce type d'apprentissage peut rendre le processus d'entraînement bien plus rapide que dans le cas d'un MLP, qui apprend de façon globale.

4.3.2. Architecture générale d'un réseau RBF:

Pour des raisons de simplicité, on a décidé de faire une petite dualité entre le réseau RBF et le PMC, en précisant les ressemblances et les différences entre les deux types de réseaux.

Ce choix est justifié par la popularité des PMCs et leur vaste utilisation dans les applications industrielles. un réseau de neurone de type RBF est un PMC spéciale, son architecture est déontique à celle d'un PMC a une seule couche cachée donc on peut dire qu'il prend toutes les caractéristiques d'un PMC simple sauf qu'il diffèrent en quelques points nous citons quelques uns :

❖ Le nombre des couches cachées :

Un réseau RBF ne peut contenir qu'une seule couche cachée, son architecture est fixée pour tous les problèmes à étudier.

❖ La fonction d'activation :

Le réseau RBF utilise toujours une fonction dite à base radiale centrée d'un point et munie d'un rayon.

❖ Les poids synaptiques :

Les poids entre la couche d'entrée et la couche cachée dans les modèles neuronaux de type RBF sont toujours d'une valeur d'unité, c'est-à-dire que l'information inscrite sur la couche d'entrée sera retransmise sans distorsion vers les neurones de la couche cachée.

En ce qui concerne les ressemblances entre un réseau RBF et un PMC, on peut mentionner quelques points :

❖ **La fonction de sortie :**

❖ Généralement une simple fonction linéaire qui renvoie une sommation pondérée des valeurs calculées par les neurones de la couche cachée. bien sur, ce n'est pas toujours le cas, parfois l'utilisation d'autres fonctions pourrait être plus adéquate dans un problème donné.

❖ **Le sens des connexions :**

Les connexions entre les couches suivent le même sens, on peut dire qu'elles ne sont pas récurrentes, et chaque neurone est entièrement connecté vers les neurones de la couche suivante.

❖ **L'apprentissage :**

Pour calculer les poids de la couche de sortie, on utilise un apprentissage super visé pour les deux types de réseaux.

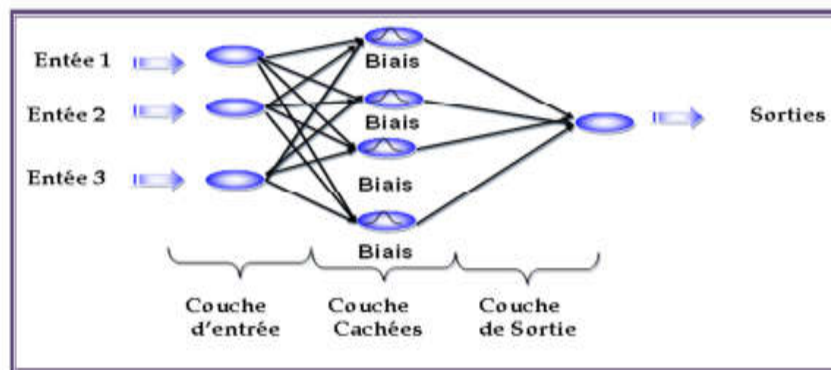


Figure III.6: Architecture d'un Réseau de Neurone RBF

4.3.3. La fonction à base radiale

Les fonctions à base radiale (ou RBF), sont apparus à la fin des années 80 comme de nouvelles variantes des réseaux de neurones. divers types de fonctions peuvent être utilisées comme noyau ou fonction de base, la fonction gaussienne reste cependant la plus utilisée. [A.Seghouane, G.Fleury, 2003], la fonction **gaussienne** ne répond qu'à une petite région de l'espace d'entrée, région sur laquelle elle est centrée. sa forme générale est donnée par :

$$H(y) = \exp(-|y-c|^2/2\mu^2) \quad (\text{III.10})$$

Avec y est une entrée scalaire, et $\mu > 0$.

Les paramètres de cette fonction sont le centre C et le rayon noté par μ .

La raison essentielle du choix de la gaussienne comme fonction de base des RBF est que cette fonction est factorisable. en effet, on peut facilement démontrer que parmi toutes les fonctions de base, la gaussienne est la seule fonction qui peut être décomposée en produit de fonctions gaussiennes unidimensionnelles.

$$H(\mathbf{y}) = \prod_i \exp(-|\mathbf{y} - \mathbf{c}_i|^2 / 2\mu_i^2)$$

Avec $\mathbf{y} = [y_i]$ et $\mathbf{C} = [C_i]$. (III.11)

Cette particularité devient intéressante pour l'adéquation biologique des réseaux de neurones artificiels.

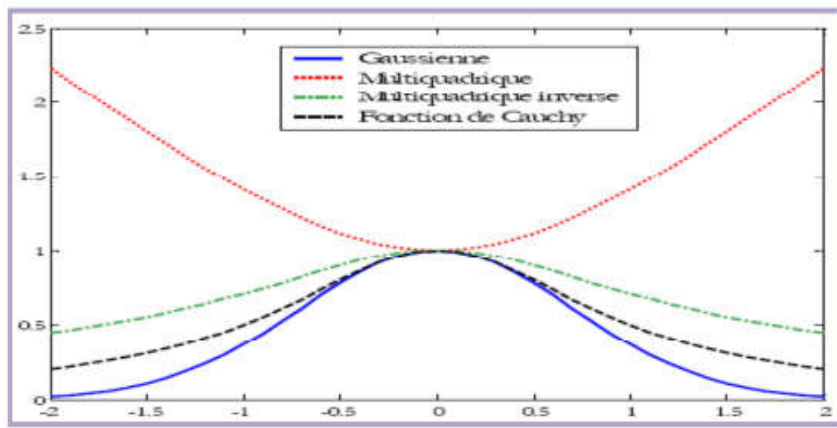


Figure III.7: Quelques fonctions radiales

4.3.4.Phase d'apprentissage :

Plusieurs méthodes d'apprentissage ont été développées pour ces types de réseaux, on dit souvent que l'apprentissage du réseau RBF est hybride : non supervisé pour l'étape de construction du réseau, et supervisé pour la détermination des poids de la couche de sortie.lors de l'étape de construction du réseau, il y a quatre paramètres principaux à régler :

- ❖ Le **nombre** de neurones RBF.
- ❖ La **position** des centres des gaussiennes de chacun des neurones.
- ❖ La **largeur** de ces gaussiennes.
- ❖ Le **poids** des connexions entre les neurones RBF et le(s) neurone(s) de sortie.

a. L'étape de construction du RBF :(non supervisé)

Toute modification d'un de ces paramètres entraîne directement un changement du comportement du réseau.

Le **nombre** de neurones RBF (n) et la **position** des gaussiennes sont deux paramètres intimement liés.

Nombre et position avec $n < I$:

n : le nombre de neurone.

I : le nombre d'exemple soumis au réseau.

Dans ce cas-ci, le nombre de neurones RBF devient un véritable paramètre. il n'existe pas de méthode pour le déterminer. il s'agit donc de trouver le nombre de centroïdes adéquat lié au problème donné.

Une fois le nombre de centroïdes choisi, il faut déterminer leur position. pour ce faire, il existe aussi plusieurs techniques parmi ces techniques on utilise la méthode k-means.

Une fois tous les centres C_j choisis, il faut déterminer la **largeur** (β) des gaussiennes.

Une règle empirique consiste à prendre :

$\beta =$ avec $M =$ nombre des centrioles et $d = \max ||c_i - c_j||$, $1 \leq i$, $j \leq M$ Si on choisit un β égal pour toutes les gaussiennes.

Mais rien n'impose de prendre la même valeur de β pour chaque centroïde. dans ce cas, une autre règle nous dit : $\beta_j = ||x_i - c_j||$

b. L'étape de détermination des poids du RBF:(supervisé)

Une fois le nombre et la position des centrioles et la largeur des gaussiennes fixés, les poids de chacune des connexions (RBF-output) peuvent être calculé par l'équation matricielle suivante : [4], [5]

$$\begin{bmatrix} f(||x_1 - c_1||) & \dots & f(||x_1 - c_M||) \\ f(||x_2 - c_1||) & \dots & f(||x_2 - c_M||) \\ \dots & \dots & \dots \\ f(||x_N - c_1||) & \dots & f(||x_N - c_M||) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_M \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} \tag{III.12}$$

Avec $f(x) = EXP(x^2 / (2 * \beta^2))$, la gaussienne.

La matrice colonne :

Y: les outputs attendus

M: le nombre de centrioles.

N: le nombre d'exemples dans la base d'apprentissage.

L'équation précédente est de la forme : $AW=Y$ avec **A** est une matrice $M*N$.

Pour une première opération, on peut simplifier cette équation par :

$$W = A^{-1} y$$

On est obligé de calculer l'inverse de la matrice A par la méthode classique. Mais la plupart des cas rencontrés, cette matrice est inversable on recherche donc a minimiser : $\|AW - Y\|$

On n'a en général aucune certitude sur le conditionnement de la matrice A, et il vaut donc mieux employer une méthode robuste d'estimation. la méthode la plus utilisé est le calcul de la matrice pseudo inverse de a la méthode nommée SVD est typiquement une bonne méthode pour ce genre de calcul.

On cherche la matrice A^+ la pseudo inverse de la matrice $A[M*N]$.

Si A est une matrice M fois N, alors il existe deux matrices orthogonales U et V :

$$U = \{u_1, u_2, u_3, \dots, u_n\}, v = \{v_1, v_2, v_3, \dots, v_n\}$$

$$\text{Tel que : } U^T A V = \text{diag}(\delta_1, \delta_2, \dots, \delta_k)$$

Avec : diag matrice diagonale et $k = \min(M, N)$.

Le pseudo inverse de A est alors donnée par :

$$A^+ = V \Sigma^+ U^T \text{ avec } \Sigma^+ = \text{diag}(1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_k, 0, \dots, 0) \quad \text{(III.13)}$$

c. L'algorithme de rétropropagation du gradient

L'algorithme de rétropropagation a été développé pour entrainer des perceptrons multicouches. l'idée est d'entrainer un réseau en propageant les erreurs de sorties en sens inverse du réseau a travers les couches. ces erreurs servent alors à évaluer les dérivés de la fonction d'erreur en fonction des poids qui peuvent ainsi être ajustés.

Le but de l'algorithme est de minimiser la fonction de cout suivante :

$$C = \frac{1}{2} \sum (\hat{y}_n - y_n)^2 \quad \text{(III.14)}$$

Avec les y_n sont les sorties désirées pour chaque exemple d'apprentissage x_i , et les \hat{y}_n sont les sorties réelles obtenues par le réseau, et S_0 est l'ensemble des sorties.

La modification des poids de la connexion reliant le jème neurone de la couche cachée

$$\Delta W_{ij} = -\eta_{ij} \frac{\partial E}{\partial W_{ij}} \quad \text{(III.15)}$$

et le i^{eme} neurone de la couche de sortie est donnée par :

η représente un petit nombre positif nommé le pas d'apprentissage .on calcule la quantité suivante : [H.Robaye, 2006], [4]

$$\frac{\partial C}{\partial w_{ij}} = \frac{1}{2} \sum_{n \in S_0} \frac{\partial (y_n - \hat{y}_n)^2}{\partial w_{ij}}$$

$$\begin{aligned}
&= \frac{1}{2} \frac{(\partial(y_i - \hat{y}_i))^2}{\partial w_{ij}} \\
&= -(y_i - \hat{y}_i) \frac{\partial \hat{y}_i}{\partial w_{ij}}
\end{aligned} \tag{III.16}$$

On calcule le terme suivant a part :

$$\begin{aligned}
\frac{\partial \hat{y}_i}{\partial w_{ij}} &= \frac{\partial f_i(\alpha_i)}{\partial w_{ij}} \\
&= \frac{\partial f_i(\alpha_i)}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial w_{ij}} \\
&= f'_i(\alpha_i) \frac{\partial \sum_{i \in s_j} w_{ij}}{\partial w_{ij}} \hat{y}_i \\
&= f'_i(\alpha_i) \hat{y}_i
\end{aligned} \tag{III.17}$$

Tel que $f'(a_i)$ représente la dérivée de la fonction d'activation au point a_i , et

$$a_i = \sum_j w_{ij} x_j.$$

On définit la quantité suivante : $\delta_i = (y_i - \hat{y}_i) f'_i(a_i)$

Et on écrit la variation du poids qui définit la règle d'apprentissage :

$$\Delta w_{ij} = \eta \delta_i \hat{y}_j$$

Signalons enfin que le processus sera terminé lorsque le taux d'erreur est acceptable (inférieur a un seuil défini préalablement) après un nombre fini d'itérations[29].

d. La Phase de Test :

Enfin, une fois l'apprentissage effectué, la partie test peut commencer. elle se résume aussi en une équation matricielle :

$$[f(\|t_i - c_i\|) \dots f(\|t_i - c_M\|)] \begin{bmatrix} w_i \\ \dots \\ w_M \end{bmatrix} = \text{Output_test}$$

Avec $f(x) = \text{EXP}(x^2 / (2 * B^2))$, la gaussienne.

La matrice colonne W, les poids des connexions (RBF-output), M, le nombre des centrioles. t_i , le $i^{\text{ème}}$ élément de la 'base de test'.

4.4. Avantages et inconvénients des réseaux RBF

Alors que l'on pouvait redouter d'avoir inventé un modèle aux belles propriétés théoriques, mais inutilisable en raison de durées d'apprentissage prohibitives, c'est tout le

contraire qui se produit : la construction d'un réseau RBF est rapide et facile, et c'est là le principal avantage de la technique .

Mais cette avantage se paye par des performances qui ne peuvent être aussi bonnes que celles de techniques plus sophistiquées (comme le Perceptron Multicouches).

En particulier, les réseaux RBF sont peu performants:

- ❖ Sur les données dans des espaces de grande dimension (beaucoup de variables d'entrée). cette faiblesse est propre à toutes les techniques locales.

- ❖ Sur des données très bruitées. la reconstruction locale de la fonction empêche le réseau de "moyenner" le bruit sur tout l'espace (comparez avec la Régression Linéaire, dont l'objectif est justement de moyenner le bruit sur les données).

5. Apprentissage

Supposons que l'on dispose d'une base d'apprentissage de taille n d'observations $(x_i^1, \dots, x_i^p, y_i)$ des variables explicatives X^1, \dots, X^p et de la variable à prévoir Y . considérons le cas le plus simple de la régression avec un réseau constitué d'un neurone de sortie linéaire et d'une couche à q neurones dont les paramètres sont optimisés par moindres carrés. ceci se généralise à toute fonction perte dérivable et donc à la discrimination à m classes. L'apprentissage est l'estimation des paramètres $\alpha_j=0;p;k=1;q$ et $\beta_k=0;q$ par minimisation de la fonction perte quadratique (ou d'une fonction d'entropie en classification) :

$$Q(\alpha, \beta) = \sum_{i=1}^n Q_i = \sum_{i=1}^n [y_i - f(x; \alpha, \beta)]^2 \quad \text{(III.18)}$$

Différents algorithmes d'optimisation sont proposés, ils sont généralement basés sur une évaluation du gradient par rétro-propagation.

5.1. Rétropropagation de l'erreur

Il s'agit donc dévaluer la dérivée de la fonction coût en une observation et par rapport aux différents paramètres.

$$\text{Soit } z_{ki} = g(\alpha k' + \alpha k' x_i) \text{ et } z_i = \{z_{1i}, \dots, z_{qi}\}. \quad \text{(III.19)}$$

Les dérivées partielles de la fonction perte quadratique s'écrivent

$$\begin{aligned} \frac{\partial Q_i}{\partial \beta_k} &= -2(y_i - \phi - (x_i)) (\beta' z_i) z_{ki} = \delta_i z_{ki} \\ \frac{\partial Q_i}{\partial \alpha_{kj}} &= -2(y_i - \phi(x_i)) (\beta' z_i) \beta_k f'(\alpha' x_i) x_{ip} = \delta_{ki} x_{ip} \end{aligned} \quad \text{(III.20)}$$

Les termes δ_i et ski sont respectivement les termes d'erreur du modèle courant à la sortie et sur chaque neurone caché. ces termes d'erreur vérifient les équations dites de rétropropagation :

$$s_{ki} = g'(\alpha'_k x_i) \beta_k \delta_i \quad (\text{III.21})$$

dont les termes sont évaluée sen deux passes. une passe avant, avec les valeurs courantes des poids : l'application des différentes entrées x_i au réseau permet de déterminer les valeurs ajustées $f(x_i)$.

5.2. Algorithmes d'optimisation

Sachant évaluer les gradients, différents algorithmes, plus ou moins sophistiqués, sont implémentés. le plus élémentaire est une utilisation itérative du gradient : en tout point de l'espace des paramètres, le vecteur gradient de Q pointe dans la direction de l'erreur croissante. pour faire décroître Q il suffit donc de se déplacer en sens contraire. il s'agit d'un algorithme itératif modifiant les poids de chaque neurone selon :

$$\begin{aligned} \beta_k^{(r+1)} &= \beta_k^{(r)} - \tau \sum_{i=1}^n \frac{\partial Q_i}{\partial \beta_k^{(r)}} \\ \alpha_{kp}^{(r+1)} &= \alpha_{kp}^{(r)} - \tau \sum_{i=1}^n \frac{\partial Q_i}{\partial \alpha_k^{(r)}} \end{aligned} \quad (\text{III.22})$$

Le coefficient de proportionnalité τ est appelé le taux d'apprentissage. il peut être fixe, à déterminer par l'utilisateur, ou encore varier en cours d'exécution selon certaines heuristiques.

Il paraît en effet intuitivement raisonnable que, grand au début pour aller plus vite, ce taux décroisse pour aboutir à un réglage plus fin au fur et à mesure que le système s'approche d'une solution. si l'espace mémoire est suffisant, une version accélérée de l'algorithme fait intervenir à chaque itération un ensemble (*batch*) d'observations pour moyennner les gradients et mises à jour des poids. Bien d'autres méthodes d'optimisation ont été adaptées à l'apprentissage d'un réseau : méthodes du gradient avec second ordre utilisant une approximation itérative de la matrice hessienne (algorithme BFGS, de Levenberg Marquardt) ou encore une évaluation implicite de cette matrice par la méthode :

Initialisation des poids b_{jk} par tirage aléatoire selon une loi uniforme sur $[0; 1]$.

Normaliser dans $[0; 1]$ les données d'apprentissage.

while $Q > \text{errmax}$ ou $\text{niter} < \text{itermax}$ **do**

Ranger la base d'apprentissage dans un nouvel ordre aléatoire.

for chaque élément $i = 1, \dots, n$ de la base **do**

Calculer $\epsilon(i) = y_i - f(x_i^1, \dots, x_i^p; (b)(i-1))$ en propageant les entrées vers l'avant.

L'erreur est rétro-propagée dans les différentes couches afin d'affecter à chaque entrée une responsabilité dans l'erreur globale.

Mise à jour de chaque poids $b_{jk}'(i) = b_{jk}'(i-1) + \Delta b_{jk}'(i)$

end for

end while

Tableau III.3: Algorithme 1 Rétropropagation élémentaire du gradient

dite du gradient conjugué. la littérature sur le sujet propose quantités de recettes destinées à améliorer la vitesse de convergence de l'algorithme ou bien lui éviter de rester collé à une solution locale défavorable. d'autres heuristiques proposent d'ajouter un terme d'inertie afin d'éviter des oscillations de l'algorithme.

D'autres algorithmes encore sont des versions adaptatives. lorsque de nouvelles observations sont proposées une à une au réseau. dans ce dernier type d'algorithme, des propriétés de dynamique markovienne (processus ergodique convergeant vers la mesure stationnaire) impliquent une convergence presque sûre : la probabilité d'atteindre une précision fixée a priori tend vers 1 lorsque la taille de l'échantillon d'apprentissage tend vers l'infini. on pourra se reporter à l'abondante littérature sur le sujet pour obtenir des précisions sur les algorithmes d'apprentissage et leurs nombreuses variantes. il est important de rappeler la liste des choix qui sont laissés à l'utilisateur. en effet, même si les logiciels proposent des valeurs par défaut, il est fréquent que cet algorithme connaisse quelques soucis de convergence.

6. L'organisation des neurones en réseaux

Les caractéristiques essentielles des réseaux de neurones réels que nous conservons dans les modèles informatiques étudiés, concernent le grand nombre de connections, la non-linéarité des relations entrée-sortie et la plasticité ou l'adaptabilité. les réseaux de neurones artificiels peuvent être définis ainsi : ensemble d'automates du type neurone formel, reliés entre eux par des interconnexions à poids variables.

Les modèles diffèrent les uns des autres par leurs architectures ou topologies et leurs règles d'adaptation ou d'apprentissage.

Conclusion

Nous venons de voir dans ce chapitre différents modèles de réseaux neurones, Autrement dit, une vue globale ou description sur les réseaux de neurones artificiels. par la suite nous avons donné une présentation générale sur quelques types de réseaux de neurones existants comme les réseaux de neurones ELM et RBF.

Contrairement aux méthodes traditionnelles qui ont montré leurs frontières, les réseaux de neurones ont montré leur tendance à s'adapter à des problèmes complexes, leur grande capacité de calcul et d'apprentissage. ils font l'objet d'une utilisation dans différents domaines tels que: la vision artificiel, le traitement de données, la robotique.....etc.

CHAPITRE IV : Résultats et discussion

Sommaire

Introduction.....	55
1.série temporelle.....	55
1.1. Utilisation.....	55
2. Présentation de la base de données.....	55
2.1. Caractéristiques des site.....	55
3.application de la technique MRMR avec les réseaux de neurones type RBF et ELM.....	56
3.1. Etape de test.....	57
4. Les résultats de simulation des deux sites en utilisant le logiciel MATLAB	57
4.1. La prédiction par la méthode de ELM.....	58
4.2. La prédiction par la méthode de RBF.....	62
Conclusion.....	66

Introduction

L'objectif de ce Chapitre est de présenter les résultats de simulation obtenus à l'aide du logiciel MATLAB sur les bases de données, dans le présent travail et de les commentés.

Les travaux de ce mémoire consistent à étudier des techniques pour la prédiction de la vitesse du vent, en utilisant la méthode de réduction de la dimension MRMR et les réseaux de neurones artificiels de type RBF et ELM en régression, ceci en se basant les critères de performances suivants: la racine carrée de l'erreur quadratique moyenne (RMSE), l'erreur moyenne de pourcentage absolue (MAPE) et le Coefficient de détermination (R).

1. Série temporelle

Une Série temporelle ,ou série chronologique ,est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps .de telles suites de variables aléatoires peuvent être exprimées mathématiquement afin d'en analyser le comportement, généralement pour comprendre son évolution passée et pour en prévoir le comportement futur.

1.1. Utilisation

Les Série temporelle sont considérées à tort comme étant une branche exclusive de l'économétrie. Cette dernière est une discipline qui est relativement jeune alors que les Série temporelles ont été utilisées bien avant ,par exemple en astronomie (1906) et en météorologie (1968).

2. Présentation de la base de données

Pour faire la prédiction du vent Nous avons pris deux bases de données pour deux sites au Etats-Unis à partir du site web du laboratoire national des énergies renouvelables d'États-Unis « NREL » (National Renewable Energy Laboratory).

2.1. Caractéristiques des sites

Le premier site se situe dans l'état de Colorado (est un état de l'Ouest des États-Unis), les coordonnées du site sont (41° 6' " N, 72° ' " W).

Le deuxième site se situe dans l'état de Connecticut (est une état du Nord-est des États Unis), les coordonnées du site sont (39° ' " N, 105° ' " W).

Site Number State 1322	LON	LAT	ELEV	AREA (sq.km)	DENS (MW/sq.km)	Total MW for site
Colorado	-102.52	37.39	1299	39.8	10.8	428.1

Tableau IV. 1: Caractéristiques géographiques du site Colorado

Site Number State 6028	LON	LAT	ELEV	AREA (sq.km)	DENS (MW/sq.km)	Total MW for site
Connecticut	-73.46	42.03	562	5.8	17.1	99.9

Tableau IV.2 : Caractéristiques géographiques du site Connecticut

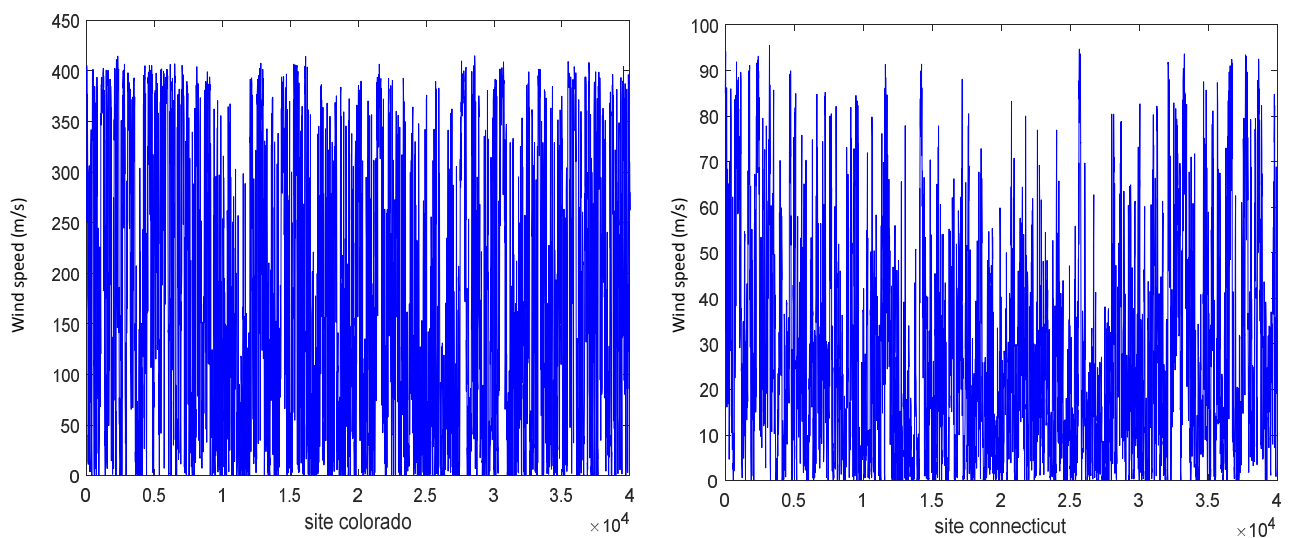


Figure IV.1 : La courbe réelle avec les séries temporelles des deux sites (base de données).

3. Application de la technique MRMR avec les réseaux de neurones type Extreme Learning machine (ELM) et Radial Bases Function (RBF)

Il existe un grand nombre de types de réseaux de neurones, dont chacun a des avantages et des inconvénients. Dans notre étude nous avons choisis deux types de réseau de neurones Extreme Learning Machine (ELM) et Radial Bases Function (RBF). Ce choix est

fait pour la facilité et la rapidité de leurs constructions et encore par le fait que notre problème présente un nombre limité de variables d'entrées.

3.1. Etape de test

Pour cette étape pour évaluer ce modèle, on a estimé l'erreur sur l'ensemble de test, en faisant une comparaison avec les valeurs réelles et prédites de la base données avec les deux critères Erreur Quadratique Moyenne (RMSE), Erreur Moyenne de Pourcentage Absolue (MAPE) et le Coefficient de Détermination (R).

➤ Racine carrée de l'erreur quadratique moyenne RMSE:

Le RMSE multiplie chaque erreur par elle-même (le carré de l'erreur), donnant ainsi un poids plus grand aux grandes erreurs qu'aux petites erreurs [24].

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (4.1)$$

y_t et \hat{y}_t sont respectivement les valeurs réelles et prédites et N est le nombre d'observations.

➤ Erreur moyenne de pourcentage absolue MAPE :

Cette erreur est une quantité souvent utilisé pour mesurer l'écart des prédictions et des mesures. Sa formule est donnée par :

$$MAPE = (|x - \hat{x}|) = \mathit{mean} [y_t - \hat{y}_t] / m (y_t) \quad (4.2)$$

$$MAPE = \frac{1}{N_f} \sum_{t=1}^{N_f} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100 \quad (4.3)$$

➤ Coefficient de détermination (R²) :

Est un indicateur qui permet de juger la qualité de régression linéaire, simple ou multiple. D'une valeur comprise entre 0 et 1, il mesure l'adéquation entre le modèle et les données observées (réelles). Dans le cadre d'une régression linéaire simple, c'est le carré du coefficient de corrélation. Le définit comme la part de variance exprimée par rapport a la variance totale, autrement dit par [1-(somme des carrés des résidus /variance totale)]

$$R^2 = 1 - \frac{v (y_t - \hat{y}_t)}{v (y_t)} \quad (4.4)$$

4. Les résultats de simulation pour les deux sites en utilisant le logiciel

Matlab:

Dans les deux tableaux suivants nous avons présenté les résultats de la prédiction de la vitesse du vent par rapport RBF et ELM avant et après MRMR.

4.1. La Prédiction par la technique ELM

Site Colorado				
	Neurones ELM	RMSE	MAPE	R
Avant MRMR	50 Neurones	26.8655	0.1034	95.6505
	100 Neurones	19.4284	0.0759	97.7101
	500 Neurones	15.7525	0.0588	98.4167
Après MRMR	50 Neurones	25.9007	0.1026	95.8858
	100 Neurones	19.5223	0.0745	97.7479
	500 Neurones	16.2846	0.0536	98.5107

Tableau IV.3: Critères de performances Pour les techniques ELM et ELM-MRMR Pour le site Colorado

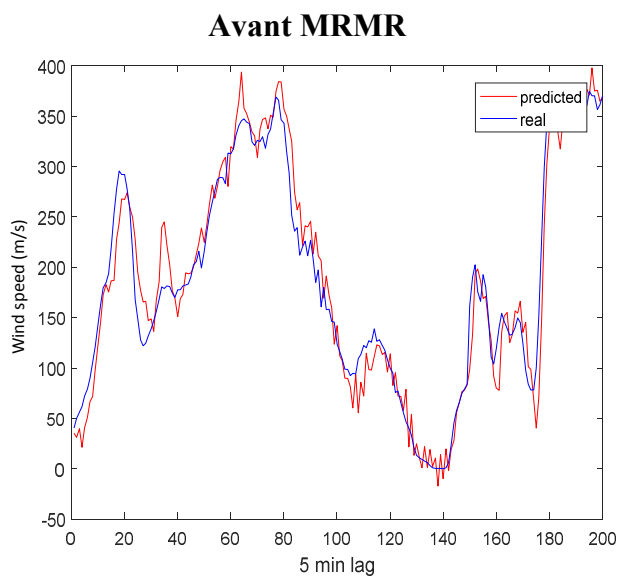


Figure IV.2: Courbe comparaison valeurs réels et prédites de ELM pour 50 neurones site Colorado.

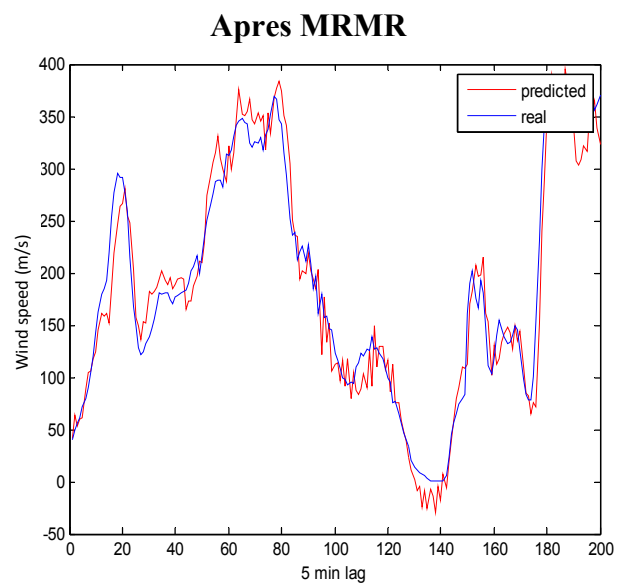


Figure IV.3: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 50 neurones site Colorado.

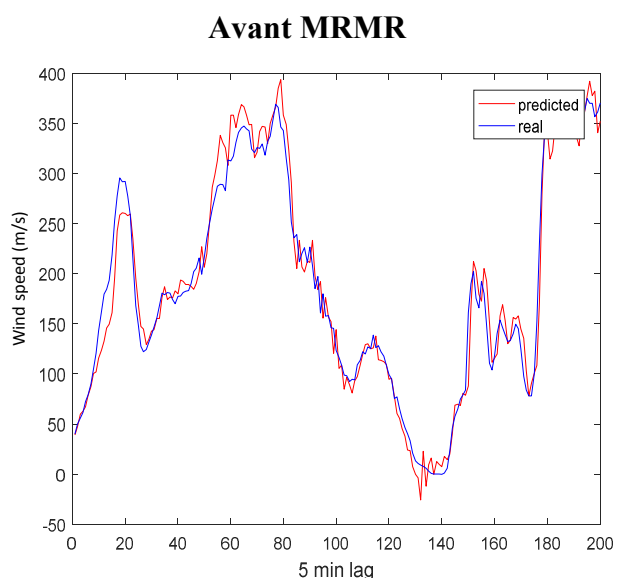


Figure IV.4: Courbe comparaison valeurs réels et prédites de ELM pour 100 neurones site Colorado.

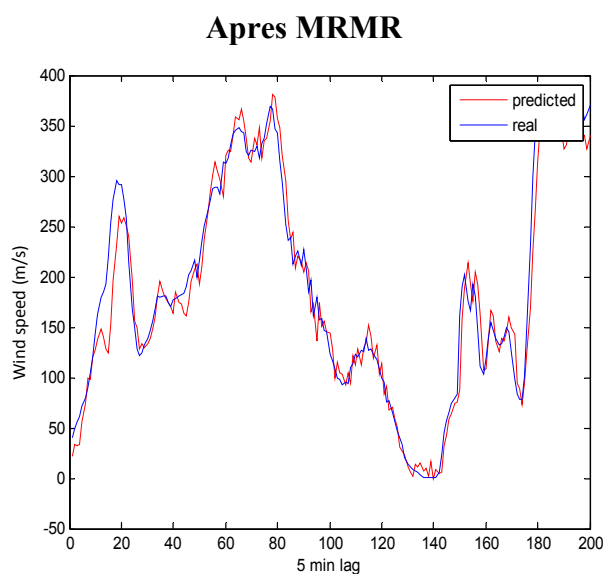


Figure IV.5: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 100 neurones site Colorado.

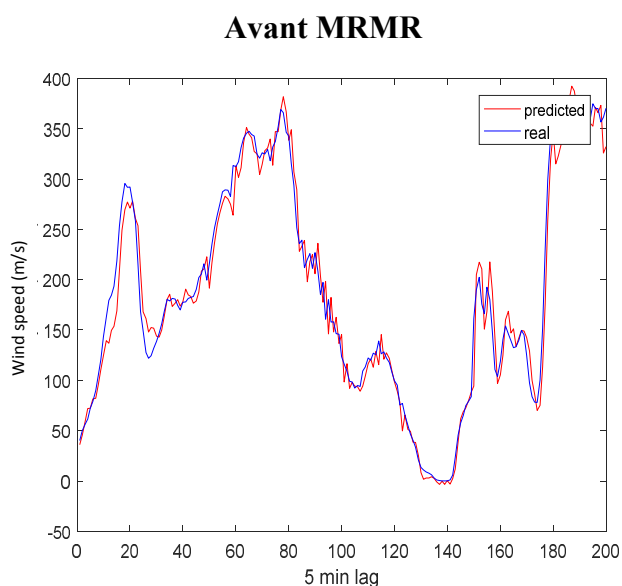


Figure IV.6: Courbe comparaison valeurs réels et prédites de ELM pour 500 neurones site Colorado.

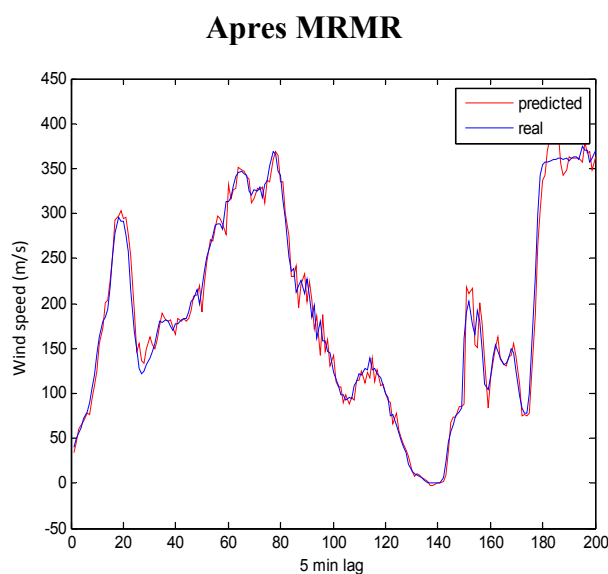


Figure IV.7: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 500 neurones site Colorado.

Commentaire sur le résultat :

Pour le site Colorado d'après les figures (IV.2), (IV.3), (IV.4), (IV.5), (IV.6), (IV.7) on constate tout d'abord que la courbe prédite par les deux techniques ELM et ELM-MRMR

s'applique sur la courbe réelle, et aussi d'après les mesures des critères de performances indiquées dans le Tableau (IV.3) on peut conclure que cette technique nous a donnée des bons résultats et qui sont améliorés par l'utilisation de la technique de la sélections des caractéristiques MRMR.

Site Connecticut				
	Neurones ELM	RMSE	MAPE	R
Avant MRMR	50 Neurones	4.7925	0.1125	96.8240
	100 Neurones	4.1548	0.0959	97.6008
	500 Neurones	4.0128	0.0873	97.8400
Après MRMR	50 Neurones	4.5076	0.1008	97.1774
	100 Neurones	3.4290	0.0724	98.4172
	500 Neurones	3.5311	0.0755	98.3308

Tableau IV.4: Critères de performances Pour les techniques ELM et ELM-MRMR Pour le Site Connecticut

Avant MRMR

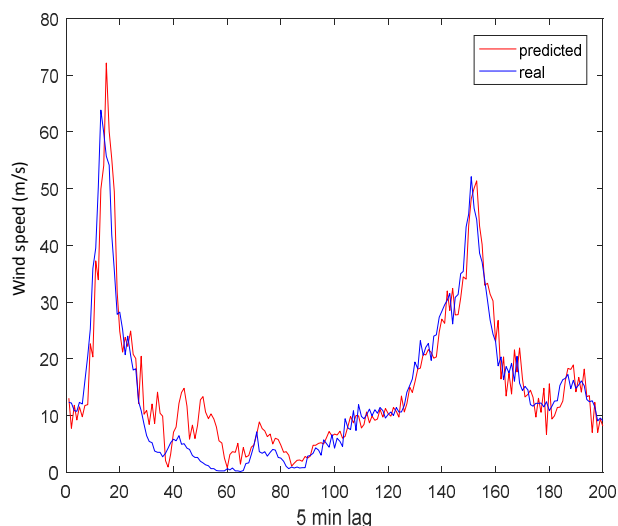


Figure IV.8: Courbe comparaison valeurs réels et prédites de ELM pour 50 neurones pour le site Connecticut.

Après MRMR

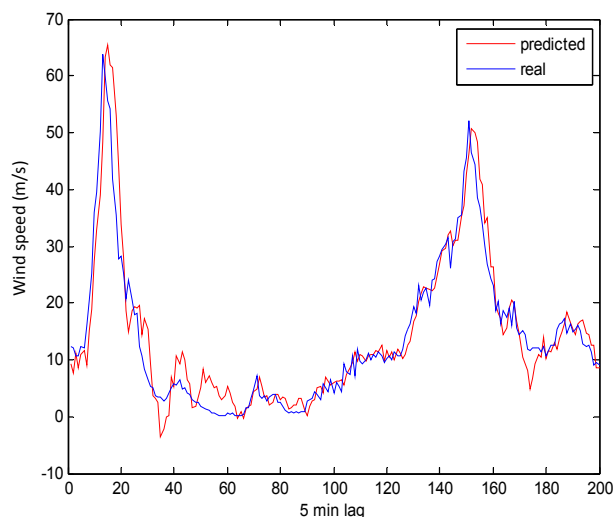


Figure IV.9: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 50 neurones site Connecticut.

Avant MRMR

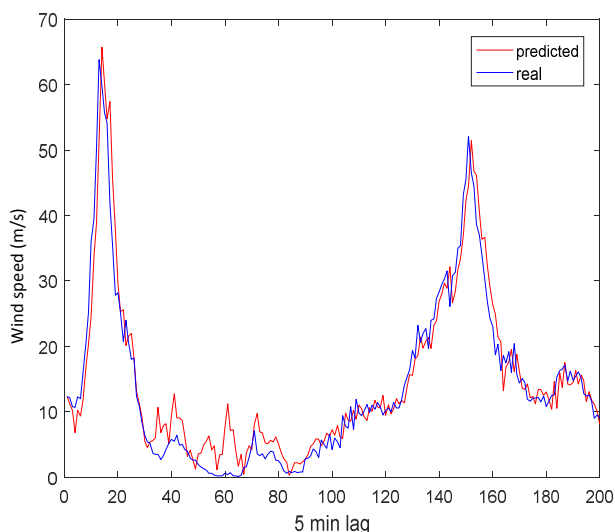


Figure IV.10: Courbe comparaison valeurs réels et prédites de ELM pour 100 neurones pour le site Connecticut

Après MRMR

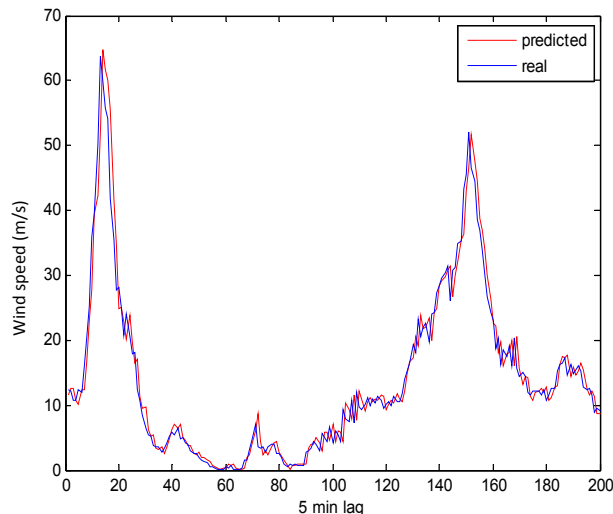


Figure IV.11: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 100 neurones site Connecticut.

Avant MRMR

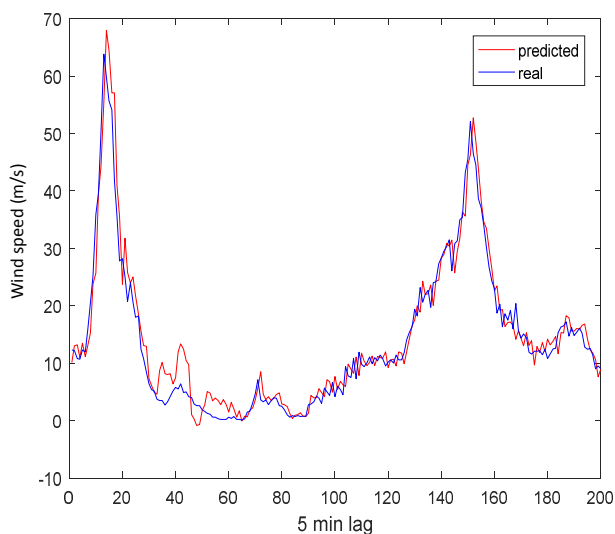


Figure IV.12: Courbe comparaison valeurs réels et prédites de ELM pour 500 neurones pour le site Connecticut

Après MRMR

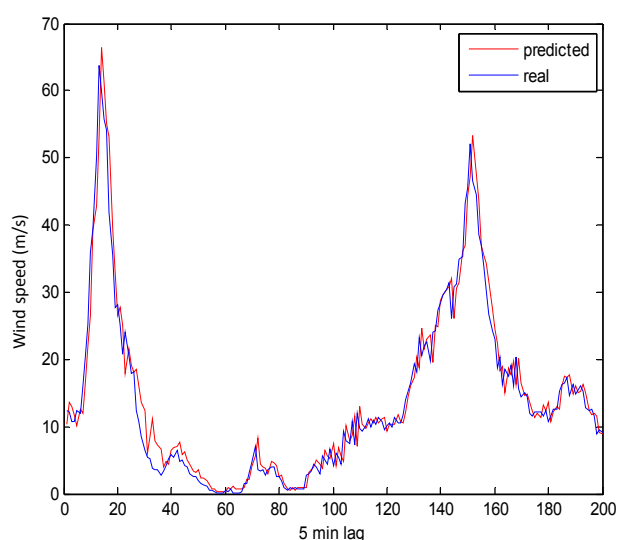


Figure IV.13: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 500 neurones site Connecticut.

Commentaire sur le résultat :

Pour le site Connecticut d'après les figures (IV.8), (IV.9), (IV.10), (IV.11), (IV.12), (IV.13) on constate tout d'abord que la courbe prédite par les deux techniques ELM et ELM-

MRMR s'applique sur la courbe réelle, et aussi d'après les mesures des critères de performances indiquées dans le Tableau (IV.4) on peut conclure que cette technique nous a donnée des bons résultats et qui sont améliorés par l'utilisation de la technique de la sélections des caractéristiques MRMR.

4.2. La prédiction par la technique RBF

Site Colorado				
	Neurones RBF	RMSE	MAPE	R
Avant MRMR	10 Neurones	17.9834	0.9984	98.0326
	50 Neurones	13.2568	0.9993	98.9462
	100 Neurones	13.2568	0.9993	98.9462
Après MRMR	10 Neurones	13.2085	0.0460	98.9545
	50 Neurones	13.1850	0.0458	98.9574
	100 Neurones	13.1855	0.0458	98.9572

Tableau IV.5: Critères de performances Pour les techniques RBF et RBF-MRMR Pour le Site Colorado

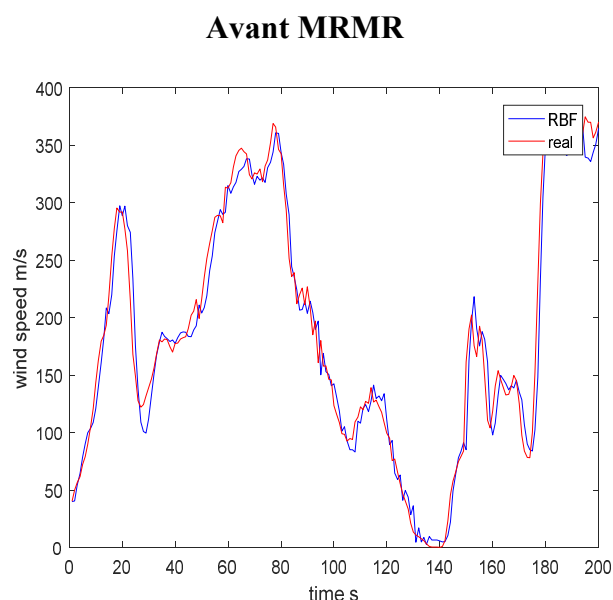


Figure IV.14 : Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Colorado.

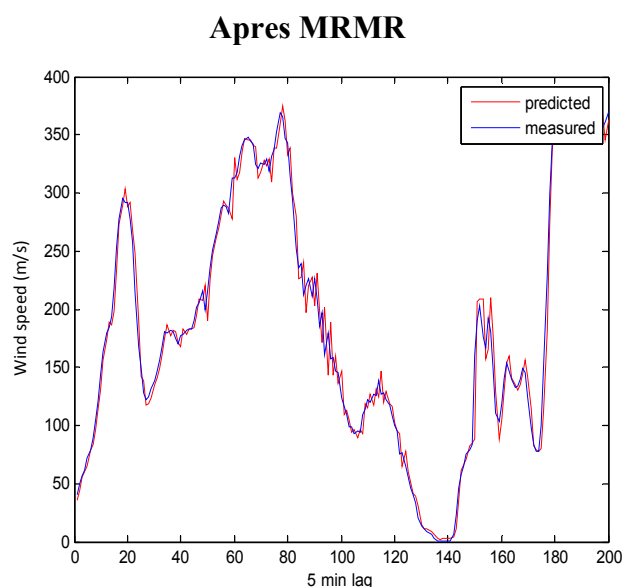


Figure IV.15 : Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 10 neurones pour le site Colorado.

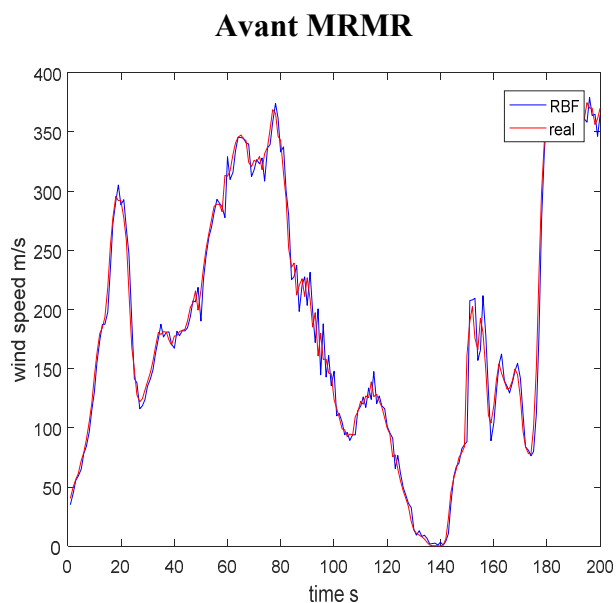


Figure IV.16: Courbe comparaison valeurs réels et prédites de RBF pour 50 neurones pour le site Colorado.

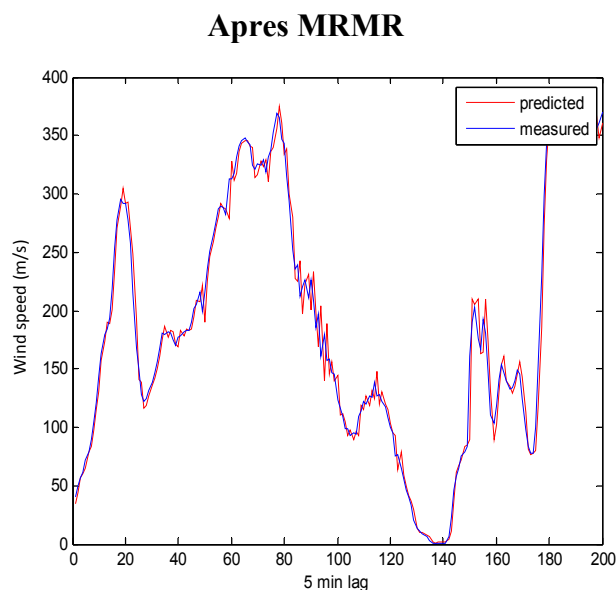


Figure IV.17: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 50 neurones pour le site Colorado.

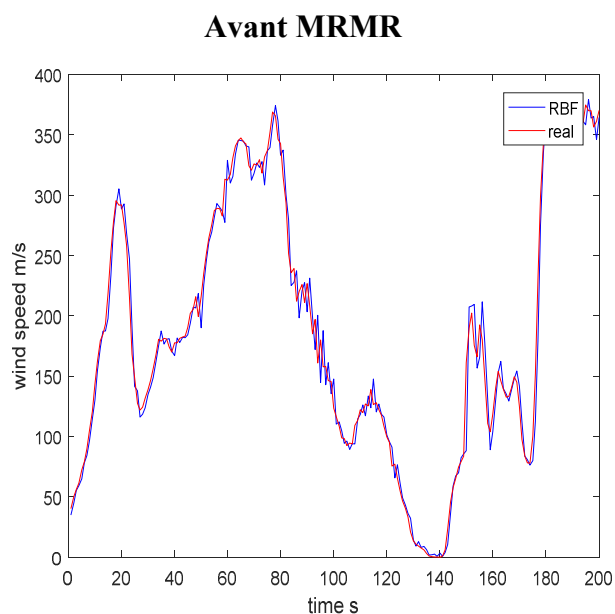


Figure IV.18: Courbe comparaison valeurs réels et prédites de RBF pour 100 neurones pour le site Colorado.

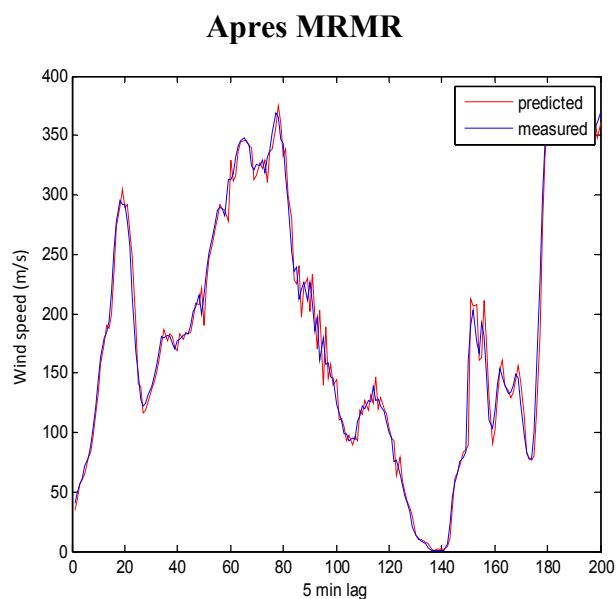


Figure IV.19: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 100 neurones pour le site Colorado.

Commentaire sur le résultat :

Pour le site Colorado d'après les figures (IV.14), (IV.15), (IV.16), (IV.17), (IV.18), (IV.19) on constate tout d'abord que la courbe prédite par les deux techniques RBF et RBF-MRMR s'applique sur la courbe réelle, et aussi d'après les mesures des critères de

performances indiquées dans le Tableau (IV.5) on peut conclure que cette technique nous a donnée des bons résultats et qui sont améliorés par l'utilisation de la technique de la sélection des caractéristiques MRMR.

Site Connecticut				
	Neurones RBF	RMSE	MAPE	R
Avant MRMR	10 Neurones	3.5772	1.0018	98.2737
	50 Neurones	13.2598	0.9993	98.4455
	100 Neurones	3.4521	1.0017	98.4040
Après MRMR	10 Neurones	3.3760	0.0712	98.4707
	50 Neurones	3.3966	0.0715	98.4527
	100 Neurones	3.3899	0.0713	98.4603

Tableau IV.6: Critères de performances Pour les techniques RBF et RBF-MRMR Pour le Site Connecticut.

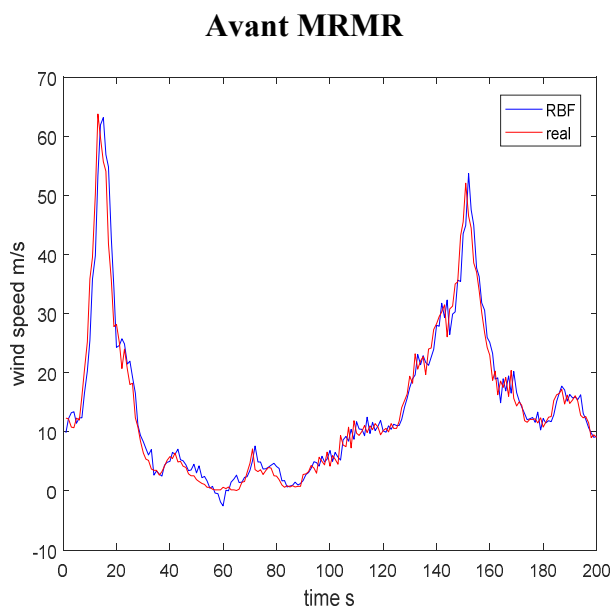


Figure IV.20: Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Connecticut.

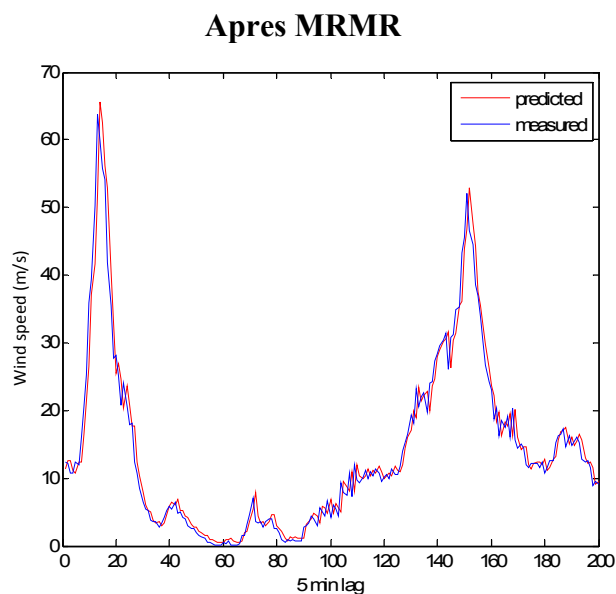


Figure IV.21: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 10 neurones pour le site Connecticut.

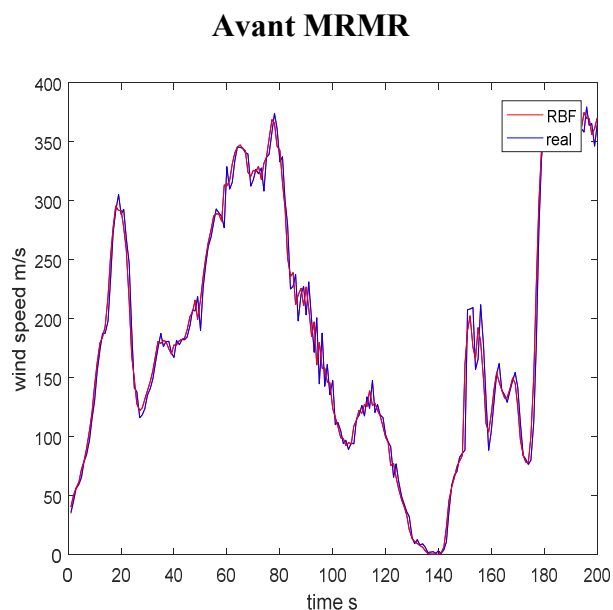


Figure IV.22: Courbe comparaison valeurs réels et prédites de RBF pour 50 neurones pour le site Connecticut.

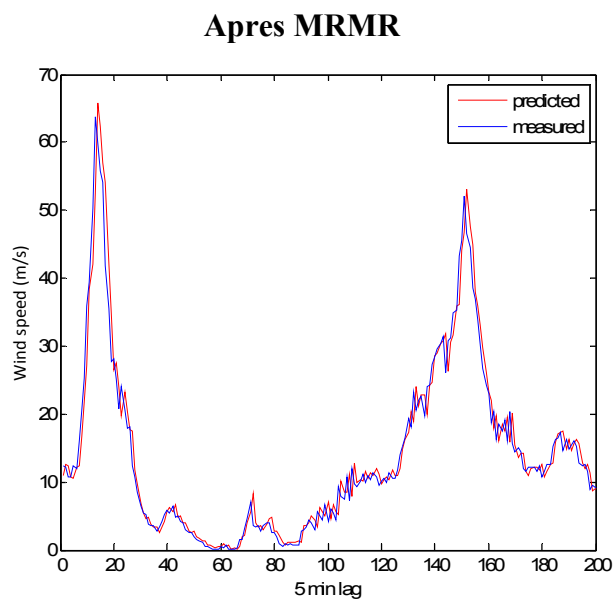


Figure IV.23: Courbe comparaison valeurs réels et prédites de RBF et réduction MRMR pour 50 neurones pour le site Connecticut.

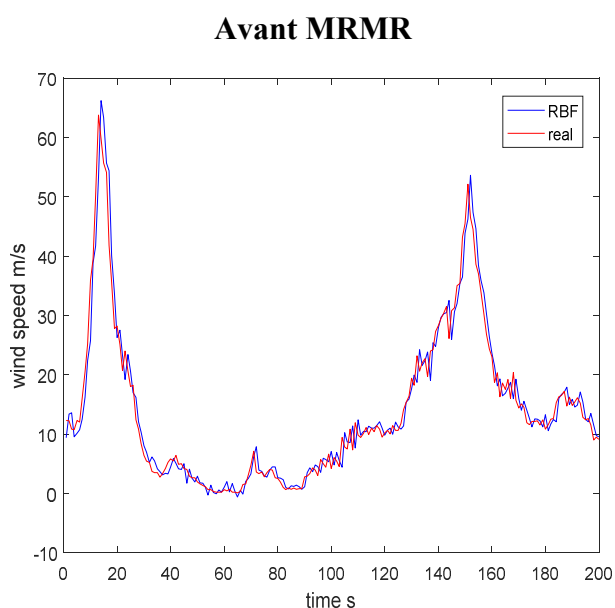


Figure IV.24: Courbe comparaison valeurs réels et prédites de ELM pour 100 neurones pour le site Connecticut.

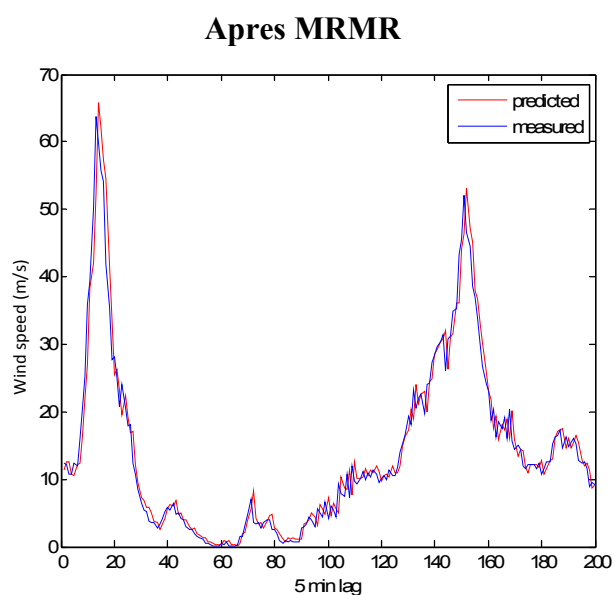


Figure IV.25: Courbe comparaison valeurs réels et prédites de ELM et réduction MRMR pour 100 neurones pour le site Connecticut.

Commentaire sur le résultat :

Pour le site Connecticut d'après les figures (IV.20), (IV.21), (IV.22), (IV.23), (IV.24), (IV.25) on constate tout d'abord que la courbe prédite par les deux techniques ELM et ELM-MRMR s'applique sur la courbe réelle, et aussi d'après les mesures des critères de performances indiquées dans le Tableau (IV.6) on peut conclure que cette technique nous a donnée des bons résultats et qui sont améliorés par l'utilisation de la technique de la sélections des caractéristiques MRMR.

Conclusion

Dans ce chapitre, nous avons tenté d'établir une synthèse globale sur le travail qui a été réalisé dans ce mémoire.

On a vu et discuté les résultats obtenus pour la prédiction de la vitesse du vent, où en a proposé une technique basé sur la réduction de la dimension qui est le MRMR et aussi deux types de réseaux de neurones ELM et RBF pour la régression.

En général, l'application de ces techniques nous a permis d'améliorer les critères de performance et la qualité de prédiction par rapport à l'utilisation de la base complète.

Nous avons utilisé la technique MRMR avec les réseaux de neurone ELM et RBF on utilise la technique MRMR pour l'obtention des valeurs des critères d'erreurs « RMSE, MAPE, et R » sur la base complète avec une comparaison entre les courbes réels et prédite pour les deux sites étudiés.

CONCLUSION GENERALE

Conclusion Générale

En guise de conclusion générale, nous allons tenter d'établir une synthèse globale sur le travail qui a été réalisé dans ce mémoire. L'objectif principal consistait à développer un modèle de calcul basé sur la méthode MRMR à deux modèle neuronal en « MATLAB », à savoir : ELM (Extrême Learning Machines) et RBF pour prédire la vitesse du vent pour deux sites.

D'une manière générale, l'utilisation de ces méthodes nous ont permit d'obtenir des valeurs futures bien précises concernant la vitesse du vent et par conséquence une meilleur prédiction de la valeur de la vitesse du vent.

On applique, en « MATLAB », la technique ELM et RBF avec la technique MRMR pour l'obtention des valeurs des critères d'erreurs « RMSE et MAPE » sur la base complète avec une comparaison entre le courbe réel et prédite pour les deux sites étudiés.

D'une manière générale, l'utilisation de ces méthodes nous permit d'obtenir des valeurs futures bien précises concernant la vitesse du vent et par conséquence une meilleur prédiction de la valeur de la vitesse du vent.

PERSPECTIVES ET FUTURES TRAVAUX

Les principaux objectifs à atteindre dans une perspective à court terme de notre étude. seront:

1. Prédire la vitesse du vent en utilisant d'autres méthodes d'intelligence artificielle (méthode à noyaux,...etc.).
2. Etendre cette études à la prédiction d'autres paramètres météorologiques, ou bien des séries appartenant à d'autres domaines (finance, commerce, biomédicale ,...etc.).

Références bibliographiques

Références Bibliographiques

- [1] A. Tiedemann ,S. Oehlschlaeger, "Intégration des énergies renouvelables au réseau," *renac (renewable Academy)*, pp. 1- 36, Berlin, 2013,
- [2] R. Dixon, M. Eckhart, D. Hales, and G. Thompson, "Rapport mondial 2012 sur les énergies renouvelables 2012," *Renewable Energy Pollicy Networkfor the 21st Century*, pp. 1-20, 2012.
- [3] O. Edenhofer, R. Pichs-Madruga, Y. Sokona " Rapport spécial sur les sources d ' énergie renouvelable et l ' atténuation du changement climatique," *Institut de recherche de Potsdam sur les effets du changement climatique*, pp. 1-226, 2011.
- [4] J MARTIN, « Energies éoliennes », Techniques de l'Ingénieur, Traités Energétique B8585, 1-, 22p,199
- [5] F. Poitiers «Étude et commande de génératrices asynchrones pour l'utilisation de l'énergie éolienne-Machine asynchrone à cage autonome-Machine asynchrone à double alimentation reliée au réseau» Thèse de Doctorat Université de Nantes, 2003.
- [6] R. Swisher, C. Real de Azua, J. Clendenin, « Strong winds on the horizon: wind power comes of age », proceedings of the IEEE, vol. 89, N°. 12, December 2001, pp. 1757-1764.
- [7] L. H. Hansen, L. Helle, F. Blaabjerg, E. Ritchie, S. Munk-Nielsen, H. Bindner, P. Sorensen and B. Bak-Jensen, « Conceptual survey of generators and power electronics for wind turbines », Riso national laboratory, Roskilde, Denmark, december 2001.
- [8] L. H. Hansen, L. Helle, F. Blaabjerg, E. Ritchie, S. Munk-Nielsen, H. Bindner, P. Sorensen and B. Bak-Jensen, Report « Conceptual survey of generators and power electronics for wind turbines », Riso national laboratory, Roskilde, Danemark, december 2001.
- [9] S El Aimani, « Modélisation des différentes technologies d'éoliennes intégrées dans un réseau de moyenne tension », thèse de doctorat, l'université des sciences et technologies de Lille (USTL).
- [10] site de constructeur d'éoliennes : <http://www.enercon.de/>;
- [11] Site web : <http://www.planetoscope.com>
- [12] E. Muljadi, C.P. Butterfield, Yih-Huei Wan, «Axial flux, modular, permanent-magnet generator with a toroidal winding for wind turbine applications », IEEE industry applications conference, St Louis, november 5-8, 1998.
- [13] H. Ben Ahmed, « Généralités sur l'énergie éolienne et convertisseurs associés et convertisseurs associés », séminaire électrotechnique de l'académie de Rennes, 9 mars 2004.

Références Bibliographiques

- [14] C. Haritza, « Minimisation de l'impact des perturbations d'origine éolienne dans la génération d'électricité par des aérogénérateurs à vitesse variable ». Doctorat, E.N.S.A.M Paris, décembre 2003.
- [15] S El Aïmani, « Modélisation des différentes technologies d'éoliennes intégrées dans un réseau de moyenne tension », thèse de doctorat, l'université des sciences et technologies de Lille (USTL).
- [16] Dash, M. et Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis*
- [17] John, G. H. (1997). Enhancements to the data mining process. Thèse de doctorat, Stanford, CA, USA. UMI Order No. GAX97-23376.
- [18] John, G. H., Kohavi, R. et Pfleger, K. (1994). Irrelevant features and the subset selection problem. In *MACHINE LEARNING : PROCEEDINGS OF THE ELEVENTH INTERNATIONAL*, pages 121. Morgan Kaufmann.
- [19] Fraser, A. M. et Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. 33.
- [20] Aha, D. W. et Bankert, R. L. (1995). A comparative evaluation of sequential feature selection algorithms. In *5th International Workshop on Artificial Intelligence and Statistics*, pages 1. Ft. Lauderdale, USA.
- [21] Reconnaissance de visages par Analyse Discriminante Linéaire(LDA) par Amira Chiheb Université du 8 mai 45 de Guelma Algérie - Licence en informatique 2003
- [22] Jain, A. et Zongker, D. (1997). Feature selection : Evaluation, application, and small sample performance. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 19:153.
- [23] Réseaux de neurones gif-21140 et gif-6432 par Marc Parizeau 2004 université laval
- [24] H. Bensaïd, « The Algerian Programme on Wind Energy ». *Proceeding of WEAC*, pp. 21-27, Oxford, 1985
- [25] Norbet Wiener. *Cybernetics; or, control and communication in the animal and the machine*. Wiley, Published by the M.I.T.press cambridge,massachusetts. 1948
- [26] Guang-Bin Huang, Qin-Yu Zhu et Chee-Kheong Siew, « Extreme learning machine: theory and applications », *Neurocomputing*, vol. 70, no 1,2006, p. 489–501
- [27] site de elm :<http://www.ntu.edu.sg/home/egbhuang/>
- [28] G.-B Huang, et al “Can Threshold Networks Be Trained Directly?” *IEEE Transactions on Circuits and Systems II*, vol. 53, no. 3, pp. 187-191, 2006

Références Bibliographies

[29]N. Kasbadji merzouk , « An Evaluation of Wind Energy Potential in Algeria ». Proceeding of EWEC'94 congress, Thessaloniki, Grèce, 1994.

Résumé :

Le développement et l'exploitation des énergies renouvelables ont connu une forte augmentation ces dernières années .d'ici 20-30 ans , tout système énergétiques durable sera basé sur l'utilisation raisonnable des sources traditionnelles et sur un recours aux énergies renouvelable.

Naturellement décentralisées, il est intéressant de les exploiter sur le lieu de consommation, en les transformant directement soit en chaleur , soit en électricité selon les besoins.

La production d'électricité décentralisée par sources d'énergie renouvelable offre une plus grande sureté d'approvisionnement des consommateurs tout en respectant l'environnement. cependant le caractère aléatoire de ces sources nous impose d'établir des règles d'utilisation de ces systèmes pour les exploiter au mieux.

Pour cette raison l'objectif de notre travail de développer un model de calcul basé sur la technique **MRMR** appliqué à un modèle intelligent qui est les **Réseaux de Neurones Artificiels (ELM et RBF)** Pour prédire la vitesse du vent.

Mots clé: Sélection des caractéristiques, Réseau de Neurone Artificiel ,Vitesse du Vent.

Abstract :

The development and exploitation of renewable energy has increased sharply in recent years. By 20-30 years, while sustainable energy system will be based on the reasonable use of traditional sources and on renewable energy.

Naturally decentralized, it is interesting to use them on the place of production, by directly transforming them to heat or electricity as needed. The decentralized electricity production by renewable energy sources provides greater security of supply to consumers while respecting the environment.

However the random nature of these sources requires us to establish rules for using these systems to exploit them.

For this reason the objective of this thesis is to develop a computational model based on **MRMR** technique applied to a neural model **Artificial Neural Network (ELM and RBF)** to predict the wind speed with MATLAB software.

Keywords: Feature selection , Neural Networks Artificial , wind speed .

ملخص

لقد عرفت الطاقات المتجددة تطورا ملحوظا بعد عديد من الاكتشافات في السنوات الأخيرة إذ يقدر أنه في 20-30 سنة القادمة المستدامة ستعتمد على الاستخدام المنطقي للطاقة الاحفورية وعودة الطاقات المتجددة . المثير للاهتمام في الطاقات المتجددة هي اللامركزية فالاهتمام منصب على التشغيل في مكان الإستهلاك عن طريق التحويل المباشر إلى طاقة حرارية أو كهربائية حسب الحاجة. إن إنتاج الكهرباء اللامركزية عن طريق مصادر الطاقة المستدامة يوفر أمن أكبر للمستهلكين و ذلك باحترام البيئة ولكن الطبيعة العشوائية للمصادر يفرض علينا تأسيس قواعد لاستعمال والتحكم بهذه الطاقة بشكل أفضل لذا فالهدف من هذه المذكرة هو تطوير نموذج حسابي يعتمد على تقنية **MRMR** (الصلة القصوى- تقليل التكرار) مطبقة على نموذج عصبي للتنبؤ بسرعة الرياح باستعمال ماتلاب .

الكلمات الرئيسية: تقنية **MRMR** (الصلة القصوى- تقليل التكرار) , طاقة الرياح , الشبكات العصبية الاصطناعية .