

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

FACULTE DE TECHNOLOGIE
DEPARTEMENT D'ELECTRONIQUE



N° :

DOMAINE : SCIENCES ET
TECHNOLOGIE

FILIERE : ELECTRIQUE

OPTION : MICRO ELECTRONIQUE

Mémoire présenté pour l'obtention
Du diplôme de Master Académique

Par: BELHOU ASMA

Intitulé

**LDA ET RNA POUR LA PREDECTION DE LA
VITESSE DU VENT**

Soutenu devant le jury composé de:

Nom et prénom Enseignant

Université

Président

Mezaache Hatem

Université de M'sila

Rapporteur

.....

Université

Examineur

Année universitaire : 2016 /2017

Remerciements

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

Je tiens à adresser mes plus vifs remerciements à Mon encadreur **Mezaache Hatem** et Aussi Je remercie aussi **Dr Hassen Bouzgou** qui m'a aidé à terminer ce travail.

Aussi étendre mes remerciements à M^{me} **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.

Je 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.

BELHOU ASMA

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 A Mon mari Ali Salamani

À mes frères Hicham, Youcef, Mohamed

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

Table de figure

Table de tableau

Liste des symboles et abréviations

Introduction Générale.....	1
----------------------------	---

Chapitre 01 : L'énergie éolienne

1.1. INTRODUCTION A L'ENERGIE EOLIENNE	3
1.1.1.Introduction.....	3
1.1.2.Historique de l'éolien.....	3
1.2. L'ENERGIE EOLIENNE EN QUELQUES CHIFFRES.....	4
1.2.1.Production mondiale d'énergie éolienne	4
1.2.2.Les plus grands producteurs d'énergie éolienne dans le monde en 2015.....	4
1.2.3.La capacité mondiale de production d'énergie éolienne dans le monde.....	5
1.2.4.La production et l'historique d'énergie éolienne en Algérie	6
1.3. DEFINITION DE L'ENERGIE EOLIENNE.....	6
1.3.1.Descriptif d'une éolienne	8
1.3.2.Principaux Composants d'une Eolienne	10
1.3.3.Principe de fonctionnement d'une éolienne	12
1.3.4.Les différents types d'éoliennes	14
1.3.4.1.Les éoliennes à axe vertical.....	14
1.3.4.2.Les éoliennes à axe horizontal.....	15
1.3.5.Classification des turbines éoliennes	17
1.3.6.Taille des Aérogénérateurs	18
1.3.7.Les avantages et les inconvénients d'énergie éolienne.....	18
1.3.7.1.Les avantages d'énergie éolienne.....	18
1.3.7.2.Les inconvénients d'énergie éolienne	20
1.4. ENERGIE CINETIQUE DU VENT – CONVERSION EN ENERGIE MECANIQUE.....	21
1.4.1.Loi de BETZ.....	21
1.5 CONCLUSION.....	25

Chapitre 02 : la réduction et la régression du dimension

2.1 INTRODUCTION	26
2.2 LA REDUCTION DE DIMENSION.....	26
2.2.1. Définition.....	26
2.2.2. Méthodes de réduction de dimension	27
2.2.1. L'Analyse en Composantes Principales.....	27
2.2.2. Factorisation de Matrice Non Négative.....	28

2.2.3	Analyse en Composantes Indépendantes.....	28
2.2.4.	Le modèle discriminant linéaire amélioré de Fisher (EFM).....	29
2.2.5	Analyse Discriminante Linéaire(LDA)	30
2.2.5.1	présentation:.....	30
2.2.5.2	La règle bayésienne:.....	31
2.2.5.3	Taux d'erreur.....	31
2.2.5.4	. Les avantages et inconvénients	32
2.3	Régression.....	32
2. 3.1	Histoire.....	32
2. 3.2	Définition.....	33
2. 3.3	Principaux modèles de régression.....	33
2. 3.4	Applications.....	34
2. 3.5	Les reseaus de neurons.....	34
2. 3.5 .1	Historique.....	34
2. 3.5 .2	Utilité.....	35
2. 3.5 .3	Le neurone formel.....	36
2. 3.5 .4	Caractéristiques d'un réseau de neurones.....	38
2.3.5.5	Principe de fonctionnement.....	38
2.3.5.6	Le Perceptron Multi Couches (M.Minsky and S.Papert, 1969).....	39
2.3.5.6 .1	Architecture.....	40
2.3.5.6 .2	L'algorithme de rétro propagation	42
2.3.5.6 .3	Fonction de transfert	45
2.3.5.6 .4	Apprentissage	46
2.3.5.6 .5	Remarques	47
2.3.5.7	Le réseau de neurone fonctions à base radial	48
2.3.5.7.1	Présentation des réseaux RBF.....	48
2.3.5.7.2	générale d'un réseau RBF	49
2.3.5.7.3	La fonction à base radiale	50
2.3.5.7.4	La phase d'apprentissage	51
2.3.5.7.5	L'étape de construction du RBF :(non supervisé)	51
2.3.5.7.6	L'étape de détermination des poids du RBF:(supervisé)	51
2.3.5.7.7	L'algorithme de rétro-propagation du gradient.....	52
2.3.5.7.8	La phase de test	52
2.3.5.7.9	Avantages et inconvénients des réseaux RBF.....	52
2.3.5.8	Comparaison MLP et RBF.....	53
2.4	CONCLUSION	54

Chapitre 03: Résultat et discussion

3.1	INTRODUCTION	55
3.2	PRESENTATION DE LA BASE DE DONNEES	55
3.2.1	Caractéristiques des sites	55

3.3.APPLICATION DE LA TECHNIQUE LDA AVEC LES RESEAUX DE NEURONES TYPE MLP et RBF	57
3.3.1 Etape apprentissage et validation sur la base complète	57
3.3.2 Etape de test	59
3.4. DISCUSSIONS DES RESULTATS	60
3.4.1 Les Résultats De Simulation Des Deux Sites En Utilisant Le Logiciel Matlab.....	61
3.4 CONCLUSION.....	73
CONCLUSION GENERAL.....	74
REFERENCES BIBLIOGRAPHIQUES	

Table de figure

Chapitre 01 : L'énergie éolienne

Figure 1. 1 Les formes d'éoliennes à travers l'histoire.....	4
Figure 1. 2 Production d'énergie éolienne dans le monde en 2013	5
Figure 1. 3 La ferme éolienne d'Adrar	6
Figure 1. 4 Conversion de l'énergie cinétique du vent	7
Figure 1. 5 Principaux organes du système de conversion éolien	7
Figure 1. 6 Conversion de l'énergie cinétique du vent	8
Figure 1. 7 Schéma d'une éolienne de type aérogénérateur	9
Figure 1. 8 Tour d'une turbine	10
Figure 1. 9 Rotor d'une turbine	11
Figure 1. 10 Constitution d'une nacelle	11
Figure 1. 11 Eoliennes à axe vertical.....	15
Figure 1. 12 Eoliennes Horizontal.....	15
Figure 1. 13 Principales composantes de l'éolienne à axe horizontal.....	16
Figure 1. 14 Taille des hélices en [m] et puissance en [kW]	18
Figure 1. 15 Coefficient de puissance en fonction rapport de vitesse du vent.....	23

Chapitre 02 : Techniques de Réduction de la dimension et la techniques de Régressions

Figure 2. 1 Projection par LDA avec séparation entre les classes.....	30
Figure 2. 2 Le neurone formel.....	36
Figure 2. 3 Principales fonctions d'activation.....	37
Figure 2. 4 Propagation de l'influx nerveux dans un réseau de type perceptron.....	38
Figure 2. 5 Exemple de perceptron multicouche élémentaire avec une couche cachée et une couche de sortie.....	41
Figure 2. 6 Table de vérité de la porte logique XOR.....	41
Figure 2. 7 Le perceptron multicouche.....	42
Figure 2. 8 L'algorithme de la rétro propagation de gradient.....	45
Figure 2. 9 structure d'un RBF	48
Figure 2. 10 Architecture d'un Réseau de Neurone RBF.....	50
Figure 2. 11 Quelques fonctions radiales.....	50

Chapitre 03 : RESULTATS ET DISCUSSION

Figure 3. 1 les courbes de la base de données pour les deux sites.....	56
Figure 3. 2 Les courbe d'erreur de validation Base complète pour les deux sites.....	58
Figure 3. 3: Courbe comparaison valeurs réels et prédites de MLP pour 1 couche cachée	

Table de figure

site Colorado.....	62
Figure 3. 4 Courbe comparaison valeurs réels et prédites de MLP pour 2 couche cachée site Colorado	62
Figure3. 5 Courbe comparaison valeurs réels et prédites de MLP pour 3 couche cachée site Colorado	63
Figure 3. 6 Courbe comparaison valeurs réels et prédites de RBF pour 20 neurones site Colorado.....	63
Figure 3.7 : Courbe comparaison valeurs réels et prédites de MLP pour 1 couche cachée pour le site Colorado.....	64
Figure 3.8 : Courbe comparaison valeurs réels et prédites de MLP pour 2 couche cachée pour le site Colorado.....	65
Figure 3.9 : Courbe comparaison valeurs réels et prédites de MLP pour 3 couches cachées pour le site Colorado.....	65
Figure 3.10 : Courbe comparaison valeurs réels et prédites de MLP pour 1 couche cachée pour le site Connecticut.....	66
Figure 3.11 : Courbe comparaison valeurs réels et prédites de MLP pour 2 couches cachées pour le site Connecticut	67
Figure 3.12 : Courbe comparaison valeurs réels et prédites de MLP pour 3 couches cachées pour le site Connecticut	68
Figure 3.13 : Courbe comparaison valeurs réels et prédites de RBF pour 30 neurones pour le site Colorado.....	69
Figure 3.14 : Courbe comparaison valeurs réels et prédites de RBF pour 20 neurones pour le site Colorado.....	70
Figure 3.15 : Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Colorado.....	70
Figure 3.16 : Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Connecticut.....	71
Figure 3.17 : Courbe comparaison valeurs réels et prédites de RBF pour 20 neurones pour le site Connecticut.....	72
Figure 3.18 : Courbe comparaison valeurs réels et prédites de RBF pour 30 neurones pour le site Connecticut.....	72

Liste des tableaux

Chapitre 01 : L'énergie éolienne

Tableau 1.1 Classification des turbines éoliennes	17
--	----

Chapitre 02 : Techniques de Réduction de la dimension et la techniques de Régressions

Tableau 2.1 MLP versus RBF.....	53
--	----

Chapitre 03 :RESULTATS ET DISCUSSION

Tableau 3.1 Caractéristiques géographiques du site Colorado.....	55
---	----

Tableau 3.2 Caractéristiques géographiques du site connecticut.....	55
--	----

Tableau 3.3 Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site colorado avant LDA.....	61
---	----

Tableau 3.4 Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site colorado par MLP.....	64
---	----

Tableau 3.5 Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site conecticut par MLP	66
--	----

Tableau 3.6 Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site colorado par RBF	69
--	----

Tableau 3.7 Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site conecticut par RBF	71
--	----

Notations

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

m	La dimension de l'espace de départ
r	La dimension de l'espace réduit
$n\lambda$	Le nombre de bandes
n_x	Le nombre de pixels en colonnes
n_λ	Le nombre de bandes spectrales S Covariance H L'entropie
Ω	La matrice de transformation
n_y	Le nombre de pixels en lignes V La variance d_{ij} La distance entre deux points dans l'espace d'origine (r dimensions).
δ_{ij}	La distance entre deux points dans l'espace réduite.
H_u	Histogramme globale associé à la bande
$B_u H_v$	Histogramme globale associé à la bande
$B_v P_l$	La distribution de probabilité associée à la bande l.
P_k	La distribution de probabilité associée à la bande k.
MT	La fiabilité
MC	La continuité $C(x,y)$ Le contraste entre deux faisceaux $I(x,y)$ La fonction de comparaison de luminance
σ_{xy}	La similarité par la perte de corrélation entre les deux vecteurs
$K(x,y)$	La fonction de noyau
W_{ij}	Le noyau de chaleur
(x, y, λ)	La matrice qui représente l'image hyperspectrale
$I(x, y, \lambda)$	La matrice moyenne qui représente l'image hyperspectrale
f	Le coefficient de partition
v_f	Indice de Fukuyama and Sugeno
v_x	Indice de Xie and Beni

Liste des Notations et abréviation

v_k	Indice de Kwon
v_T M	Indice de Pakhira
v_S	Indice de Kim
v_C b	Indice de rezaee
DMLE	Maximum de vraisemblance
I	La fonction indicatrice de l'événement A
P	Exposant de proportion
N (r)	Le nombre de hyper sphères de rayon r
S	La sortie
f	La fonction d'activation
O	L'ensemble des états possibles des neurones.
o_i	L'état du neurone i , où $o_i \in O$
f_i	La fonction de transfert associée au neurone i .
s_i	L'entrée du neurone i , si $i \in R$
w_{ji}	Le poids de la connexion du neurone j vers le neurone i
$w_{ji} \in W$	L'ensemble des poids du PMC
x	Le vecteur de longueur d (l'entrée)
V	La matrice $k \times d$ (les poids de la couche d'entrée vers la couche cachée)
c	Le vecteur de longueur k (les biais de la couche cachée)
W	La matrice $m \times h$ (les poids de la couche cachée vers la couche de sortie).
b_k	Le vecteur de biais
W_k	La matrice de poids
y	La classe cible

Abréviations

RNA	Réseaux Neuronaux Artificiels
ACP	L'Analyse en Composantes Principale
NMF	Factorisation de Matrice Non Négative
ACI	Analyse en Composantes Indépendantes

Liste des Notations et abréviation

EFM	Le modèle discriminant linéaire amélioré de Fisher
LDA	Analyse Discriminante Linéaire
PMC	Perceptron Multi Couches
MLP	Multilayer perceptron
RBF	Le réseau de neurone fonctions à base radial
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
<i>MABE</i>	Le biais Moyen Absolue

Introduction Générale

Introduction Générale

Les énergies renouvelables sont des énergies propres inépuisables et ne dégagent pas les gaz à effet de serre .parmi ces énergies, on cite l'énergie éolienne qui transforme cinétique du vent en énergie rotative. L'outil principal de transformation de cette énergie est l'aérogénérateur.

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.

Pourquoi prédire la vitesse du vent ?

La prédiction de la vitesse du vent permet le contrôle et la gestion du réseau en connaissant la quantité d'énergie produite, la protection du système des grandes vitesses et la détermination de la possibilité de construire des centrales éoliennes dans un site en connaissant son potentiel.

L'objectif principal des systèmes de production d'énergie devrait être de répondre à tout moment à la demande d'énergie tout en minimisant les impacts environnementaux associés. Ceci nécessite de développer des formes propres d'énergie, tout en s'assurant de la cohérence et de la fiabilité de leur approvisionnement et de leur utilisation..

Pour cette raison nous proposons de développer un model de calcul basé sur la technique de LDA et Réseau de neurone Artificiel pour prédire la vitesse du vent

ORGANISATION DU MANUSCRIT

Dons ce mémoire, nous proposons la méthodologie de LDA et RESEAU DE NEURONE ARTIFICIEL pour la modélisation et la prédiction de la vitesse du vent en ce basant sue des modèles dites les séries temporelles.

Ce document est constitué de la présente introduction, de trois chapitres et d'une conclusion générale. Il est organisé comme suit :

) Le Chapitre 1: dans ce chapitre nous exposons un aperçu sur la production générale de l'électricité, l'énergie renouvelable et la consommation d'énergie éolienne.

Introduction Générale

-) Chapitre2: on présente les outils essentiels afin de comprendre la technique LDA avec les techniques MLP et RBF.
-) Le Chapitre3: dans ce chapitre on donne des rappels théoriques sur les séries temporelles ainsi que les modèles associés et on présente les résultats de simulation par MATLAB de nos modèles de calcul.

Chapitre 01

L'énergie Eolienne

Sommaire

1.1.1. Introduction	3
1.1.2. Historique de l'éolien	3
1.2. L'ENERGIE EOLIENNE EN QUELQUES CHIFFRE	4
1.2.1. Production mondiale d'énergie éolienne	4
1.2.2. Les plus grands producteurs d'énergie éolienne dans le monde en 2015	4
1.2.3. La capacité mondiale de production d'énergie éolienne dans le monde	5
1.2.4. La production et l'historique d'énergie éolienne en Algérie.....	6
1.3. DEFINITION DE L'ENERGIE EOLIENNE.....	6
1.3.1. Descriptif d'une éolienne	7
1.3.2. Principaux Composants d'une Eolienne	10
1.3.3. Principe de fonctionnement d'une éolienne	12
1.3.4. Les différents types d'éoliennes	14
1.3.5. Classification des turbines éoliennes.....	17
1.3.6. Taille des Aérogénérateurs	18
1.3.7. Les avantages et les inconvénients d'énergie éolienne	18
1.4. ENERGIE CINETIQUE DU VENT – CONVERSION EN ENERGIE MECANIQUE..	21
1.4.1. Loi de BETZ.....	21
1.5. CONCLUSION	25

1.1. INTRODUCTION A L'ENERGIE EOLIENNE.

1.1.1. Introduction

Une éolienne est un instrument qui convertit la puissance du vent en énergie électrique, donc le maximum d'énergie émis ne dépend pas seulement de la machine mais aussi des limites de la vitesse du vent. D'autre part, elles convertissent l'énergie éolienne en énergie mécanique. L'électricité produite par les éoliennes est utilisée pour les circuits de charge de batterie, les systèmes de production d'énergie résidentiels, les systèmes de production distribuée, et les grands réseaux de services publics. Les différents éléments des éoliennes sont conçus pour maximiser cette conversion énergétique.

Plusieurs technologies sont utilisées pour capter l'énergie du vent (capteur à axe vertical ou axe horizontal), les structures de ces capteurs sont de plus en plus performantes.

Dans ce chapitre, on s'intéresse essentiellement aux différents types d'éoliennes avec leurs constitutions et leurs principes de fonctionnement et la conversion d'énergie.

1.1.2. Historique de l'éolien

Parmi toutes les énergies renouvelables, à part l'énergie du bois, c'est l'énergie du vent qui a été exploitée en premier par l'homme. Depuis l'antiquité, elle fut utilisée pour la propulsion des voiliers et ensuite les moulins à blé et les constructions permettant le pompage d'eau. Les premières utilisations connues de l'énergie éolienne remontent à 2 000 ans avant J-C environ.

Hammourabi, fondateur de la puissance de Babylone, avait conçu tout un projet d'irrigation de la Mésopotamie utilisant la puissance du vent. La première description écrite de l'utilisation des moulins à vent en Inde date d'environ 400 ans av. J.-C. En Europe, les premiers moulins à vent ont fait leur apparition au début du moyen âge. Utilisés tout d'abord pour moulinier le grain, d'où leur nom de " moulins ", ils furent aussi utilisés aux Pays-Bas pour assécher des lacs ou des terrains inondés.

Dès le XIV^e siècle, les moulins à vent sont visibles partout en Europe et deviennent la principale source d'énergie. Seulement en Hollande et au Danemark, vers le milieu du XIX^e siècle, le nombre des moulins est estimé respectivement à plus de 30000 et dans toute l'Europe à 200000. A l'arrivée de la machine à vapeur, les moulins à vent commencent leur disparition progressive.

L'arrivée de l'électricité donne l'idée à Paule La Cour en 1891 d'associer à une turbine éolienne une génératrice. Ainsi, l'énergie en provenance du vent a pût être « redécouverte » et de nouveau utilisée (dans les années 1940 au Danemark).

Au début du siècle dernier, les aérogénérateurs ont fait une apparition massive (6 millions de pièces fabriquées) aux Etats-Unis où ils étaient le seul moyen d'obtenir de l'énergie électrique dans les campagnes isolées. Dans les années 60, fonctionnait dans le monde environ 1 million d'aérogénérateurs. Le premier choc pétrolier de 1973 a relancé de nouveau la recherche et les réalisations éoliennes dans le monde de manière très dynamique.

La crise pétrolière de 1973 a relancé de nouveau la recherche et les réalisations éoliennes dans le monde.



Figure 1. 1 : *Les formes d'éoliennes à travers l'histoire*

1.2. L'ENERGIE EOLIENNE EN QUELQUES CHIFFRE

1.2.1. 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 € [1]

1.2.2. Les plus grands producteurs d'énergie éolienne dans le monde en 2015

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 élargi 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, Envision et CSIC. [1]

1.2.3. La 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).

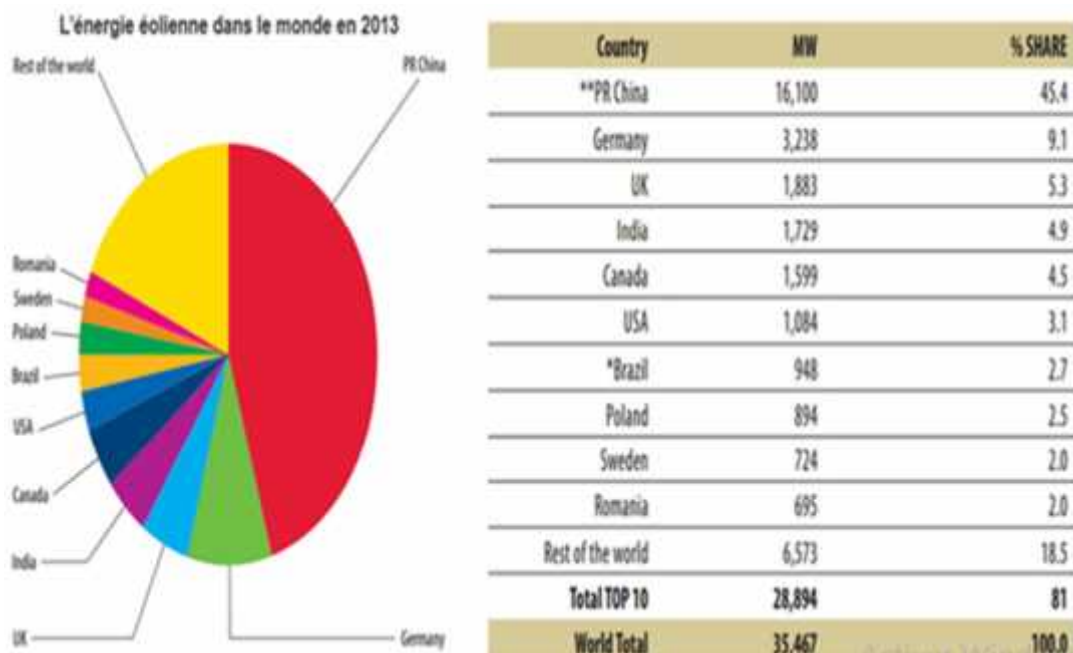


Figure 1. 2 : Production d'énergie éolienne dans le monde en 2013 [1]

1.2.4. La production et l'historique d'énergie éolienne en Algérie

En ce qui concerne l'Algérie, la ressource éolienne varie beaucoup d'un endroit à un autre. Ceci est dû à une topographie et un climat diversifiés. Les premiers travaux Algériens sur le potentiel éolien ont été publiés par Said et all, [2] en 1984, suivis par Benssad, en 1985, [3], Hammouche, en 1990 [4], Kasbadji Merzouk en 1994 [5] et F. Youcef Ettoumi en 2002, [6].

Après plus d'une année de retard, l'Algérie a mis en service sa première ferme éolienne en Adrar, qui doit servir de projet pilote pour le plan éolien du gouvernement de 420MW à l'horizon 2024. La ferme éolienne d'Adrar constituée de 12 éoliennes d'une puissance unitaire de 0.85MW chacune (pour une capacité globale de de 10 mégawatts), implantée sur une superficie de 30 hectares dans la zone de Kabertène à 72 km au nord du chef-lieu de la wilaya. Cette nouvelle centrale de production d'électricité fait partie d'un vaste plan de projets de génération de 22 GW, d'énergie renouvelable à l'horizon 2030 dont 1,7 GW en éolien



Figure 1. 3 : La ferme éolienne d'Adrar

1.3. DEFINITION DE L'ENERGIE EOLIENNE

Un aérogénérateur, plus communément appelé éolienne, est un dispositif qui transforme une partie de l'énergie cinétique du vent en énergie mécanique disponible sur un arbre de transmission puis en énergie électrique par l'intermédiaire d'une génératrice.[7] (figure 1.4)



Figure 1. 4 : Conversion de l'énergie cinétique du vent

Le SCE (Système de Conversion Eolien) est constitué d'un générateur électrique, entraîné 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.

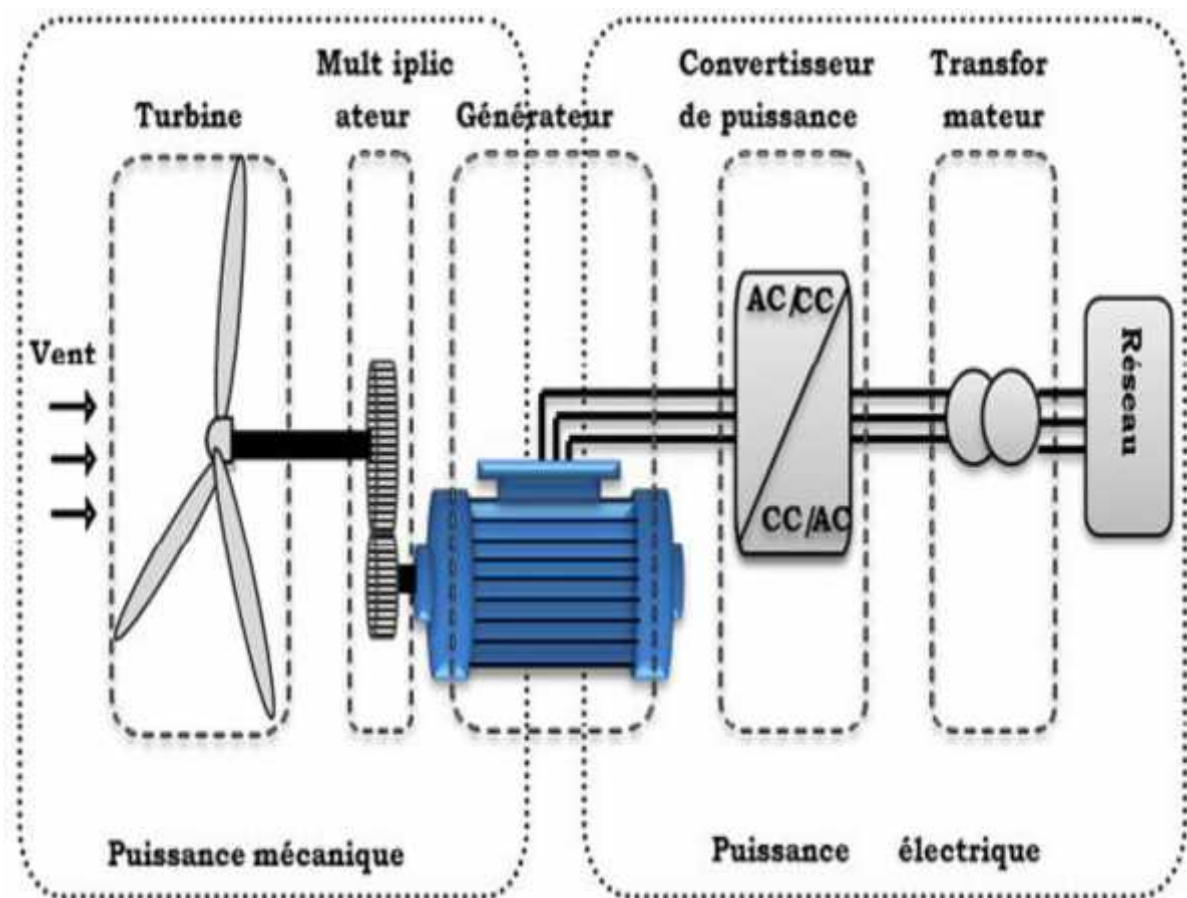


Figure 1. 5 : Principaux organes du système de conversion éolien [7]

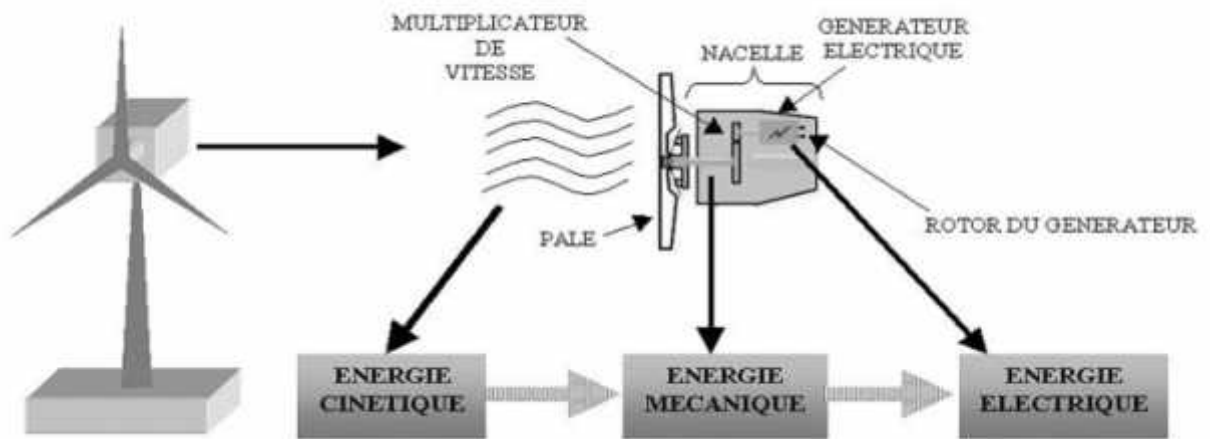


Figure 1. 6 : Conversion de l'énergie cinétique du vent [7]

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.[8]

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.

1.3.1. 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 [9].

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, treillis, autoporteur...etc. Pour qu'une éolienne produise...Quelques notions importantes doivent être pris avant l'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 10m. 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.

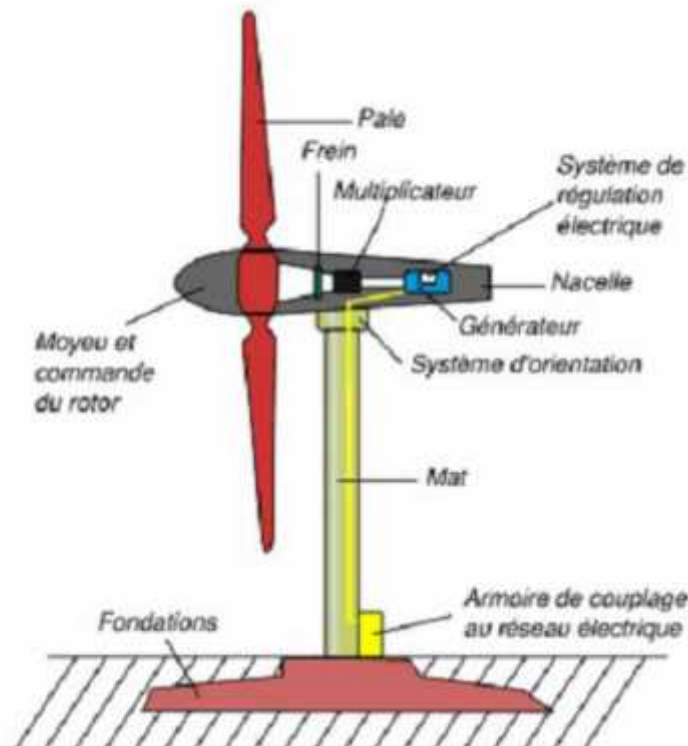


Figure 1.7 : Schéma d'une éolienne de type aérogénérateur [9]

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, les éoliennes doivent résister à des tempêtes aussi violentes. 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 [9].

1.3.2. Principaux Composants d'une Eolienne

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 80m de haut pour les plus puissantes (exceptionnellement jusqu'à 100 m). 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 [10].



Figure 1. 8 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, et d'autre part la puissance transmise par le rotor [10].

- **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 suffisante : 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 [10].



Figure 1. 9 Rotor d'une turbine.

Le rotor est relié à la nacelle par le moyeu, il transforme l'énergie cinétique du vent en énergie mécanique [10].

- **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 [10]. 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.

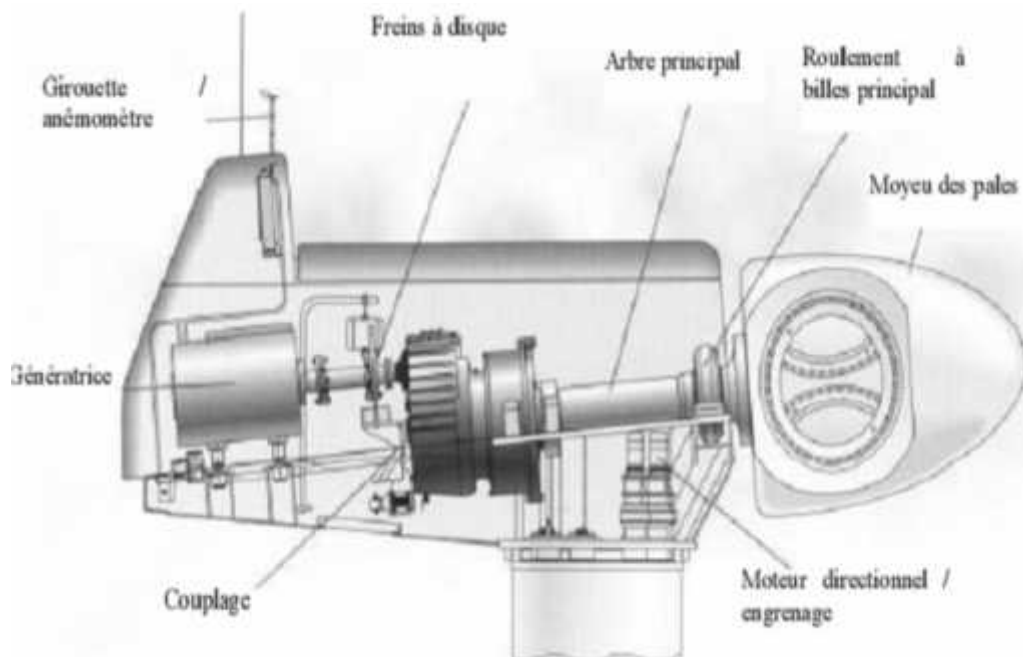


Figure 1. 10 : Constitution d'une nacelle [10]

Les différents composants d'une nacelle sont [9]:

-) 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 injecter l'intégralité de l'énergie produite [9].

1.3.3. 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 Algérie), 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.

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, 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 probabilité 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 d'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 ne 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. 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 et en Algérie), 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, 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.

1.3.4. Les différents types d'éoliennes

1.3.4.1. Les éoliennes à axe vertical

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.

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.

Les avantages théoriques d'une machine à axe vertical sont les suivantes :

-) Elle vous permet de placer la génératrice, le multiplicateur, etc. à terre, et vous n'avez pas besoin de munir la machine d'une tour.
-) Un mécanisme d'orientation n'est pas nécessaire pour orienter le rotor dans la direction du vent.

Les inconvénients principaux sont les suivants :

-) L'efficacité globale des éoliennes à axe vertical n'est pas impressionnante.
-) L'éolienne ne démarre pas automatiquement. Cependant, ceci ne constitue qu'un inconvénient mineur dans le cas d'une éolienne raccordée au réseau, étant donné qu'il est alors possible d'utiliser la génératrice comme un moteur absorbant du courant du réseau pour démarrer l'éolienne).



Figure 1. 11 *Eoliennes à axe vertical.*

1.3.4.2. 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 aéro-dynamiquement 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



Figure 1. 12 :*Eoliennes Horizontale.*

Les différents composants d'une éolienne à axe horizontal sont :

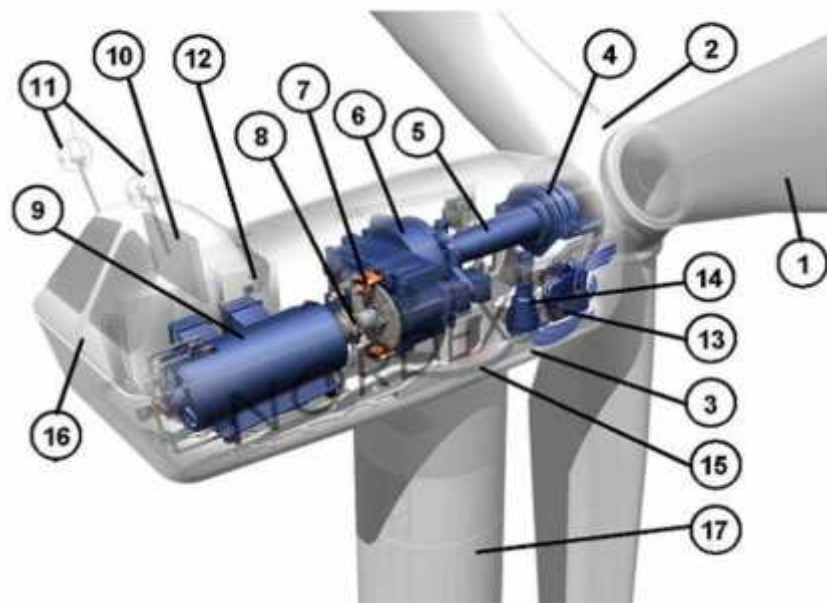


Figure 1. 13 : Principales composantes de l'éolienne à axe horizontal.

1. Pales en composite fibre de verre. Pas fixe, profil de type stall (décrochage aérodynamique).
Freins aérodynamiques d'extrémité pivotants.
2. Noyau du rotor en fonte.
3. Structure de la turbine en fonte ductile ou en acier soudé galvanisé à chaud.
4. Paliers du rotor à double rangée de billes.
5. Arbre du rotor en acier de haute résistance.
6. Multiplicateur de vitesse à 3 étages (1 train épicycloïdal et 2 trains parallèles).
7. Frein à disque sur l'arbre rapide avec témoin d'usure.
8. Accouplement avec la génératrice de type flexible.
9. Génératrice (800/200 kW) asynchrone refroidie à l'eau.
10. Radiateur de refroidissement intégré au système multiplicateur - génératrice.
11. Système de mesure du vent (anémomètre et girouette) transmettant les signaux au système de contrôle de la turbine.
12. Système de contrôle surveillant et pilotant la turbine.
13. Système hydraulique pour les freins d'extrémité de pale et le système d'orientation.
14. Entraînement d'orientation de la tourelle à deux trains d'engrenages planétaires entraînés par des moteurs alimentés à fréquence variable.
15. Paliers du système d'orientation équipés d'un frein à disque.
16. Capot de la nacelle en structure acier recouverte de composite fibre de verre.

17. Tour en acier tubulaire (plusieurs hauteurs possibles).

1.3.5. Classification des turbines éoliennes

Les solutions techniques permettant de recueillir l'énergie du vent sont très variées. Le tableau 1 présente une classification des turbines éoliennes :

ECHELLE	DIMÉTRE DE L'HÉLICE	PUISSANCE DELIVRÉE
Petit	moins de 12 m	moins de 40 KW
Moyen	12 m à 45 m	De 40 KW à 1 MW
Grand	46 m et plus	1 MW et plus

Tableau 1. 1 *Classification des turbines éoliennes [11]*

1.3.5.1 Les raisons pour choisir une grande éolienne :

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) sont plus ou moins les mêmes, quel que soit la taille de l'éolienne.

1. 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.
2. 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.

1.3.5.2 Les raisons pour choisir une éolienne plus petite :

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.

1. 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, le choix d'éoliennes plutôt petites peut se révéler avantageux dans un réseau électrique faible.
2. 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.
3. 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).

1.3.6. Taille des Aérogénérateurs

Les constructeurs et chercheurs ne finissent pas de mettre au point des éoliennes de plus en plus puissantes donc plus grandes, ce que représente la figure 1.14:

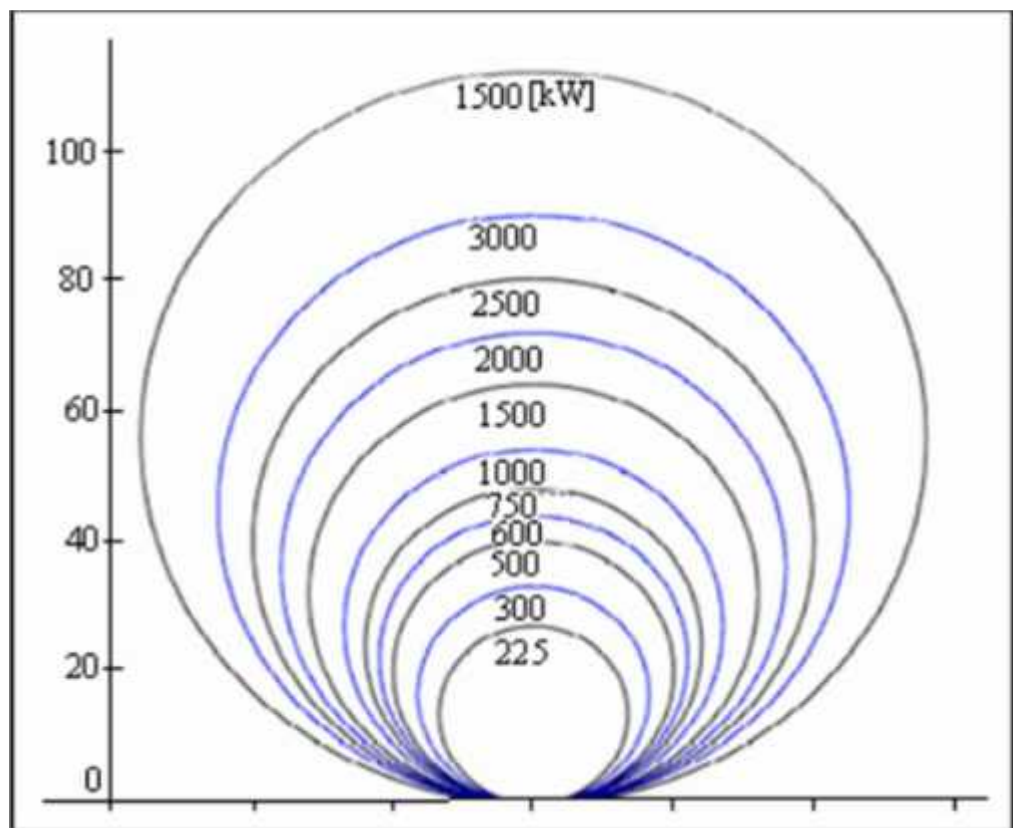


Figure 1. 14 Taille des hélices en [m] et puissance en [kW] [11]

1.3.7. Les avantages et les inconvénients d'énergie éolienne

1.3.7.1. Les avantages d'énergie éolienne

L'énergie éolienne est avant tout une énergie qui respecte l'environnement :

-) L'impact néfaste de certaines activités de l'homme sur la nature est aujourd'hui reconnu par de nombreux spécialistes. Certaines sources d'énergie, contribuent notamment à un changement global du climat, aux pluies acides ou à la pollution de notre planète en général. La concentration de CO₂ a augmenté de 25% depuis l'ère préindustrielle et on augure qu'elle doublera pour 2050 . Ceci a déjà provoqué une

augmentation de la température de 0,3 à 0,6° C depuis 1900 et les scientifiques prévoient que la température moyenne augmentera de 1 à 3,5°C d'ici l'an 2100, ce qui constituerait le taux de réchauffement le plus grand des 10000 dernières années . Toutes les conséquences de ce réchauffement ne sont pas prévisibles, mais on peut par exemple avancer qu'il provoquera une augmentation du niveau de la mer de 15 à 95 cm d'ici l'an 2100 . « L'exploitation d'énergie éolienne ne produit pas directement de CO₂ ».

- J L'énergie éolienne, propre, fiable, économique, et écologique, c'est une énergie qui respecte l'environnement.

Bien que, on pouvait pas 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.

- J L'énergie éolienne est une énergie renouvelable propre, gratuit, et inépuisable.
- J 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.
- J Parmi toutes les sources de production d'électricité, celle d'origine éolienne subit de très loin le plus fort taux de croissance.
- J 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.
- J 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 lorsqu'on n'en a pas besoin, entraînant ainsi d'importantes pertes et par conséquent un mauvais rendement énergétique.
- J Les parcs éoliens se démontent très facilement et ne laissent pas de trace.
- J 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.
- J 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, 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.
-)] L'énergie éolienne se révèle une excellente ressource d'appoint à d'autres énergies, notamment durant les pics de consommation, en hiver par exemple.

1.3.7.2. Les inconvénients d'énergie éolienne

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

-)] L'impact visuel : Ça 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 un site éolien.
-)] 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). 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.

1.4. ENERGIE CINETIQUE DU VENT – CONVERSION EN ENERGIE MECANIQUE

L'éolienne est un système complet permettant de convertir l'énergie (mécanique) du vent en énergie électrique. Trois facteurs déterminent le rapport entre l'énergie du vent et l'énergie mécanique récupérée par le rotor : la densité de l'air, la surface balayée par le rotor et la vitesse du vent. La densité de l'air et la vitesse du vent sont des paramètres climatologiques qui dépendent du site [12].

1.4.1. Loi de BETZ

Les pales d'une éolienne tire l'écoulement d'énergie à partir de l'air mobile, qui convertit cette dernière en énergie de rotation et la fournit par l'intermédiaire d'une unité mécanique d'entraînement au rotor d'un générateur électrique. L'énergie cinétique en air d'un objet de la masse m se déplaçant avec la vitesse v est égale à [13], [14] :

$$E = \frac{1}{2} m v^2 \quad (1.1)$$

La puissance de l'air en mouvement, si nous supposons la vitesse constante du vent, est la suivante :

$$P_m = \frac{d}{dt} = \frac{1}{2} \dot{m} \cdot v^2 \quad (1.2)$$

Où : \dot{m} - le taux de écoulement (débit) de la masse par seconde.

Lorsque l'air passe à travers une zone A, telle que la surface balayée par les pales du rotor, la puissance dans l'air peut être estimé par :

$$P_m = \frac{1}{2} \rho \cdot A \cdot v^2 \quad (1.3)$$

Où : ρ - la densité de l'air. Elle varie avec la pression de l'air et la température.

L'équation ci-dessus estime combien de puissance se trouve dans le vent, mais quelle quantité de cette puissance peut être obtenue à partir du flux d'air avec un convertisseur de puissance. L'énergie mécanique qui est obtenu du convertisseur à partir du flux d'air sera **égale à la différence des puissances du flux d'air avant et après le convertisseur** :

$$P_m = \frac{1}{2} \rho \cdot A_1 \cdot v_1^3 - \frac{1}{2} \rho \cdot A_2 \cdot v_2^3 = \frac{1}{2} \rho (A_1 v_1^3 - A_2 v_2^3) \quad (1.4)$$

Où : A_2 et A_1 - les sections transversales avant et après le convertisseur

v_1 et v_2 : la vitesse du vent avant et après le convertisseur.

Comme le flux d'air passe par le convertisseur, la vitesse du vent doit diminuer derrière le convertisseur de d'énergie éolienne et le flux massif reste invariable. Par conséquent.

$$\rho \cdot v_1 \cdot A_1 = \rho \cdot v_2 \cdot A_2 \quad (1.5)$$

Ainsi,

$$P_m = \frac{1}{2} \rho \cdot A_1 \cdot v_1 \cdot (v_1^2 - v_2^2) \quad (1.6)$$

On peut remarquer que la puissance mécanique maximale se produit lorsque v_2^2 est nulle. Si v_2^2 est égale à zéro, l'air a été porté à un arrêt complet par le convertisseur, qui n'est pas physiquement possible. Si la vitesse de sortie derrière le convertisseur est zéro, alors la vitesse d'apport doit être zéro, que implique qu'il n'y a pas de débit par le convertisseur. Par conséquent, la puissance mécanique débitée du flux d'air doit être exprimé d'une autre équation. En utilisant la loi de la conservation des impulsions, la force exercée par le vent sur le convertisseur est :

$$F = m \cdot (v_1 - v_2) \quad (1.7)$$

La transmission mécanique extraite est :

$$P_m = F \cdot v = m \cdot (v_1 - v_2) \cdot v \quad (1.8)$$

En comparant l'équation 1.4 et 1.8 on peut obtenir la relation pour la vitesse du flux

$$v = \frac{1}{2} (v_1 + v_2); \quad (1.9)$$

Ainsi la vitesse du flux d'air à travers le convertisseur est égale à la moyenne de v_1 et de v_2 .

La puissance mécanique du convertisseur peut alors être exprimé en tant que:

$$P_m = \frac{1}{2} \rho \cdot A \cdot (v_1^2 - v_2^2) \cdot (v_1 + v_2) \quad (1.10)$$

Si l'on compare cette puissance mécanique de sortie avec la puissance dans le flux d'air qui circule à travers la même surface transversale de la section A, le rapport entre la puissance mécanique débitée par le convertisseur et la puissance contenue dans le flux d'air qui traverse la même zone est appelée le "coefficient de puissance" C_p et peut être représentée comme suit:

$$C_p = \frac{P_m}{P_m} = \frac{\frac{1}{2} \rho \cdot A \cdot (v_1^2 - v_2^2) \cdot (v_1 + v_2)}{\frac{1}{2} \rho \cdot A \cdot v^3} \quad (1.11)$$

Le coefficient de puissance peut également être exprimé en termes de rapport de vitesse v_1/v_2 :

$$C_p = \frac{P_m}{P_m} = \frac{1}{2} \left(1 - \left(\frac{v_2}{v_1} \right)^2 \right) \left| 1 - \frac{v_2}{v_1} \right| \quad (1.12)$$

Si on trace l'équation (1,12) on remarque que la valeur maximale idéale du coefficient de puissance c_p , donc la puissance maximale qui peut être récupérée avec une turbine éolienne se produit quand

$$\frac{v_2}{v_1} = \frac{1}{3}, \text{ donc le } c_p \text{ devient : } c_p = \frac{1}{2} = 0.593$$

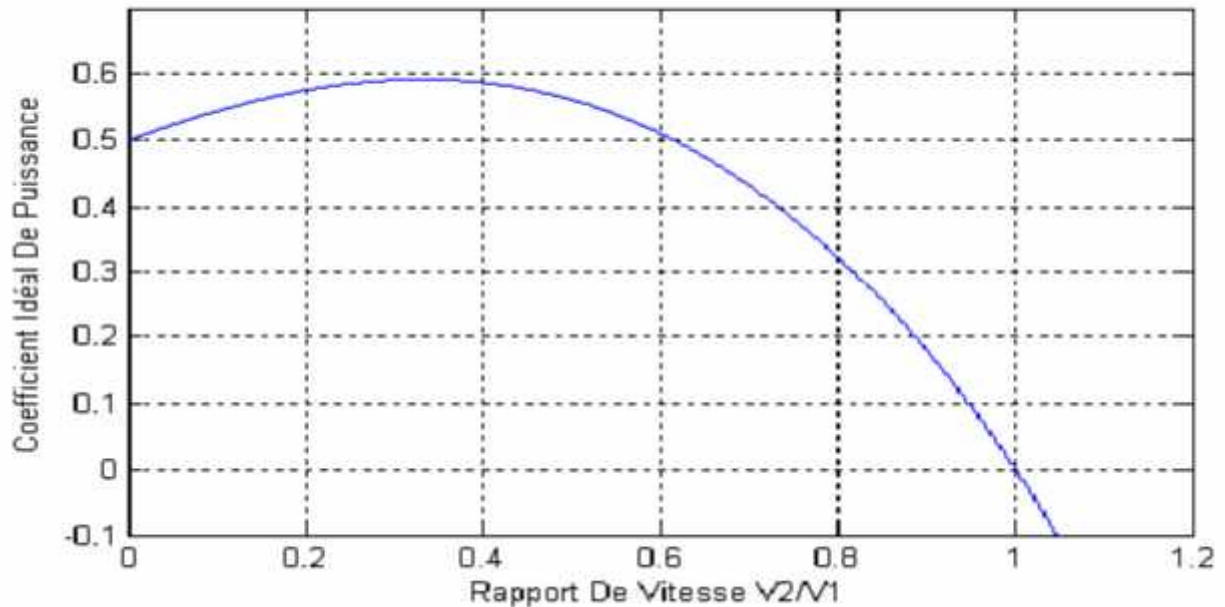


Figure 1. 15 : Coefficient de puissance en fonction du rapport de vitesse du vent [13]

Ceci a été dérivé la première fois par A. Betz, c'est pourquoi on l'appelle le " facteur de Betz " ou " limite de Betz ". Cette valeur est la valeur théorique maximale du coefficient de puissance, donc, l'efficacité théorique maximale qu'un convertisseur d'énergie éolienne peut avoir est 59,3%. Il est bien de mentionner que cette valeur du coefficient de puissance a été obtenue pour un idéal convertisseur de flux sans frottement.

Dans des cas réels, l'éolienne aura toujours un plus petit coefficient maximum de puissance que le facteur de Betz; à cause des pertes aérodynamiques qui dépendent de la conception du rotor et de la construction (nombre de pales, poids, rigidité, etc..). Le coefficient de puissance et l'efficacité d'un système d'éoliennes sont différents. L'efficacité d'une éolienne inclut la perte dans la transmission mécanique, la production d'électricité, les pertes du convertisseur, etc., où le coefficient de puissance est l'efficacité de convertir la puissance du vent en puissance mécanique dans l'arbre du rotor. La figure 1.16 montre un diagramme avec les pertes d'un système de la turbine éolienne.

Le coefficient de puissance est généralement donné comme fonction du rapport de vitesse en bout de pale d'hélice et l'angle d'attaque de pale . L'angle d'attaque est l'angle entre le plan de la rotation et l'accord du plan latitudinal de la pale. Le rapport de vitesse en

bout d'une éolienne est défini comme :

$$\lambda = \frac{u}{v_1} = \frac{\Omega_1 \cdot R}{v_1} \quad (1.13)$$

Où :

u : La vitesse tangentielle d'attaque de la pale ;

Ω_1 : La vitesse angulaire du rotor ;

R : Le rayon du rotor en mètres ;

v_1 : La vitesse du vent.

1.5. CONCLUSION

Une description de l'énergie éolienne a été présentée dans ce chapitre. Dans ce contexte, quelques notions principales de la technologie éolienne ont été données comme : les méthodes de description du gisement éolien ; les exemples d'architectures d'éoliennes (à axe horizontal, vertical), et les caractéristiques technologiques des éoliennes à axe horizontal ; le rappel des notions élémentaires nécessaires à la compréhension de la chaîne de conversion de l'énergie cinétique du vent en énergie électrique a été également donné dans ce chapitre.

Chapitre 02

Techniques de Réduction de la dimension et techniques de Régression

Sommaire

II.2 La réduction du dimension	26
II.2.1 Définition :	26
II.2.2 Méthodes de réduction de dimension	27
II.3 Régression	32
II.3.1 Histoire	32
II.3.2 Définition :	32
II.3.3 Principaux modèles de régression	33
II.3.4 Applications.....	34
II.3.5 Les reseaus de neurons	34
II.3.5.6.Développement des modèles de type Perceptron Multi Couches (PMC).....	39
II.3.5.7 Le réseau de neurone fonctions à base radial (RBF).....	48
II.3.5.8 Comparaison MLP et RBF.....	53
II.4 Conclusion:	54

II.1 Introduction :

Tout système de reconnaissance biométrique comporte une phase très importante basée sur la réduction d'espace. Pour cela nous consacrons un chapitre pour présenter les différentes techniques de projection d'espace.

Dans les différents domaines de recherches scientifiques, le développement technologique et le besoin de superviser des systèmes de plus en plus complexes nécessitent l'analyse de bases de données de taille importante (signaux, images, documents, ...). [17]

En apprentissage automatique, on distingue les problèmes de régression des problèmes de classification. Ainsi, on considère que les problèmes de prédiction d'une variable quantitative sont des problèmes de régression tandis que les problèmes de prédiction d'une variable qualitative sont des problèmes de classification. Certaines méthodes, comme la régression logistique, sont à la fois des méthodes de régression au sens où il s'agit de prédire la probabilité d'appartenir à chacune des classes et des méthodes de classification.

Les réseaux de neurones sont généralement optimisés par des méthodes d'apprentissage de type probabiliste, en particulier bayésien. Ils sont placés d'une part, dans la famille des applications statistiques, qu'ils enrichissent avec un ensemble de paradigmes permettant de créer des classifications rapides (réseaux de Kohonen en particulier), et d'autre part dans la famille des méthodes de l'intelligence artificielle auxquelles ils fournissent un mécanisme perceptif indépendant des idées propres de l'implémenter, et fournissant des informations d'entrée au raisonnement logique formel.

II.2 La réduction de la dimension

II.2.1 Définition :

Les méthodes de réduction de la dimension de l'espace de représentation peuvent être divisées en méthodes d'extraction d'attributs et méthodes de sélection d'attributs. L'extraction d'attributs transforme l'espace d'attributs de départ en un nouvel espace formé de la combinaison linéaire ou non linéaire des attributs initiaux. La sélection d'attributs choisit les attributs les plus pertinents selon un critère donné. Les données sont alors analysées après projection dans un espace de représentation composé des attributs les plus pertinents. Toutefois, l'interprétation des attributs extraits est plus délicate que l'interprétation des attributs sélectionnés. Le point clé de la sélection d'attributs est la définition d'un score mesurant la pertinence de chacun des attributs. Cette sélection s'appuie sur la connaissance explicite et implicite sur les données. Quand on ne dispose d'aucune information a priori sur le regroupement des données en classes, le contexte d'apprentissage est dit non supervisé.

La pertinence d'un attribut est alors mesurée en évaluant ses capacités à préserver la structure des données. Pour de nombreuses applications, on dispose des informations à priori sur la répartition des données en classes. Ainsi, pour ces données, les labels des classes ont été fournis. Dans ce cas, la sélection supervisée consiste à mesurer la corrélation entre l'attribut et les labels des classes des données. Nous n'aborderons que ce dernier cas de classification car tous système biométrique est fondé sur une phase d'apprentissage où la création de signature des données est effectuée d'où la création d'une base de données indexée.

II.2.2 Méthodes de réduction de dimension

Dans ce chapitre nous nous attachons à décrire plusieurs méthodes de réduction de dimension. La réduction de dimension consiste à transformer des données représentées dans un espace de grande dimension en une représentation dans un espace de dimension plus faible. Idéalement, la nouvelle représentation a une dimension égale au nombre de paramètres nécessaires pour décrire les données observées. La réduction de dimension est importante dans de nombreux domaines étant donné qu'elle facilite la classification, la visualisation ou encore la compression de données de grande dimension. Elle permet également souvent de limiter l'effet de la malédiction de la dimension et d'autres propriétés non désirées des espaces de grande dimension.

Nous décrivons ici quatre des méthodes linéaires les plus couramment utilisées :

II.2.2.1 L'Analyse en Composantes Principale

La description des liaisons entre deux variables par des techniques statistiques bidimensionnelles conduit à se poser la question de la représentation simultanée de données en dimension plus grande que 2. Quelle graphique permettrait de "généraliser" le nuage de points tracé dans le cas de deux variables permettant d'aborder la structure de corrélation présente entre plus de 2 variables. L'outil utilisé est alors l'analyse en composantes principales. Mathématiquement, l'analyse en composantes principales est un simple changement de base : passer d'une représentation dans la base canonique des variables initiales à une représentation dans la base des facteurs définis par les vecteurs propres de la matrice des corrélations [16].

L'ACP servira à mieux connaître les données sur lesquelles on travaille, à détecter éventuellement des valeurs suspectes, et aidera à formuler des hypothèses qu'il faudra étudier à l'aide de modèles et d'études statistiques différentielles. On pourra aussi, a posteriori, se servir des représentations fournies par l'ACP pour illustrer certains résultats dans un but pédagogique [15].

Algorithme :

Calcul de l'ACP

Entrées : matrice X

Sorties : vecteur moyen \bar{x} , base de vecteurs propres U , valeurs propres associées λ_i

Calcul du vecteur moyen :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Normalisation des images d'entrées :

$$\hat{x} = x_i - \bar{x}$$

$$\hat{X} = (\hat{x}_1, \dots, \hat{x}_n)$$

si Données de grande dimension alors

$$M = \frac{1}{n-1} \hat{X} \hat{X}^T$$

Calcul des éléments propres de M :

$$E = [e_1; \dots; e_n]$$

$$u_i = X e_i \quad U = [u_1; \dots; u_n]$$

$$\lambda_i = \delta_i \quad \lambda = [\lambda_1; \dots; \lambda_n]$$

Sinon

$$C = \frac{1}{n-1} \hat{X} \hat{X}^T$$

Calcul des éléments propres de C :

$$U = [u_1; \dots; u_n]$$

$$\lambda = [\lambda_1; \dots; \lambda_n]$$

retourner \bar{x}, U, λ

II.2.2.2 Factorisation de Matrice Non Négative

Les techniques de factorisation matricielle approchée sont d'un grand intérêt dans des domaines aussi variés que la reconnaissance de formes, le traitement d'antennes et aussi le traitement de données environnementales. La NMF, appelée Non negative Matrix Factorization, y est dédiée, elle a plutôt été introduite dans les domaines du traitement de signal et des images. Différentes formes de la NMF ont été développées selon le critère de minimisation choisie. [17]

II.2.2.3 Analyse en Composantes Indépendantes

Depuis plus d'une dizaine d'années, une petite communauté (venant à l'origine principalement du domaine du traitement du signal) développe des algorithmes pour l'analyse en composantes indépendantes (ACI) et la séparation de sources. Il s'agit de méthodes traitant des observations vectorielles (multi variées) afin d'en extraire des composantes linéaires qui soient « aussi indépendantes que possible ». Cette simple idée s'est révélée très fructueuse pour le traitement des

signaux multi-capteurs dans de nombreux domaines : réseaux d'antennes pour les télécommunications, prise de son, signaux biomédicaux multi-électrodes, et de manière plus générale dans tous les cas où un système de plusieurs capteurs, fournissant des signaux cohérents, est à l'écoute d'un ensemble discret de sources de signal que l'on cherche à extraire des observations et que l'on peut, pour des raisons physiques, supposer mutuellement statistiquement indépendantes[20].

L'Analyse en Composantes Indépendantes (ou ICA pour Indépendant Component Analysis) a été introduite par Héroult. Le but est d'exprimer un ensemble de n variables aléatoires $x_1; \dots; x_n$ comme une combinaison linéaire de n variables aléatoires statistiquement indépendantes S_j :

$$S_n \quad a_j = a_{j;1}S_1 + \dots + a_{j;n} \quad \forall j \quad (2.1)$$

ou sous forme matricielle : $x = As$ (2.2)

Le but de l'Analyse en Composantes Indépendantes est l'estimation des composantes originales s_i , ou de manière équivalente des coefficients a_{ij} . Par définition, les variables aléatoires s_i sont mutuellement indépendantes et la matrice de mélange est donc inversible. Ainsi le problème de l'ICA peut être formulé :

$$u = Wx = WAs \quad (2.3)$$

II.2.2.4 Le modèle discriminant linéaire amélioré de Fisher (EFM)

Ce modèle discriminant linéaire de Fisher améliore la capacité de généralisation de la FLD en décomposant la procédure FLD en diagonalisation simultanée des deux matrices de dispersion intra-classe et inter-classe. La diagonalisation simultanée est une étape sagement équivalente à deux opérations comme l'a souligné Fukunaga. Blanchiment de la matrice de dispersion intra-classe et l'application de l'ACP sur la matrice de dispersion intra-classe en utilisant les données transformées. Durant l'opération du blanchiment de la matrice de dispersion intra-classe apparaît dans le dénominateur de la séparabilité des petites valeurs propres qui tendent à capturer du bruit. Pour atteindre des performances améliorées l'EFM préserve un équilibre approprié entre la sélection des valeurs propres (correspondant à la composante principale de l'espace de l'image originale) qui tiennent compte de la plupart de l'énergie spectrale des données brutes, c'est à dire, représentation adéquate et l'exigence que les valeurs propres de la matrice de dispersion intra-classe (de l'espace ACP réduit) ne sont pas trop petites, c'est-à-dire une meilleure généralisation. Le choix de rang des composantes principales (m) pour la réduction de la dimension, prend en compte de l'ordre de grandeur de l'énergie spectrale. Les valeurs propres de la matrice de covariance fournissent un bon indicateur pour répondre au critère de l'énergie. Il faut ensuite calculer les valeurs propres de la matrices de dispersion intra-classe dans l'espace ACP réduit pour faciliter le choix du rang des composantes principales de sorte que l'exigence de grandeur est respectée [15].

II.2.2.5 Analyse Discriminante Linéaire(LDA) :

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 within-class scatter)[19].

II.2.2.5.1. présentation:

L'Analyse Discriminante Linéaire(LDA) est utilisée pour trouver la combinaison linéaire des caractéristiques qui séparent le mieux les classes d'objet ou d'événement. Les combinaisons résultantes peuvent être employées comme classificateur linéaire, ou généralement dans la réduction de caractéristiques avant la classification postérieure.

LDA est étroitement lié à l'ACP, du fait que tous les deux recherchent les combinaisons linéaires des variables qui représentent au mieux les données. Elle essaie explicitement de modéliser la différence entre les classes de données contrairement à l'ACP qui ne tient pas compte des différences entre les classes.

Chacune des nouvelles dimensions est une combinaison linéaire des valeurs de pixel, qui forment un *Template*. Les combinaisons obtenues en utilisant LDA s'appellent les fisher faces.

LDA est une technique qui cherche les directions qui sont efficaces pour la discrimination entre les données.

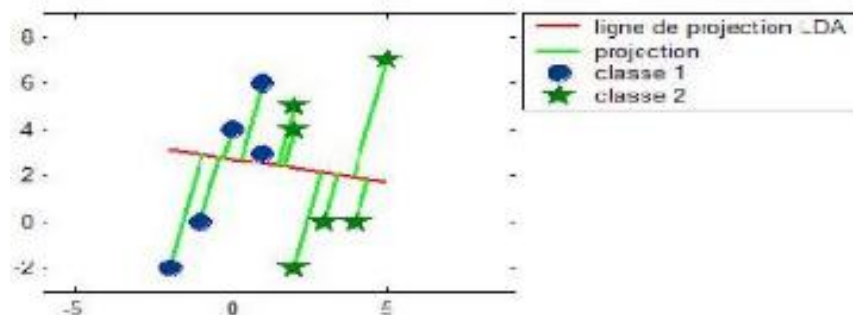


Figure 1.1 :Projection par LDA avec séparation entre les classes[21]

II.2.2.5.2 La règle bayésienne:

L'objectif est de produire une règle d'affectation $X(\omega) \rightarrow Y(\omega)$ qui permet de prédire, pour une observation ω donnée, sa valeur associée de Y à partir des valeurs prises par X .

La règle bayésienne consiste à produire une estimation de la probabilité *a posteriori* d'affectation

$$P(Y=y_k | X) = \frac{P(Y=y_k) \times P(X=Y=y_k)}{\sum_{i=1}^k P(Y=y_i) \times P(X=Y=y_i)} \quad (2.4)$$

$P(Y = y_k)$ est la probabilité *a priori* d'appartenance à une classe. $P(X = Y | y_k)$ représente la fonction de densité des X conditionnellement à la classe y_k

La règle d'affectation pour un individu ω à classer devient alors $Y(\omega) = \arg \max P(X = y_k | X(\omega))$.

Toute la problématique de l'analyse discriminante revient alors à proposer une estimation de la quantité $P(X = Y | y_k)$

II.2.2.5.3 Taux d'erreur:

De manière classique en apprentissage supervisé, pour évaluer les performances d'une fonction de classement, nous confrontons ses prédictions avec les vraies valeurs de la variable à prédire sur un fichier de données. Le tableau croisé qui en résulte s'appelle une matrice de confusion avec : en ligne les vraies classes d'appartenance, en colonne les classes d'appartenance prédites. Le taux d'erreur ou taux de mauvais classement est tout simplement le nombre de mauvais classement, lorsque la prédiction ne coïncide pas avec la vraie valeur, rapporté à l'effectif du fichier de données.

Le taux d'erreur a de séduisant qu'il est d'interprétation aisée, il s'agit d'un estimateur de la probabilité de se tromper si l'on applique la fonction de classement dans la population.

Attention cependant, on parle de taux biaisé ou taux d'erreur en résubstitution, le taux d'erreur mesuré sur les données qui ont servi à construire la fonction de classement. Tout simplement parce que les données sont juges et parties dans ce schéma. La bonne procédure serait de construire la fonction de classement sur une fraction des données, dites d'apprentissage ; puis de l'évaluer sur une autre fraction de données, dite de test. Le taux d'erreur en test ainsi mesuré est un indicateur digne de foi.

La pratique veut que la répartition des données en apprentissage et test soit de $2/3 - 1/3$. Mais en réalité, il n'y a pas de règle véritable. Le plus important est de concilier deux exigences contradictoires : en avoir suffisamment en test pour obtenir une estimation stable de l'erreur, tout en réservant suffisamment en apprentissage pour ne pas pénaliser la méthode d'apprentissage.

Le taux d'erreur permet d'évaluer et de comparer des méthodes, quelles que soient leurs hypothèses sous-jacentes. Dans le cas de l'analyse discriminante linéaire, nous pouvons exploiter le modèle probabiliste pour réaliser des tests d'hypothèses.

II.2.2.5.4 Les avantages et inconvénients :

Parmi les avantages engendrés par une utilisation de la méthode LDA, nous citons les points suivants :

- LDA maximise l'éparpillement inter-classes.
- Réduit l'éparpillement intra-classes.
- La méthode de Fisher faces résout le problème de la robustesse face aux variations de pose, et d'expressions faciales.

Malgré ces avantages, dans la littérature un ensemble de points négatives existe encore comme :

- Couteuse en temps de calcul.
- Couteuse en espace mémoire.
- Rend de mauvais résultats quand le nombre d'images d'apprentissage est grand.[21]

II.3 Régression

II.3.1 Histoire

Ruder Josip Bošković est le premier scientifique à calculer les coefficients de régression linéaire, en 1755-1757, quand il entreprit de mesurer la longueur de cinq méridiens terrestres en minimisant la somme des valeurs absolues. Pierre-Simon de Laplace utilise cette méthode pour mesurer les méridiens dans « Sur les degrés mesurés des méridiens et sur les longueurs observées sur pendule » en 1789. La première utilisation de la méthode des moindres carrés est attribuée à Adrien-Marie Legendre en 1805 ou à Carl Friedrich Gauss qui dit l'avoir utilisée à partir de 1795.

Carl Friedrich Gauss démontre, en 1821, le théorème connu aujourd'hui sous le nom de théorème de Gauss-Markov qui exprime sous certaines conditions la qualité des estimateurs, Andrei Markov le redécouvre en 1900.

C'est à Francis Galton qu'est accordée la paternité de l'expression « régression linéaire » en 1886³. Dans son article, Galton exprime la taille des fils en fonction de la taille des pères. Il constate un phénomène de « régression vers la moyenne ».

Plus tard la colinéarité des variables explicatives est devenue un sujet de recherche important. En 1970, Arthur E. Hoerl et Robert W. Kennard proposent la régression pseudo-orthogonale (*Ridge Regression*), une des méthodes d'estimation conçues pour pallier la présence de colinéarité de certaines variables explicatives en imposant des contraintes sur les coefficients.

La méthode du lasso (*Lasso Regression*), ayant le même objectif en utilisant une technique analogue, a été créée en 1996 par Robert Tibshirani.

Avec les méthodes de régression sur composantes (régression des moindres carrés partiels (PLS) et régression sur composantes principales), les algorithmes recherchent des variables

explicatives indépendantes liées aux variables initiales, puis estiment les coefficients de régression sur les nouvelles variables.

II.3.2 Définition :

La régression est un ensemble de méthodes statistiques très utilisées pour analyser la relation d'une variable par rapport à une ou plusieurs autres.

Parmi les modèles de régression linéaire, le plus simple est l'ajustement affine. Celui-ci consiste à rechercher la droite permettant d'expliquer le comportement d'une variable y comme étant une fonction affine d'une autre variable statistique x .

En général, le modèle de régression linéaire désigne un modèle dans lequel l'espérance conditionnelle de y sachant x est une transformation affine en les paramètres. Cependant, on peut aussi considérer des modèles dans lesquels c'est la médiane conditionnelle de y sachant x ou n'importe quel quantile de la distribution de y sachant x qui est une transformation affine en les paramètres.

Le modèle de régression linéaire est souvent estimé par la méthode des moindres carrés mais il existe aussi de nombreuses autres méthodes pour estimer ce modèle. On peut par exemple estimer le modèle par maximum de vraisemblance ou encore par inférence bayésienne.

Bien qu'ils soient souvent présentés ensemble, le modèle linéaire et la méthode des moindres carrés ne désignent pas la même chose. Le modèle linéaire désigne une classe de modèles qui peuvent être estimés par un grand nombre de méthodes, et la méthode des moindres carrés désigne une méthode d'estimation. Elle peut être utilisée pour estimer différents types de modèles.

II.3.3 Principaux modèles de régression

- Le modèle de régression le plus connu est le modèle de régression linéaire.
- Lorsque le modèle n'est pas linéaire, on peut effectuer une régression approchée par des algorithmes itératifs, on parle de régression non linéaire.
- Si on s'intéresse au quantile conditionnel de la distribution de la variable aléatoire y sachant le vecteur de variables aléatoires x , on utilise un modèle de régression quantile.
- Si la variable expliquée est une variable aléatoire binomiale, il est courant d'utiliser une régression logistique ou un modèle probit.
- Si la forme fonctionnelle de la régression est inconnue, on peut utiliser un modèle de régression non paramétrique.

II.3.4 Applications

Comme les autres modèles de régression, le modèle de régression linéaire est aussi bien utilisé pour chercher à prédire un phénomène que pour chercher à l'expliquer.

Après avoir estimé un modèle de régression linéaire, on peut prédire quel serait le niveau de y pour des valeurs particulières de x .

Il permet également d'estimer l'effet d'une ou plusieurs variables sur une autre en contrôlant par un ensemble de facteurs. Par exemple, dans le domaine des sciences de l'éducation, on peut évaluer l'effet de la taille des classes sur les performances scolaires des enfants en contrôlant par la catégorie socio-professionnelle des parents ou par l'emplacement géographique de l'établissement. Sous certaines hypothèses restrictives, cet effet peut être considéré comme un effet causal.

En apprentissage statistique, la méthode de régression linéaire est considérée comme une méthode d'apprentissage supervisé utilisée pour prédire une variable quantitative.

Dans cette perspective, on entraîne généralement le modèle sur un échantillon d'apprentissage et on teste ensuite les performances prédictives du modèle sur un échantillon de test.

II.3.5 Les réseaux de neurones

Un réseau de neurones artificiels est un modèle de calcul dont la conception est très schématiquement inspirée du fonctionnement des neurones biologiques.

Les réseaux de neurones sont généralement optimisés par des méthodes d'apprentissage de type probabiliste, en particulier bayésien. Ils sont placés d'une part, dans la famille des applications statistiques, qu'ils enrichissent avec un ensemble de paradigmes permettant de créer des classifications rapides (réseaux de Kohonen en particulier), et d'autre part dans la famille des méthodes de l'intelligence artificielle auxquelles ils fournissent un mécanisme perceptif indépendant des idées propres de l'implémenter, et fournissant des informations d'entrée au raisonnement logique formel

En modélisation des circuits biologiques, ils permettent de tester quelques hypothèses fonctionnelles issues de la neurophysiologie, ou encore les conséquences de ces hypothèses pour les comparer au réel

II.3.5.1. Historique

Les réseaux de neurones sont une métaphore biologique issue des recherches en cybernétique [22]. On doit les premiers travaux sur les réseaux de neurones ainsi que la première modélisation du neurone formel à W.S. McCulloch et W. Pitts [23]. Ceux-ci démontrent dès qu'en théorie, un réseau de neurones peut résoudre des problèmes logiques, arithmétiques et symboliques complexes. Toutefois, le premier vrai modèle de réseau de neurones n'apparaît qu'en 8 : Le perceptron de Rosenblatt [24] est le premier système artificiel capable d'apprendre par expérience. La communauté scientifique commence alors à s'intéresser très sérieusement aux réseaux de neurones, en se basant sur les travaux et les études précédentes de D Hebb sur l'apprentissage et les mémoires associatives [25]. Cependant, les limites du perceptron à traiter des problèmes complexes, et notamment l'impossibilité à résoudre des problèmes non linéaires font qu'il faudra attendre 8 pour

que HopField et son réseau multicouche et rebouclé [26] donne aux réseaux de neurones la légitimité qu'on lui connaît aujourd'hui

II.3.5.2. Utilité

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. Autre exemple, une banque peut créer un jeu de données sur les clients qui ont effectué un emprunt constitué : de leur revenu, de leur âge, du nombre d'enfants à charge... et s'il s'agit d'un bon client Si ce jeu de données est suffisamment grand, il peut être utilisé pour l'entraînement d'un réseau de neurones La banque pourra alors présenter les caractéristiques d'un potentiel nouveau client, et le réseau répondra s'il sera bon client ou non, en généralisant à partir des cas qu'il connaît Si le réseau de neurones fonctionne avec des nombres réels, la réponse traduit une probabilité de certitude. Par exemple : 1 pour « sûr qu'il sera un bon client », -1 pour « sûr qu'il sera mauvais client », 0 pour « aucune idée », 0,9 pour « presque sûr qu'il sera bon client ».

Le réseau de neurones ne fournit pas toujours de règle exploitable par un humain. Le réseau reste souvent une boîte noire qui fournit une réponse quand on lui présente une donnée, mais le réseau ne fournit pas de justification facile à interpréter.

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

II.3.5.3. Le neurone formel

Le neurone formel est une abstraction de la réalité biologique. Il synthétise mathématiquement toutes les informations apportées par les observations d'un vrai neurone à l'exception du critère de temporalité qu'il n'est pas utile de modéliser puisque les valeurs manipulées par les neurones représenteront les fréquences d'émission des stimuli.

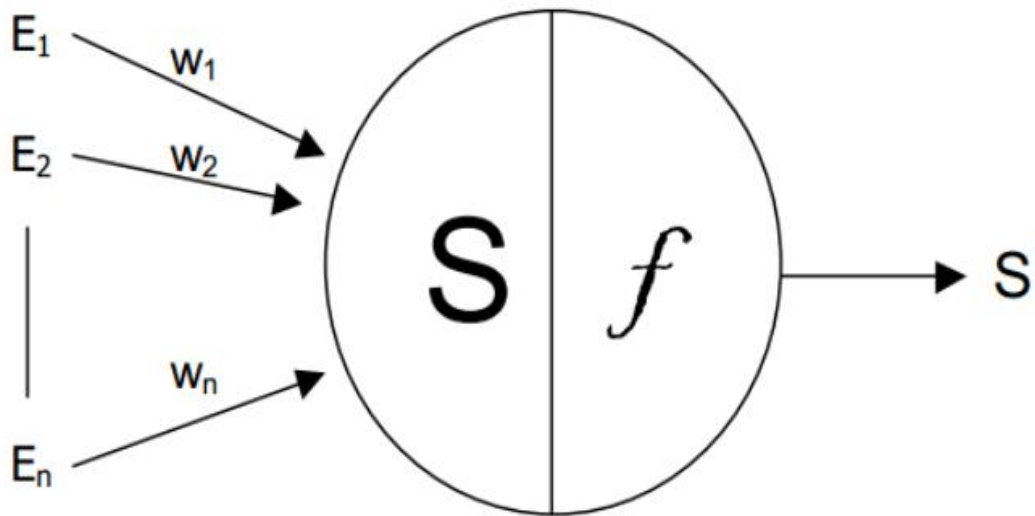


Figure 2.2 : *Le neurone formel .*

Le neurone formel, comme son homologue biologique, a donc pour unique fonction de transmettre un influx nerveux sous certaines conditions. Pour cela, il possède une ou plusieurs entrées (les dendrites) et une sortie (l'axone) Le corps quant à lui ne sert qu'à évaluer si le signal reçu sur les entrées doit être propagé sur la sortie ou non. Sur la figure. 2.3, les représentent les n entrées du neurone, dont les liens sont pondérés par les poids synaptiques Ces poids traduisent de quelle manière l'entréeconcernée pèsera par rapport aux autres. Si le poids est positif, on parle de synapse excitatrice alors que s'il est négatif, on parle de synapse inhibitrice S représente donc la somme pondérée des entrées du neurone.

La sortie S est donnée par l'application de la fonction d'activation f à cette somme, autrement dit : $S = f(S)$.

Les modèles linéaires et sigmoïdaux sont bien adaptés aux algorithmes d'apprentissage impliquant une rétro-propagation du gradient car leur fonction d'activation est différentiable ; ce sont les plus utilisés Il existe en pratique plusieurs types de fonctions d'activation classées suivant les valeurs manipulées par le neurone en la page suivant :

Valeurs binaires	Fonction Heaviside	$f(\Sigma) = H(\Sigma) \begin{cases} 1 & \text{si } \Sigma \geq 0 \\ 0 & \text{si } \Sigma < 0 \end{cases}$	
	Fonction signe	$f(\Sigma) = s(\Sigma) = \begin{cases} 1 & \text{si } \Sigma \geq 0 \\ -1 & \text{si } \Sigma < 0 \end{cases}$	
	Fonction linéaire à seuil	$f(\Sigma) = LS(\Sigma) = \begin{cases} 1 & \text{si } \Sigma < a \\ \frac{1}{a} \Sigma & \text{si } \Sigma \in [-a, a] \\ -1 & \text{si } \Sigma < -a \end{cases}$	
Valeur réelles	Fonction sigmoïde	$f(\Sigma) = \text{sigm}(\Sigma) = \frac{\Sigma}{1 + \Sigma }$	

Figure 2.3 : Principales fonctions d'activation .

II.3.5.4. Caractéristiques d'un réseau de neurones

Malgré la simplicité de fonctionnement d'un neurone, c'est l'interconnexion d'un certain nombre d'entre eux qui rend l'ensemble très complexe. De manière générale, on envisage un réseau comme ayant plusieurs couches de neurones, dont une couche d'entrée, une couche de sortie et un certain nombre de couches intermédiaires appelé en anglais "hidden layers" (couches cachées). Si les liens vont des neurones d'entrée vers les neurones de sortie, on définit le réseau comme "feedforward". Si les liens peuvent boucler à l'intérieur du réseau, on l'appelle «feedback». Le fait de pouvoir avoir des rebouclages permet au réseau d'avoir une sorte de mémoire interne. Si chaque neurone est relié à tous les neurones de la couche suivante, on dit du réseau qu'il est « complet ».

II.3.5.5. Principe de fonctionnement

Le fonctionnement d'un réseau est un peu plus simple, et dépend uniquement du fonctionnement de ses neurones. Celui-ci doit être considéré comme un circuit électrique contrôlé par une horloge. A chaque cycle, tous les neurones vont, de manière parallèle, calculer une valeur de sortie en fonction de la somme de leurs valeurs d'entrées, sachant que, pour un neurone A , les valeurs en entrée au cycle n sont les valeurs en sortie au cycle $n-1$ des neurones connectés à A . On va donc assister, à chaque "clockage" du réseau, à une propagation de l'influx nerveux de ses entrées vers ses sorties.

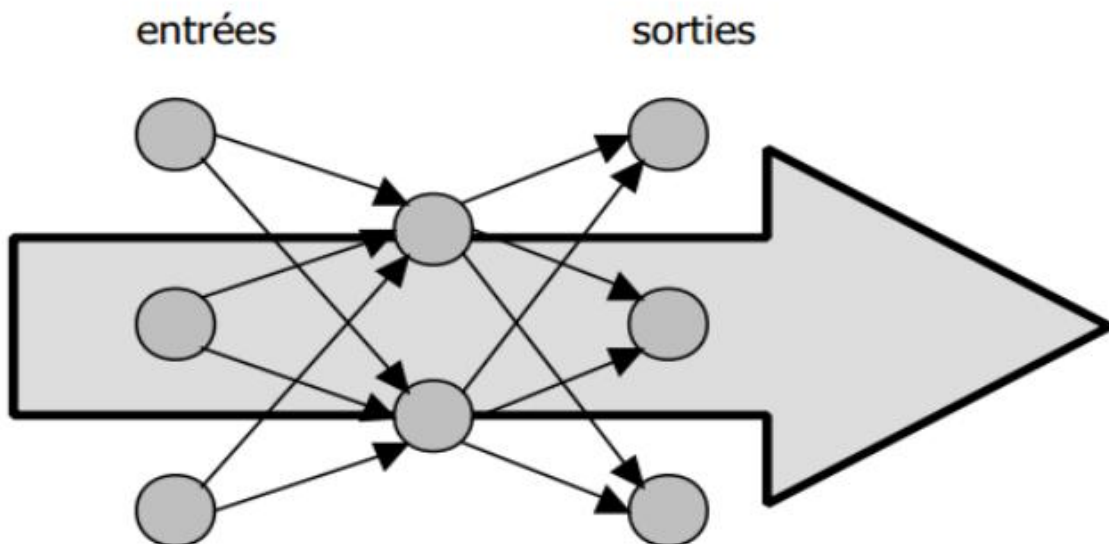


Figure 2.4 : Propagation de l'influx nerveux dans un réseau de type perceptron.

II.3.5.6. Développement des modèles de type Perceptron Multi Couches (PMC)

Il existe un grand nombre de types de réseaux de neurones, dont chacun a des avantages et des inconvénients. le réseau choisit dans notre cas est un réseau multicouches .ce choix est fait pour la facilité et la rapidité de sa construction et encore par le fait que notre problème présente un nombre limité de variables d'entrées .

Un réseau de neurones est un ensemble de processeurs élémentaires, les neurones qui sont largement connectés les uns aux autres et qui sont capables d'échanger des informations au moyen des connexions qui les relient. Les connexions sont directionnelles et à chacune d'elle est associé un réel appelé poids de la connexion. L'information est ainsi transmise de manière unidirectionnelle du neurone j vers le neurone i, affectée du coefficient pondérateur w_{ij} . Un neurone calcule son état à partir d'informations venues de l'extérieur, ou bien il détermine son entrée à partir des neurones auxquels il est connecté et calcule son état comme une transformation souvent non linéaire de son entrée. Il transmet à son tour son état vers d'autres neurones ou vers l'environnement extérieur.

Un neurone est donc défini par trois caractéristiques: son état, ses connexions avec d'autres neurones et sa fonction de transfert. Nous utiliserons dans la suite, les notations suivantes:

- O : l'ensemble des états possibles des neurones.
- o_i : l'état du neurone i, où $o_i \in O$.
- f_i : la fonction de transfert associée au neurone i.
- s_i : l'entrée du neurone i, si $s_i \in \mathbb{R}$
- w_{ij} : le poids de la connexion du neurone j vers le neurone i; $w_{ij} \in W$ l'ensemble des poids du PMC. Ainsi le neurone i recevant les informations de n_i neurones effectue l'opération suivante:

$$o_i = f(s_i) \text{ avec } s_i = \sum_{j=1}^{n_i} w_{ij} o_j \quad (2.5)$$

Les fonctions de transfert les plus souvent utilisées sont la fonction identité, la fonction sigmoïde, la fonction exponentielle et plus rarement les fonctions ondelettes [62].

la fonction identité: un neurone dont la fonction de transfert est la fonction identité appelé neurone linéaire. Pour un tel neurone, l'état est calculé à l'aide de l'équation suivante:

$$o_i = s_i = \sum_j w_{ij} o_j \quad (2)$$

la fonction sigmoïde: est la plus utilisée car elle introduit de la non-linéarité, mais c'est aussi une fonction continue, différentiable. Une fonction sigmoïde peut être définie par l'une des deux formes suivantes :

$$f(s) = A \frac{e^{Ks}-1}{e^{Ks}+1} = A \tanh\left(\frac{k}{2} s\right) \quad (2.6)$$

Un perceptron multi-couches (PMC, ou *multi-layer perceptron*, *MLP*), ou réseau de neurones multi-couches, définit une famille de fonctions. Considérons d'abord le cas le plus classique, celui d'un réseau de neurones à une couche cachée, qui associe à un vecteur de dimension d un vecteur de longueur m (par exemple, pour effectuer une régression) :

$$g(x) = b + W = \tan h(c + Vx) \quad (2.7)$$

où :

- x est un vecteur de longueur d (l'entrée) ;
- V est une matrice $k \times d$ (les poids de la couche d'entrée vers la couche cachée) ;
- c est un vecteur de longueur k (les biais de la couche cachée) ;
- b est un vecteur de longueur m (les biais de la couche de sortie) ;
- W est une matrice $m \times k$ (les poids de la couche cachée vers la couche de sortie).

La fonction $h(x) = \tan h(x + Vx)$, qui retourne un vecteur, est appelée la sortie de la couche cachée. Notons que dans le réseau ci-dessus, la sortie est une transformation affine de la couche cachée. Une non-linéarité peut être appliquée par dessus dans certaines architectures. Les éléments de la couche cachée sont appelés *unités cachées*.

Il est possible d'appliquer à nouveau le même genre de transformation à $h(x)$ elle-même, avec des paramètres (biais et poids) différents. Cela donnerait un réseau de neurones multi-couches (à propagation avant) avec deux couches cachées. De manière plus générale, on peut construire un réseau de neurones profond en empilant plusieurs de ces couches. Chacune des couches peut avoir une taille différente (le k utilisé ci-dessus). Une variante commune est l'ajout de *connexions directes*, c'est-à-dire qu'une couche peut prendre comme entrée à la fois la couche précédente et d'autres couches plus basses (plus proches de l'entrée).

II.3.5.6.1 Architecture

Le perceptron multicouche (PMC) est un réseau composé de couches successives. Une couche est un ensemble de neurones n'ayant pas de connexion entre eux. Une couche d'entrée lit les signaux entrant, un neurone par entrée x_j , une couche en sortie fournit la réponse du système. Selon les auteurs, la couche d'entrée qui n'introduit aucune modification n'est pas comptabilisée. Une ou plusieurs couches cachées participent au transfert. Un neurone d'une couche cachée est connecté en entrée à chacun des neurones de la couche précédente et en sortie à chaque neurone de la couche suivante :

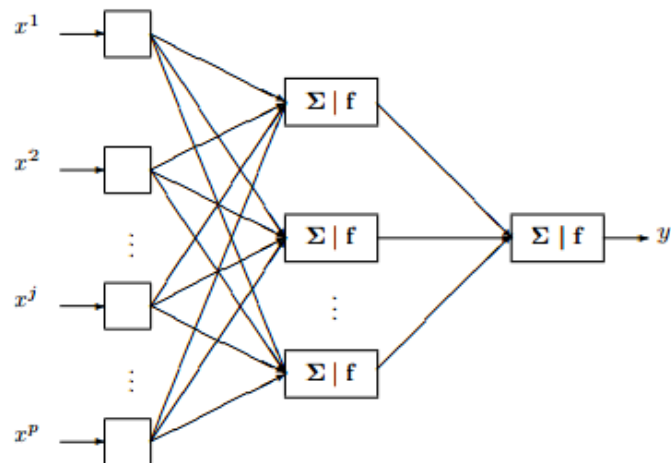


Figure 2.5 : Exemple de perceptron multicouche élémentaire avec une couche cachée et une couche de sortie

Descendant direct du perceptron, il est développé par Minsky et Papert. Reprenant les caractéristiques principales de son prédécesseur: feedforward et complet, il adopte un « hidden layer » non vide lui permettant de palier l'inaptitude de son prédécesseur à traiter de problèmes non linéairement séparables comme la porte logique complexe XOR.

	0	1
0	0	1
1	1	0

Figure 2.6 : Table de vérité de la porte logique XOR.

Il n'est plus exclusivement binaire et peut désormais utiliser des fonctions d'activations telles que la sigmoïde voire des gaussiennes dans le cas du RBF (Radial Basic Functions) qui en est une spécialisation L'apprentissage quant à lui est effectué par rétropropagation.

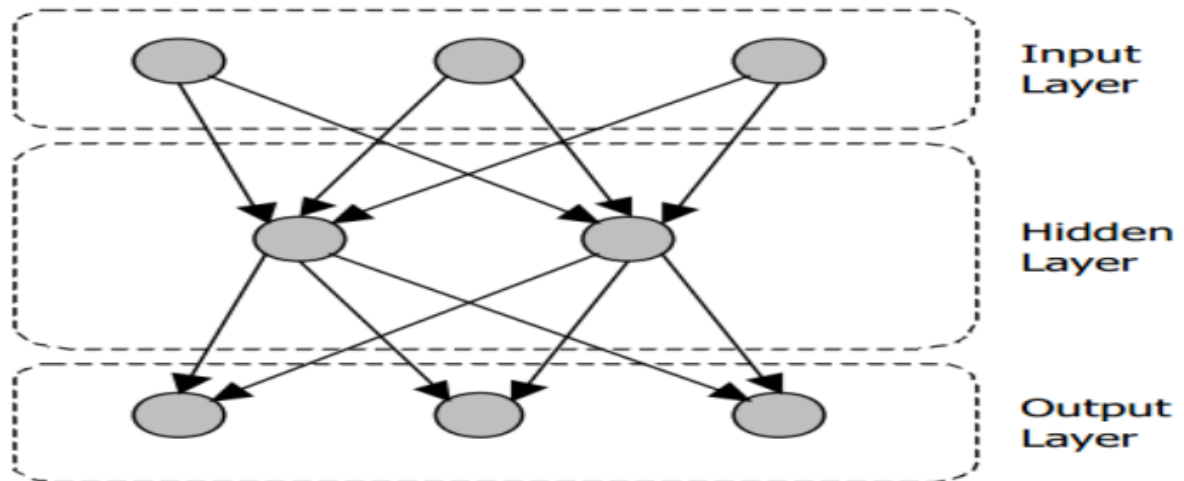


Figure 2.7 : Le perceptron multicouche.

II.3.5.6.2.L'algorithme de rétro propagation

Voici quelques conventions de notation pour les PMC avec plus d'une couche cachée. On note h_i le vecteur de sortie de la couche i , en commençant par $h_0(x) = x$ (l'entrée), jusqu'à la couche de sortie, h_k , qui produit la prédiction, ou sortie, du réseau.

Avec des unités tanh (dont la fonction d'activation est la tangente hyperbolique) dans les couches cachées, nous avons (en utilisant la notation matricielle) :

- Pour $k = 1$ à $K-1$:
- $h_k = \tanh(b_k + W_k h_{k-1})$
- où b_k est un vecteur de biais, et W_k est une matrice de poids, connectant la couche $k - 1$ à la couche k . Le calcul scalaire associé à une unité donnée i de la couche k est

$$h_{k,i} = \tanh(b_{k,i} + \sum_j W_{k,i,j} h_{k-1,j}) \quad (2.8)$$

Dans le cas d'un classifieur probabiliste, nous avons un softmax sur la couche de sortie, c'est-à-dire :

$$p = h_k = \text{softmax}(b_k + W_k h_{k-1}) \quad (2.9)$$

où nous notons la sortie p parce que c'est un vecteur indiquant une distribution de probabilité sur les classes. L'erreur est

$$L = -\log p_y \quad (2.10)$$

où y est la classe cible, c'est-à-dire que l'on veut maximiser $p_y = P(Y = y|x)$, qui est un estimateur de la probabilité conditionnelle de la classe y étant donnée l'entrée x .

Voyons maintenant comment se fait, dans cette structure, l'application récursive de la règle de dérivation en chaîne dans un graphe de flot. Soit

$$a_k = b_k + W_k h_{k-1} \quad (2.11)$$

l'argument de la non-linéarité à chaque couche, que nous appellerons le potentiel, notons que (par une dérivation rapide)

$$\frac{\partial(-\log p_y)}{\partial a_{k,i}} = (p_i - 1_{y=i}) \quad (2.12)$$

et que

$$\frac{\partial \tanh(u)}{\partial u} = (1 - \tanh(u)^2) \quad (2.13)$$

Maintenant, appliquons la méthode de la rétro propagation dans le graphe de flot correspondant. Chaque paramètre (chaque poids et chaque biais) est un nœud, le potentiel $a_{k,i}$ et la sortie $h_{k,i}$ de chaque neurone est aussi un nœud.

- On part du nœud de sortie:

$$\frac{\partial L}{\partial L} = 1$$

- Ensuite, on calcule le gradient de L par rapport à chacune des sommes précédant le softmax, $a_{k,i}$:

$$\frac{\partial L}{\partial b_{k,i}} = \frac{\partial L}{\partial a_{k,i}} \frac{\partial a_{k,i}}{\partial b_{k,i}} = \frac{\partial L}{\partial a_{k,i}} \quad (2.14)$$

- Maintenant, on peut répéter la procédure pour chaque couche. Pour $k = K$ descendant jusqu'à 1 :
- On obtient facilement le gradient de L par rapport aux biais :

$$\frac{\partial L}{\partial b_{k,i}} = \frac{\partial L}{\partial a_{k,i}} \frac{\partial a_{k,i}}{\partial b_{k,i}} = \frac{\partial L}{\partial a_{k,i}} \quad (2.15)$$

- On calcule le gradient par rapport aux poids :

$$\frac{\partial L}{\partial W_{k,i,j}} = \frac{\partial L}{\partial a_{k,i}} \frac{\partial a_{k,i}}{\partial W_{k,i,j}} = \frac{\partial L}{\partial a_{k,i}} h_{k-1,j} \quad (2.16)$$

- On rétropropage le gradient vers la couche précédente, si $k > 1$:

$$\frac{\partial L}{\partial W_{k,i,j}} = \sum_i \frac{\partial L}{\partial a_{k,i}} \frac{\partial a_{k,i}}{\partial h_{k-1,j}} \sum_i \frac{\partial L}{\partial a_{k,i}} W_{k,i,j} \quad (2.17)$$

$$\frac{\partial L}{\partial a_{k-1,j}} = \frac{\partial L}{\partial h_{k-1,j}} \frac{\partial h_{k-1,i}}{\partial a_{k-1,j}} = \frac{\partial L}{\partial h_{k-1,j}} (1 - h_{k-1,j}^2) \quad (2.18)$$

Cet algorithme évalue l'erreur de gradient pour chaque neurone dans le réseau depuis la dernière couche jusqu'à la première. Les poids produisant une erreur significative seront modifiés en grande valeur par rapport aux poids produisant une erreur marginale. La rétropropagation par descente de gradient a été efficacement employée pour l'apprentissage d'un MLP ; son principe est de converger de manière itérative vers un ensemble optimal des poids synaptiques. Dans le cas de l'apprentissage supervisé, les sorties sont fournies à l'avance ; la différence entre ces sorties et les sorties du réseau (après propagation de l'information par une fonction d'activation) constitue l'erreur à réduire au minimum par rétropropagation tout en ajustant les poids [R.A.Rojas, 1996], [D.F.A.Santiago et al, 2002]. Cette étape est répétée plusieurs fois jusqu'à ce que le réseau puisse offrir la meilleure prévision. Dans certains cas l'algorithme de la rétropropagation de gradient ne peut pas

échapper des optimums locaux, pour cette raison, un poids momentum a été ajouté à la formule de rétropropagation permettant à l'algorithme d'échapper aux optimums locaux.

Pour des raisons de simplicité, nous avons choisi de schématiser ce processus de recherche à travers la figure suivante qui résume les étapes principales de l'algorithme de la rétropropagation de gradient :

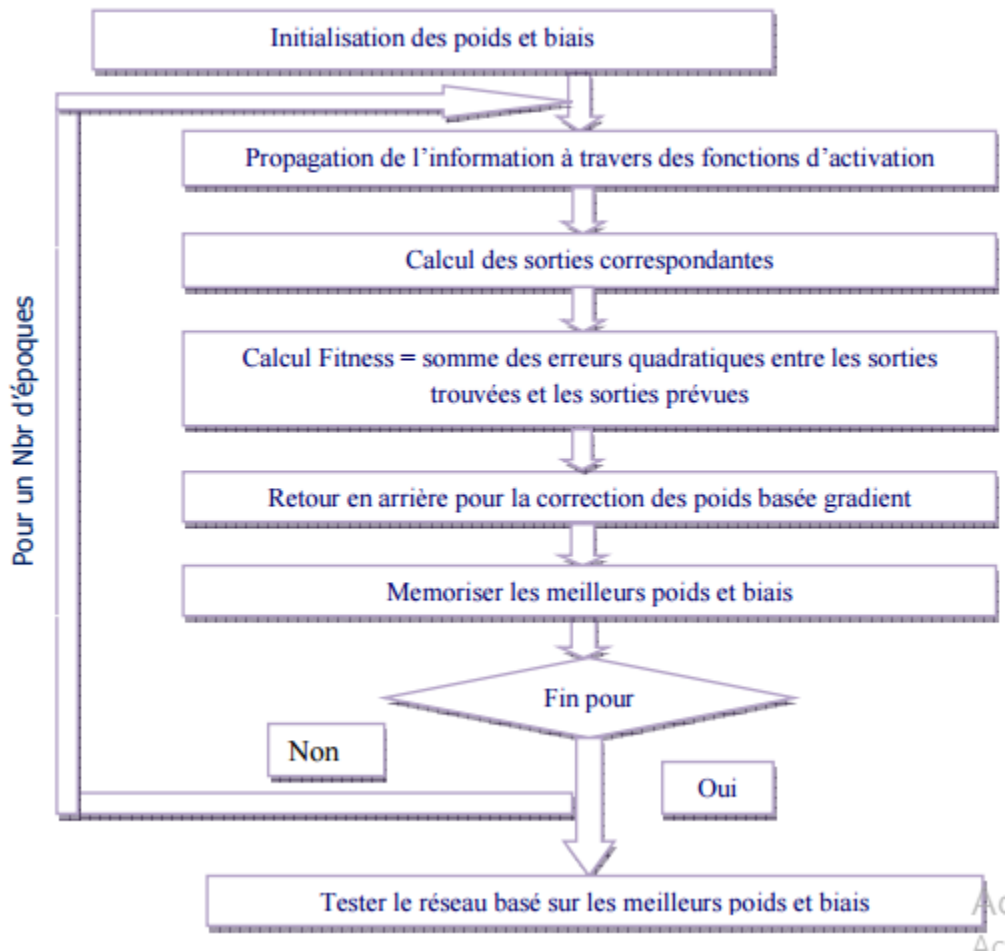


Figure 2.8:L'algorithme de la rétro-propagation de gradient

II.3.5.6.3.Fonction de transfert

Par souci de cohérence, les mêmes notations ont été conservées à travers les différents chapitres. Ainsi, les entrées d'un réseau sont encore notées X_1, \dots, X_p comme les variables explicatives d'un modèle tandis que les poids des entrées sont des paramètres α, β à estimer lors de la procédure d'apprentissage et que la sortie est la variable Y à expliquer ou cible du modèle. Un perceptron multicouche réalise donc une transformation des variables d'entrée :

$$Y = \varphi(X_1, \dots, X_p; \alpha) \quad (2.19)$$

où α est le vecteur contenant chacun des paramètres α_{jkl} de la j ème entrée du k ème neurone de la l ème couche ; la couche d'entrée ($l = 0$) n'est pas paramétrée, elle ne fait que distribuer les entrées sur tous les neurones de la couche suivante.

Un théorème dit d'approximation universelle" montre que cette structure élémentaire à une seule couche cachée est bien suffisante pour prendre en compte les problèmes classiques de modélisation ou apprentissage statistique. En effet, toute fonction régulière peut être approchée uniformément avec une précision arbitraire et dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini possédant tous la même fonction d'activation et un neurone de sortie linéaire.

De façon usuelle et en régression (Y quantitative), la dernière couche est constituée d'un seul neurone muni de la fonction d'activation identité tandis que les autres neurones (couche cachée) sont munis de la fonction sigmoïde. En classification binaire, le neurone de sortie est muni également de la fonction sigmoïde tandis que dans le cas d'une discrimination à m classes (Y qualitative), ce sont m neurones avec fonction sigmoïde, un par classe, qui sont considérés en sortie.

Ainsi, en régression avec un perceptron à une couche cachée de q neurones et un neurone de sortie, cette fonction s'écrit :

$$y = \varphi(x; \alpha, \beta) = \beta_0 + \beta' z \quad (2.20)$$

Avec
$$z_k = f(\alpha_{k0} + \alpha_k' x); k = 1, \dots, q \quad (2.21)$$

II.3.5.6.4.Apprentissage

Supposons que l'on dispose d'une base d'apprentissage de taille n d'observations $(x_1^i, \dots, x_p^i; y_i)$ 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) :

$$Q(\alpha, \beta) = \sum_{i=1}^n Q_i = \sum_{i=1}^n [y_i - \varphi(x; \alpha, \beta)]^2. \quad (2.22)$$

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

II.3.5.6.5 Rétro-propagation de l'erreur

Il s'agit donc dévaluer la dérivée de la fonction coût en une observation et par rapport aux différents paramètres. Soit $z_{ki} = f(\alpha_k 0 + \alpha_k 0 x_i)$ et $z_i = \{z_{1i}, \dots, z_{qi}\}$. Les dérivées partielles de la fonction perte quadratique s'écrivent :

$$\frac{\partial Q_i}{\partial \beta_k} = -2(y_i - \varphi(x_i))(\beta' z_i)z_{ki} = \delta_i z_{ki} \quad (2.23)$$

$$\frac{\partial Q_i}{\partial \alpha_{kj}} = -2(y_i - \varphi(x_i))(\beta' z_i)\beta_k f'(\alpha' k x_i)x_{ij} = s_{kij} \quad (2.24)$$

Les termes δ_i et s_{kij} 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 rétro-propagation : $s_{kij} = f'(\alpha' k x_i)\beta_k \delta_i$ 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 $\varphi_b(x_i)$. La passe "retour" permet ensuite de déterminer les δ_i qui sont "rétro-propagés" afin de calculer les s_{kij} et ainsi obtenir les évaluations des gradients.

II.3.5.6.6 Remarques

Les champs d'application des PMC sont très nombreux : discrimination, prévision d'une série temporelle, reconnaissance de forme. . . Ils sont en général bien explicités dans les documentations des logiciels spécialisés. Les critiques principales énoncées à l'encontre du PMC concernent les difficultés liés à l'apprentissage (temps de calcul, taille de l'échantillon, localité de l'optimum obtenu) ainsi que son statut de boîte noire. En effet, contrairement à un modèle de discrimination ou un arbre, il est a priori impossible de connaître l'influence effective d'une entrée (une variable) sur le système dès qu'une couche cachée intervient. Néanmoins, des techniques de recherche de sensibilité du système à chacune des entrées permettent de préciser les idées et, éventuellement de simplifier le système en supprimant certaines des entrées. En revanche, ils possèdent d'indéniables qualités lorsque l'absence de linéarité et/ou le nombre de variables explicatives rendent les modèles statistiques traditionnelles inutilisables. Leur flexibilité alliée à une procédure d'apprentissage intégrant la pondération (le choix) des variables comme de leurs interactions peuvent les rendre très efficaces .[18]

II.3.5.7 Le réseau de neurone fonctions à base radial

II.3.5.7.1 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. [1] 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.

Le réseau de fonctions à base radial (RBF) est un réseau de neurones feed d'une couche d'entrée, une seule couche cachée sortie. Le réseau RBF est conçu pour effectuer une combinaison non à l'espace caché, suivi d'une combinaison linéaire de l'espace caché à l'espace de sortie [Dumitrescu et al, 2005].

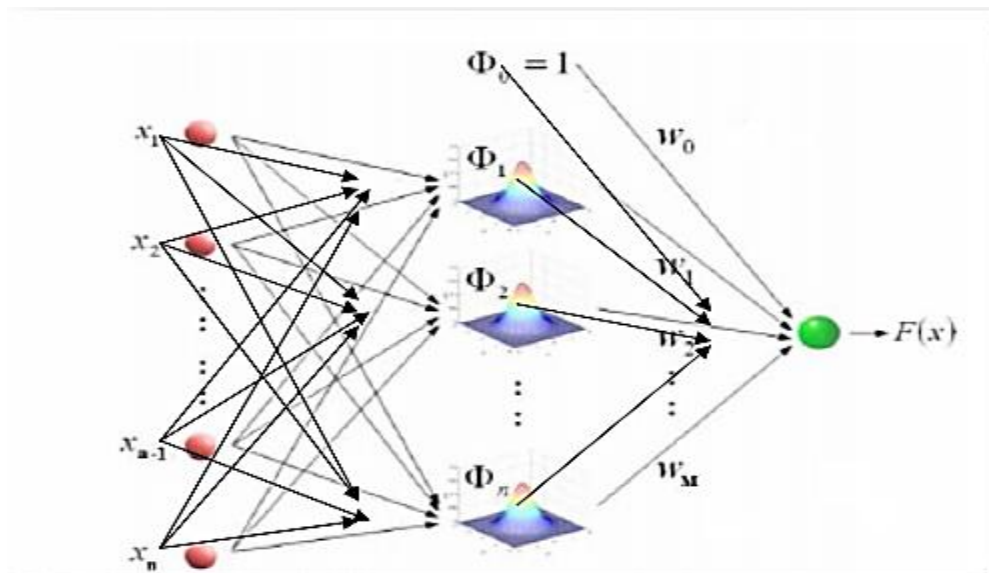


Figure 2.9 :structure d'un RBF

Les réseaux (RBF) peuvent être employés pour un grand nombre d'applications parce qu'ils peuvent optimiser n'importe quelle fonction régulière et leur apprentissage est plus rapide qu'un perceptron multicouche. Cette vitesse vient du fait qu'un RBF a seulement deux couches de poids qui peuvent être évalués d'une façon déterministe.

La performance d'un réseau RBF dépend du nombre et position des fonctions à base radiales, de leur forme et de la méthode employée pour son apprentissage [C. Wu et al, 2003].

Les stratégies de sélection des centres d'un RBF peuvent être classifiées comme suit :

- les stratégies choisissant de manière aléatoire les centres d'un RBF à partir des exemples d'apprentissage
- les stratégies non supervisées comme les algorithmes génétiques [B.A. Whitcheas et al, 1994],
- Les stratégies supervisées comme le calcul des moindres carrés [J.B. Gomm et al, 2000], affectation de ressources ou ceux qui emploient des algorithmes du clustering comme K organisatrices [Y. Hwang et al, 1997].

II.3.5.7.2 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 identique à celle d'un PMC à une seule couche cachée donc on peut dire qu'il prend toutes les caractéristiques d'un PMC simple sauf qu'il diffère 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 sûr, 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 supervisé pour les deux types de réseaux.

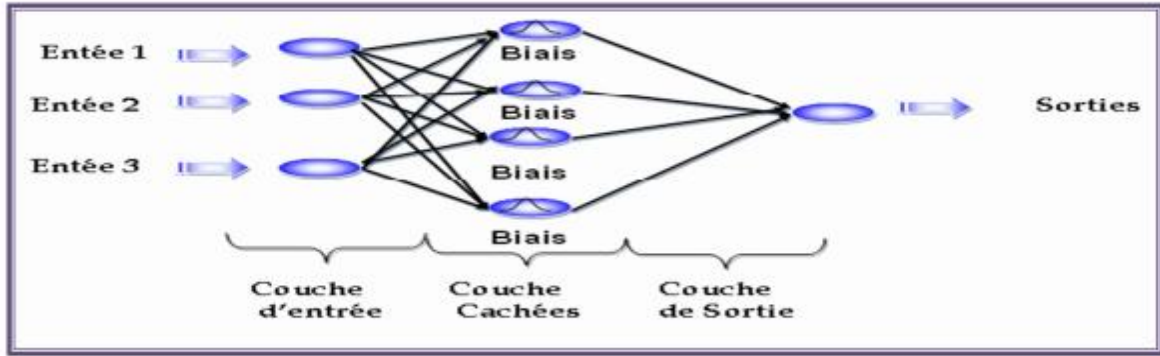


Figure 2.10 : Architecture d'un Réseau de Neurone RBF

II.3.5.7.3 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],[2]. 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 :

$$H(y)=\exp (-|y-c|^2/2\mu^2) \quad (2.25)$$

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 RBFs 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. $H(y)=\prod \exp (-|y_i-c_i|^2/2\mu_i^2)$ Avec $y = [y_i]$ et $C = [C_i]$. Cette particularité devient intéressante pour l'adéquation biologique des réseaux de neurones artificiels.

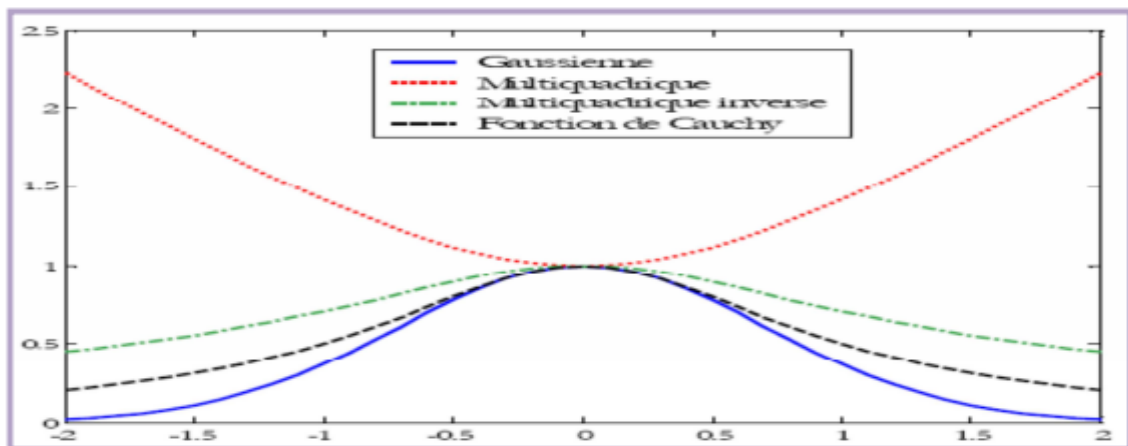


Figure 2.11 : Quelques fonctions radiales

II.3.5.7.4 La 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.

II.3.5.7.5 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 :

$\beta = \frac{d}{\sqrt{M}}$ avec M = nombre des centroïdes 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 : $\beta_j = \|x_i - c_j\|$

II.3.5.7.6 L'étape de détermination des poids du RBF:(supervisé)

Une fois le nombre et la position des centroïdes et la largeur des gaussiennes fixés, les poids de chacune des connexions (RBF-output) peuvent être calculé par l'équation matricielle suivante : [28], [29]

$$\begin{bmatrix} f(\|x_1 - c_1\|) & f(\|x_1 - c_M\|) \\ f(\|x_2 - c_1\|) & f(\|x_2 - c_M\|) \\ f(\|x_N - c_1\|) & f(\|x_N - c_M\|) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_M \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_N \end{bmatrix} \quad (2.26)$$

Avec $f(x) = \text{EXP}(x^2 / (2 * \beta^2))$, la gaussienne. La matrice colonne Y, les outputs attendus M, le nombre de centroïdes. 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 :

$$W=A^{-1}y \quad (2.27)$$

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 à 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 utilisée est le calcul de la matrice pseudo inverse de la méthode nommée SVD est typiquement une bonne méthode pour ce genre de calcul. On cherche la matrice 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 : $U = \{u_1, u_2, u_3, \dots, u_n\}$, $v = \{v_1, v_2, v_3, \dots, v_n\}$

Tel que :

$$u^T A V = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k) \quad (2.28)$$

Avec diag matrice diagonale et $k = \min(M, N)$. Le pseudo inverse de A est alors donnée par

$$A^+ = V \Sigma u^T \text{ avec } \Sigma = \text{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_k}, 0, \dots, 0\right) \quad (2.29)$$

II.3.5.7.7 L'algorithme de rétro-propagation du gradient

L'algorithme de rétro-propagation a été développé pour entraîner des perceptrons multicouches. L'idée est d'entraîner un réseau en propageant les erreurs de sortie en sens inverse du réseau à 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 coût suivante :

$$C = 1/2 \sum (y_n^2 - y_n)^2 \quad (2.30)$$

Avec les y_n sont les sorties désirées pour chaque exemple d'apprentissage x_i , et les \hat{y}_n

II.3.5.7.8 La phase de test :

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

$$[f(|t_1 - c_1|) \dots f(|t_1 - c_M|)] \text{Output_test} \quad (2.31)$$

Avec $f(x) = \text{EXP}(x^2 / (2 * \beta^2))$, la gaussienne. La matrice colonne W, les poids des connexions (RBF-output), M, le nombre des centroïdes. t_i , le i ème élément de la 'base de test'.

II.3.5.7.9 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 .[28],[29]

Mais cet 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).

II.3.5.8 Comparaison MLP et RBF

La différence principale d'un RBF par rapport à un MLP est l'absence des poids de la -couche cachée. Aussi, les sorties cachées ne sont pas calculées en utilisant une fonction d'activation sigmoïde, elles se sont basées sur des fonctions à base radiale. L'apprentissage d'un réseau RBF est radicalement différent de l'apprentissage classique d'un réseau MLP. Dans ce cas, il n'y a aucun poids caché à changer par la méthode du gradient. Dans les réseaux RBF, l'apprentissage est basé sur le choix des centres et leurs largeurs aussi bien que les poids des neurones de sortie.

	Le perceptron multicouche	Les réseaux de neurones à base radiale
Les unités cachées	Fonction hyper-plan	Fonction hyper-sphère
Les couches cachées	Plusieurs	Une seule couche cachée
Apprentissage	Supervisé Lent	Non supervisé+combinaison linéaire rapide

Table 2.12: MLP versus RBF

Les réseaux de neurones du type MLP et RBF sont capables de converger avec n'importe quel type de problème ; la principale différence porte sur l'architecture qui consiste de deux couches seulement pour les RBF et peut comporter plus pour les MLP. Le réseau RBF souffre du problème d'identification de ces paramètres de contrôle qui sont nombreux notamment, la détermination des centres des fonctions radiales. Leurs rayons (largeurs) et les poids de la couche de sortie. Si l'un de ces paramètres n'est pas identifié précisément, Le réseau RBF peut produire des résultats néfastes. En outre, la taille du réseau RBF augmente considérablement avec le nombre des exemples d'apprentissage, augmentant ainsi le temps de calcul. Cependant un réseau MLP avec Une seule couche cachée peut avoir une grande capacité de généralisation sur un grand nombre d'exemples d'apprentissage. [30]

II.4 Conclusion:

Dans ce chapitre , après une définition des notations , nous avons faite un état de l'art de différentes techniques de réduction de la dimension ,puis nous avons présenter la méthode que nous avons utilisées :LDA , en fait nous avons expliquées le fondement mathématique de la LDA ; et son utilisation avec exposition de quelques avantages et inconvénients de cette méthode.

Ainsi que , nous avons parlé en détails de notre modèle de calcul de base: « les réseaux neurone » leurs applications, leurs classes et leurs caractéristiques... en particulier la technique MLP , qui est une technique spéciale des réseaux de neurones caractérisée par son modèle simple et sont temps de calcul qui est extrêmement rapide par rapport aux autres types de réseaux afin de prédire la valeur future de la vitesse du vent.

Contrairement aux méthodes classiques qui ont montré leurs limites, les réseaux de neurones ont montré leurs tendances à s'adapter à des problèmes complexes grâce à leur grande capacité de calcul et d'apprentissage. Ils sont l'objet d'utilisation dans les différents domaines tels que : La reconnaissance des formes et le traitement des images.

Le grand avantage caractérisé dans les réseaux de neurones à base radiale est que ces derniers n'ont qu'une seule couche cachée, ce qui les rend les plus simples à utiliser et les plus rapides.

CHAPITRE 03

Résultats et discussion

Sommaire

3.1 INTRODUCTION.

3.2 PRESENTATION DE LA BASE DE DONNEES .

3.3. APPLICATION DE LA TECHNIQUE LDA AVEC LA METHODE MLP et RBF.

3.4. DISCUSSIONS DES RESULTATS.

3.5 CONCLUSION.

3.1 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. Nous illustrons l'intérêt de la technique développée pour la prédiction du vent.

Les travaux de ce mémoire consistent à étudier les méthodes de la prédiction de la vitesse du vent, en utilisant la technique de réduction de la dimension LDA et les réseaux de neurones de types MLP et RBF en régression, ceci en se basant sur deux critères de performances, à savoir : la racine carrée de l'erreur quadratique moyenne (RMSE) et l'erreur moyenne de pourcentage absolue (MAPE) , le Biais moyen absolue (MABE)et le Coefficient de détermination (R).

3.2 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 Rewable Energy Laboratory).

3.2.1 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 ($41^{\circ} 6' N, 72^{\circ} W$).

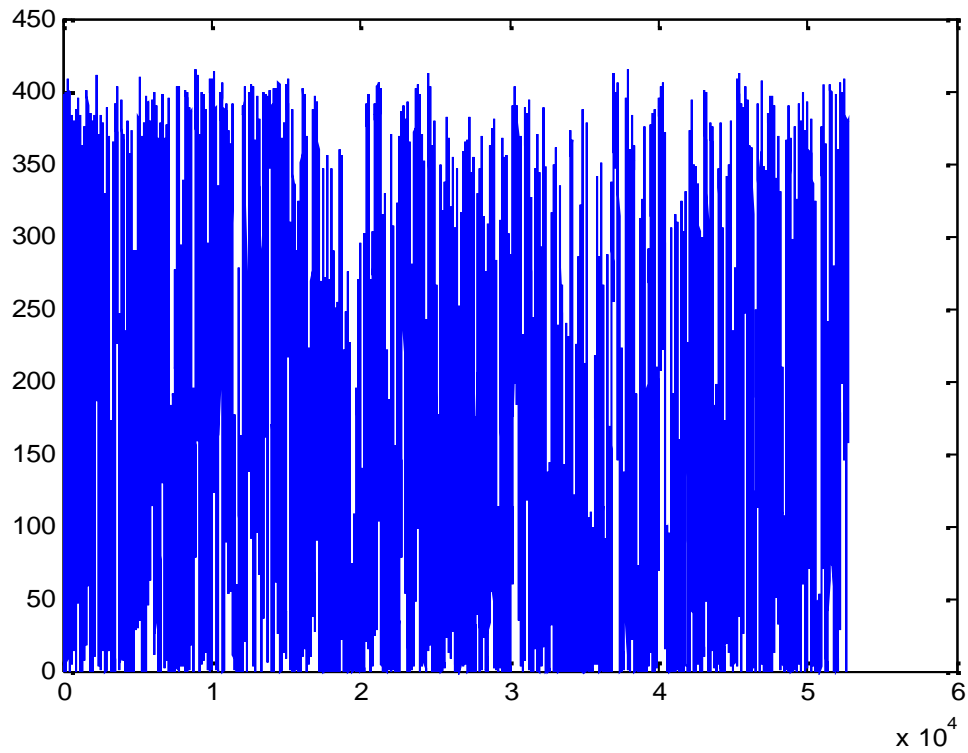
Le deuxième site se situe dans l'état de Connecticut (est une état du Nord-est des États Unis), les coordonnées du site sont ($39^{\circ} N, 105^{\circ} W$).

Site	Longitude	Latitude	L'altitude	Zone(km ²)	Densité (MW/km ²)	Total MW pour site
Colorado	-103.51	40.8	1358	58.4	14.4	839.7

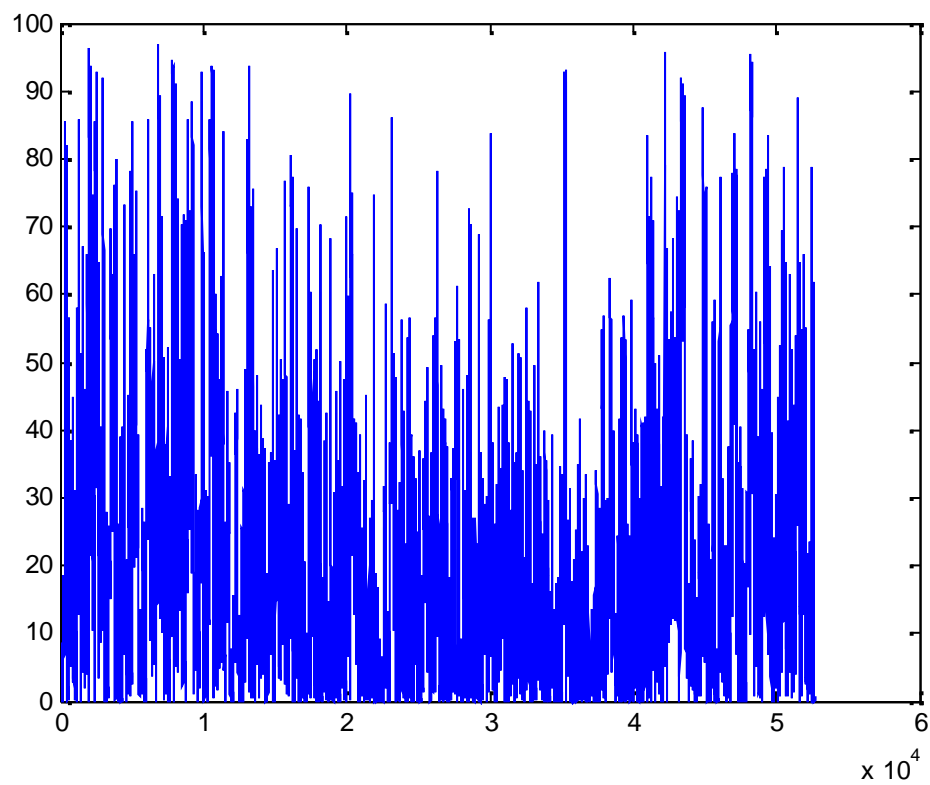
Tableau 1. Caractéristiques géographiques du site Colorado.

Site	longitude	latitude	L'altitude	Zone(km ²)	Densité (MW/km ²)	Total MW pour site
Connecticut	-73.52	41.54	295	5.4	18.6	100.4

Tableau 2. Caractéristiques géographiques du site Connecticut.



Site Colorado



Site Connecticut

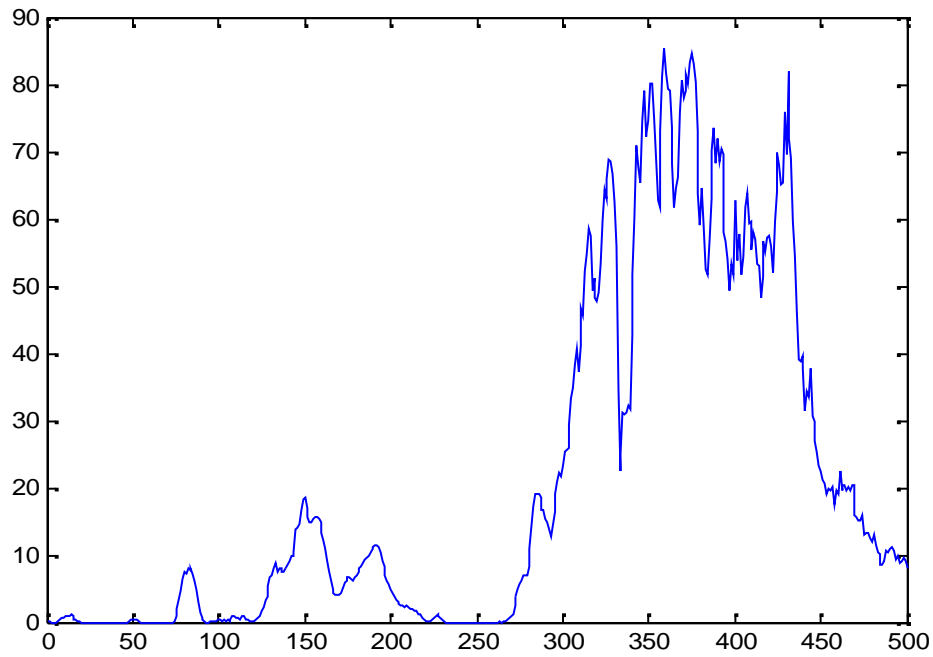
Figure 3.1 : les courbes de la base de données pour les deux sites.

3.3.APPLICATION DE LA TECHNIQUE LDA AVEC LES RESEAUX DE NEURONES TYPE MLP et RBF

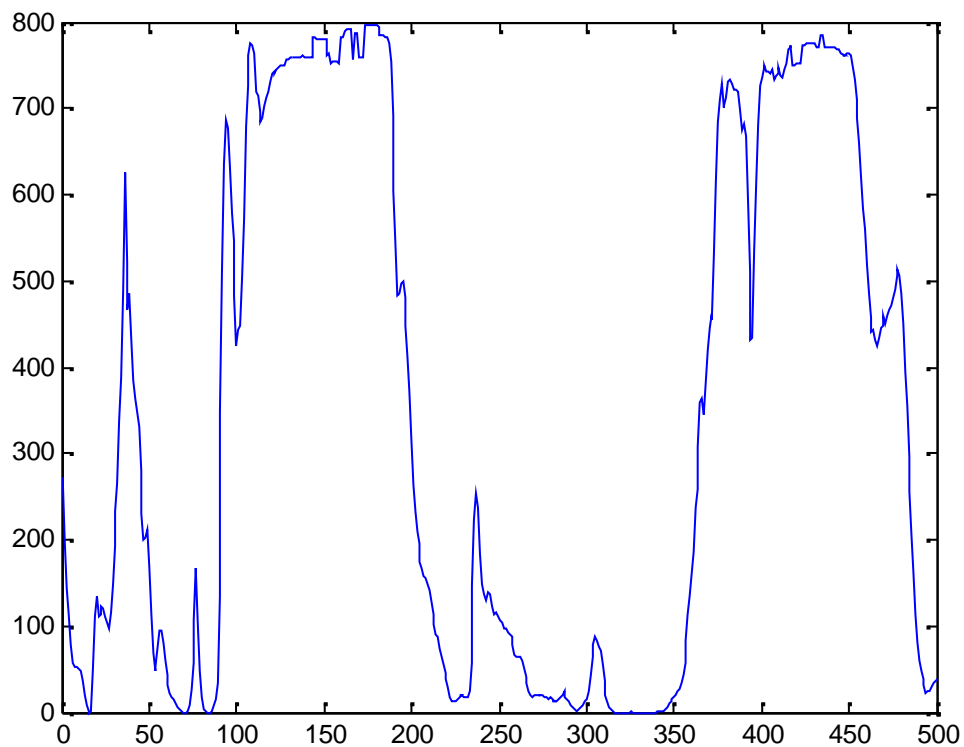
Il existe un grand nombre de types de réseaux de neurones, dont chacun a des avantages et des inconvénients. Le réseau choisit dans notre cas est un réseau multicouches(MLP). Ce choix est fait pour la facilité et la rapidité de sa construction et encore par le fait que notre problème présente un nombre limité de variables d'entrées.

3.3.1 Etape apprentissage et validation sur la base complète

Nous obtenons l'erreur de validation sur la base complète (LDA), le rôle ici est de chercher le nombre de neurones dans la couche cachée du réseau MLP.



Site Colorado



Site Connecticut.

Figure 3. 2 : Les courbes d'erreur de validation Base complète pour les deux sites

3.3.2 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 RMSE et (MAPE) et (MABE) et (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].

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (3.1)$$

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 MPAAE :**

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

$$\text{MAPE} = \left(\frac{1}{m} \sum_{t=1}^m |y_t - \hat{y}_t| \right) \times 100 \quad (3.2)$$

$$\text{MAPE} = \frac{1}{N_f} \sum_{t=1}^{N_f} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100 \quad (3.3)$$

➤ **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 R^2 se définit comme la part de variance exprimée par rapport à la variance totale, autrement dit par [1-(somme des carrés des résidus / variance totale)]

$$R^2 = 1 - \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y})^2} \quad (3.4)$$

➤ **Biais Moyen Absolue**

Il est défini comme étant l'écart algébrique moyen entre la simulation et la mesure, sa formule est :

$$M = \frac{1}{N} \sum (y - \hat{y}) \quad (3.5)$$

Une valeur positive de *MABE* signifie que sur le long terme, le cumul des données de prédiction sous-estime le cumul des mesures réelles alors qu'une valeur négative signifiera l'inverse. L'utilisation de ce genre d'outil n'est pas triviale. En effet, il faut savoir qu'une valeur nulle de *MABE*, ne signifie pas forcément que la prédiction est parfaite.

3.4. DISCUSSIONS DES RESULTATS

Dans le cadre de ce travail, nous avons utilisé les réseaux de neurones non récurrents de types PMC et RBF. Le choix d'exploiter un type de réseaux de neurones par rapport à un autre n'est pas arbitraire, mais repose sur les fondements qu'exige le problème en question. Sachant que chaque couche type de réseaux, en termes d'apprentissage, est conseillé pour une application donnée. Nous avons opté donc pour le PMC puisqu'il répond aux spécificités du problème de prédiction. Ce réseau apporte des résultats très significatifs et le taux de précision de la prédiction étant meilleur.

Le réseau RBF fait partie des réseaux de neurones supervisés. Il est constitué de trois couches : une couche d'entrée qui retransmet les entrées sans distorsion, une seule couche cachée qui contient les neurones sont généralement animés par une fonction d'activation linéaire.

Le réseau de neurone artificiel RBF se compose d'une couche d'entrée contenant sept neurones, d'une couche cachée et d'une couche de sortie contenant un seul neurone. Les variables d'entrées se composent du vecteur I_1 , indépendant et normalisé entre -1 et 1 qui est : la vitesse du vent.

La sortie de chaque neurone de la couche cachée sera amplifié par les poids. Ces derniers ont été calculés par l'équation de pseudo-inverse, puis rajoutés au biais b_i de la couche de sortie et transformés enfin par une fonction linéaire. Nous obtenons alors le modèle du réseau RBF, tout en fixant les poids et les biais relevés dans la phase d'apprentissage. La partie test consiste à valider l'architecture neuronale RBF de configuration en jugeant sa capacité de prédiction. Nous avons par ailleurs comparé les performances des modèles statistiques neuronaux PMC et RBF.

Pour déterminer l'architecture des réseaux à utiliser, nous avons varié le nombre de couche cachée de MLP, le nombre de neurones dans une couche cachée de RBF, les fonctions de transfert, le nombre d'itération et le pas d'apprentissage. Pour cela nous avons divisé aléatoirement notre base de données en trois parties : 60% pour l'apprentissage, 20% pour le test et 20% pour la validation.

3.4.1 Les Résultats De Simulation Des 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 MLP et RBF avant et après LDA.

➤ **Avant LDA**

MLP	RMSE	MAPE	MABE	R
1 couche cachée	23.5895	5.7708	15.3939	0.9909
2 couche cachée	21.5533	4.8813	13.0212	0.9923
3 couche cachée	86.0418	25.6613	68.4534	0.8840
RBF	RMSE	MAPE	MABE	R
20 neurones	21.2134	4.7727	12.7316	0.9926

Tableau 3. Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site calorado.

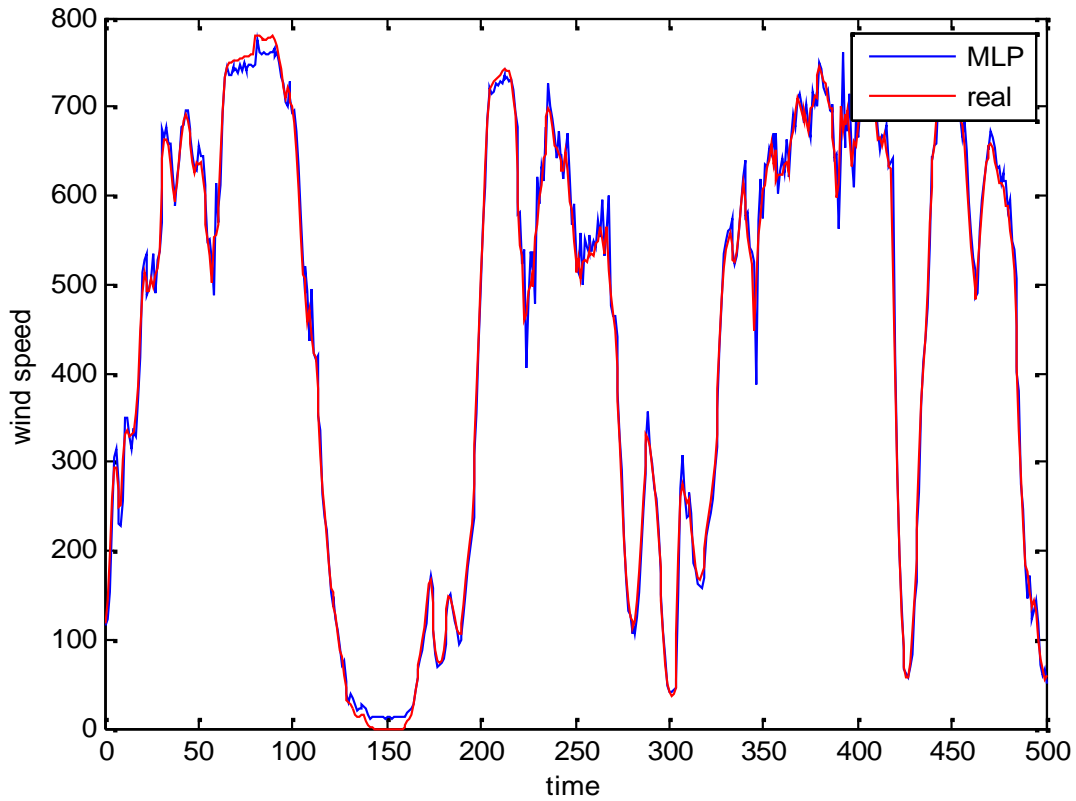


Figure3. 4 : Courbe comparaison valeurs réels et prédites de MLP pour 1 couche cachée site Colorado.

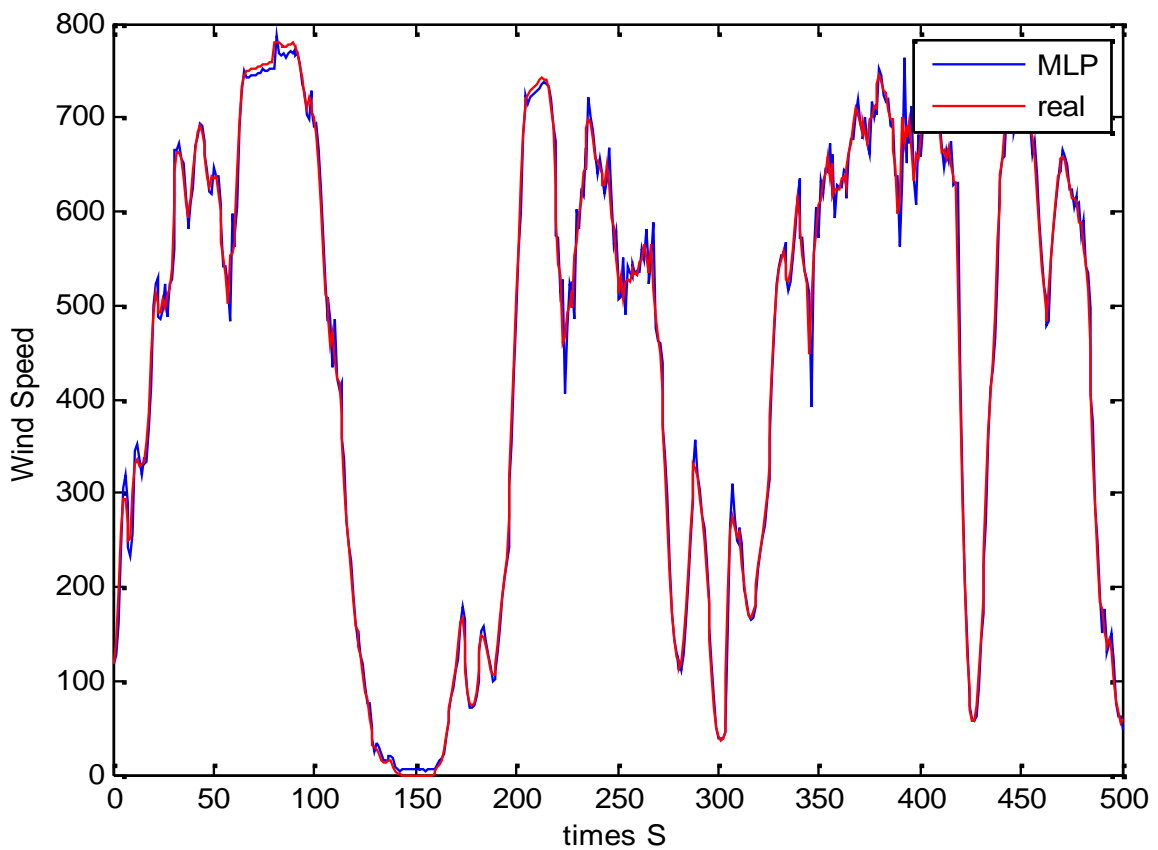


Figure 3. 5 : Courbe comparaison valeurs réels et prédites de MLP pour 2 couche cachée site Colorado.

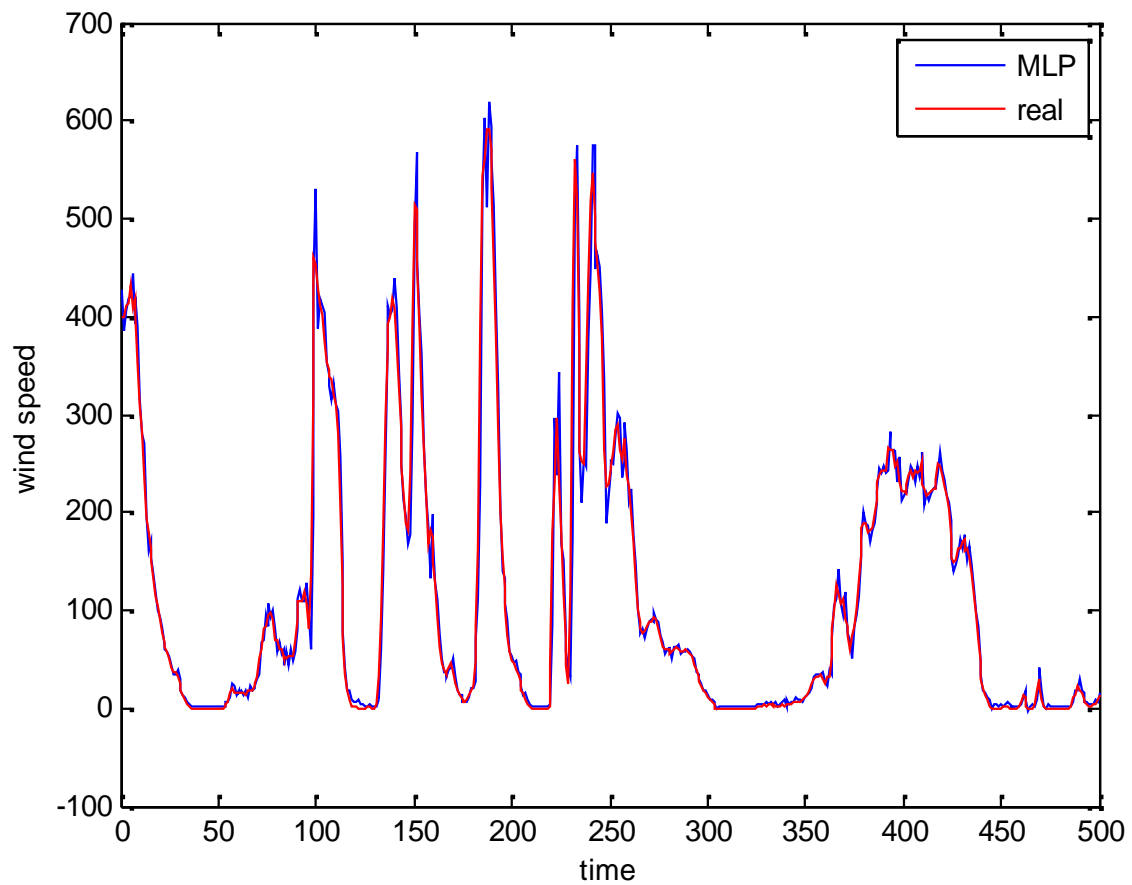


Figure 3.6 : Courbe comparaison valeurs réels et prédites de MLP pour 3 couche cachée site Colorado.

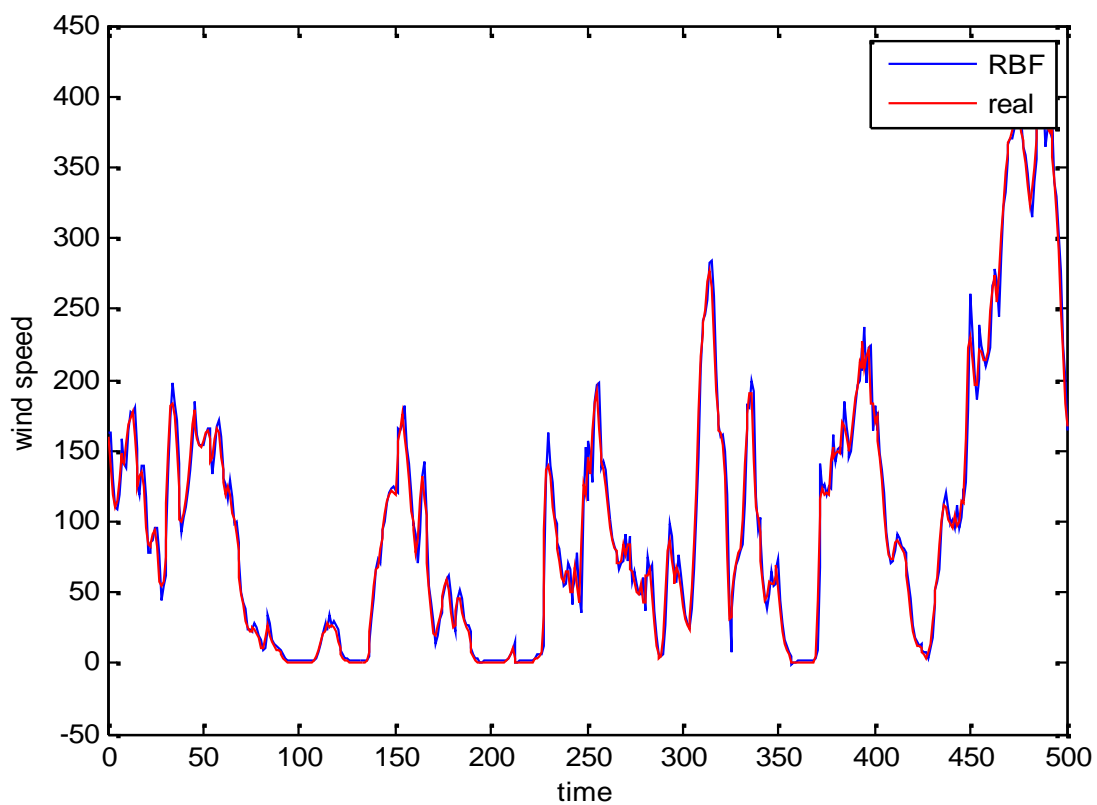


Figure 3.7 : Courbe comparaison valeurs réels et prédites de RBF pour 20 neurones site Colorado.

➤ **La prédiction par la méthode de MLP :**

Dans les tableaux suivants ,nous avons présenté les calculs des RMSE , MAPE, MABE et R pour 1,2,3 couches cachées augmente la charge des calculs sans aucun gain de performance. Nous pouvons donc, affirme que l'utilisation d'une seule couche cachée est le meilleur pour le modèle de type MLP.

	RMSE	MAPE	MABE	R
1 couche cachée	23.26679	6.2262	16.3613	0.9912
2 couche cachée	21.9743	5.3582	14.0804	0.0921
3 couche cachée	22.2201	5.6431	14.8292	0.9920

Tableau 4. Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site colorado.

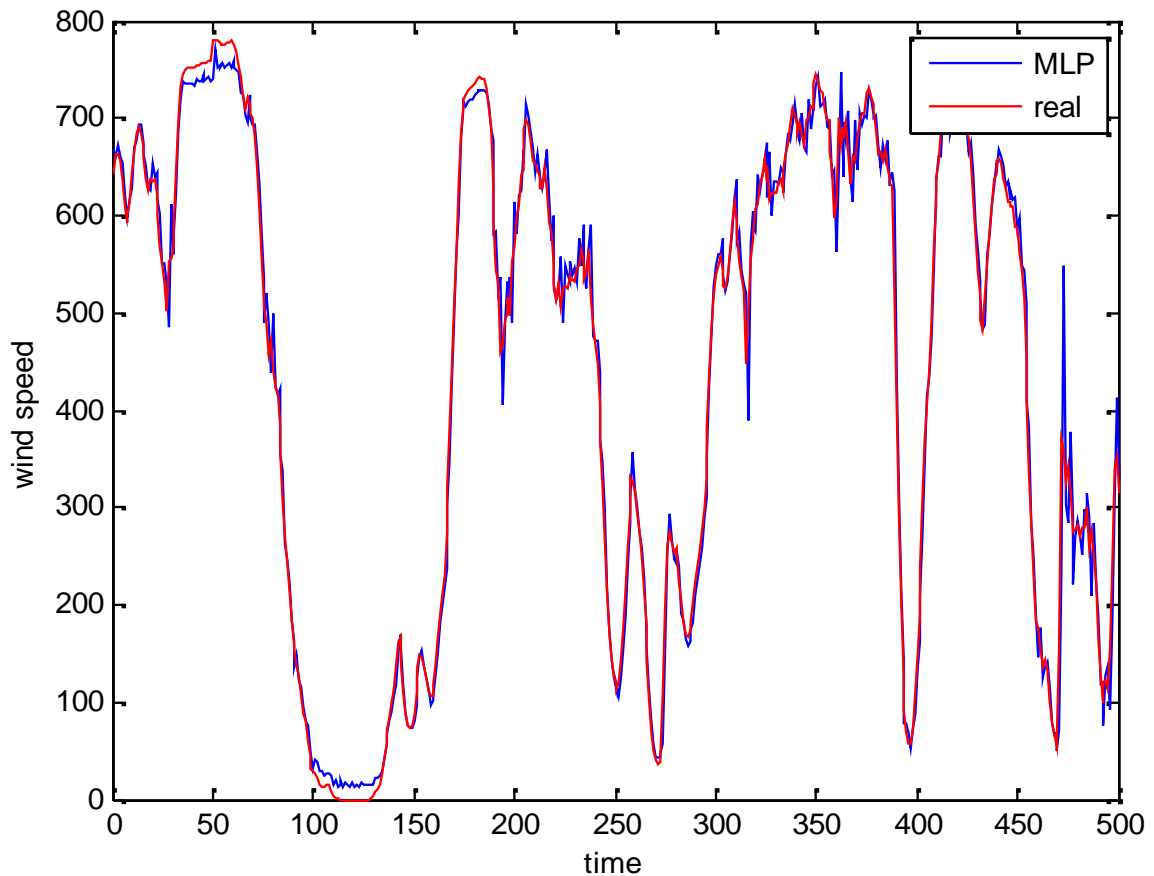


Figure 3.8 : Courbe comparaison valeurs réels et prédites de MLP pour 1 couche cachée pour le site Colorado.

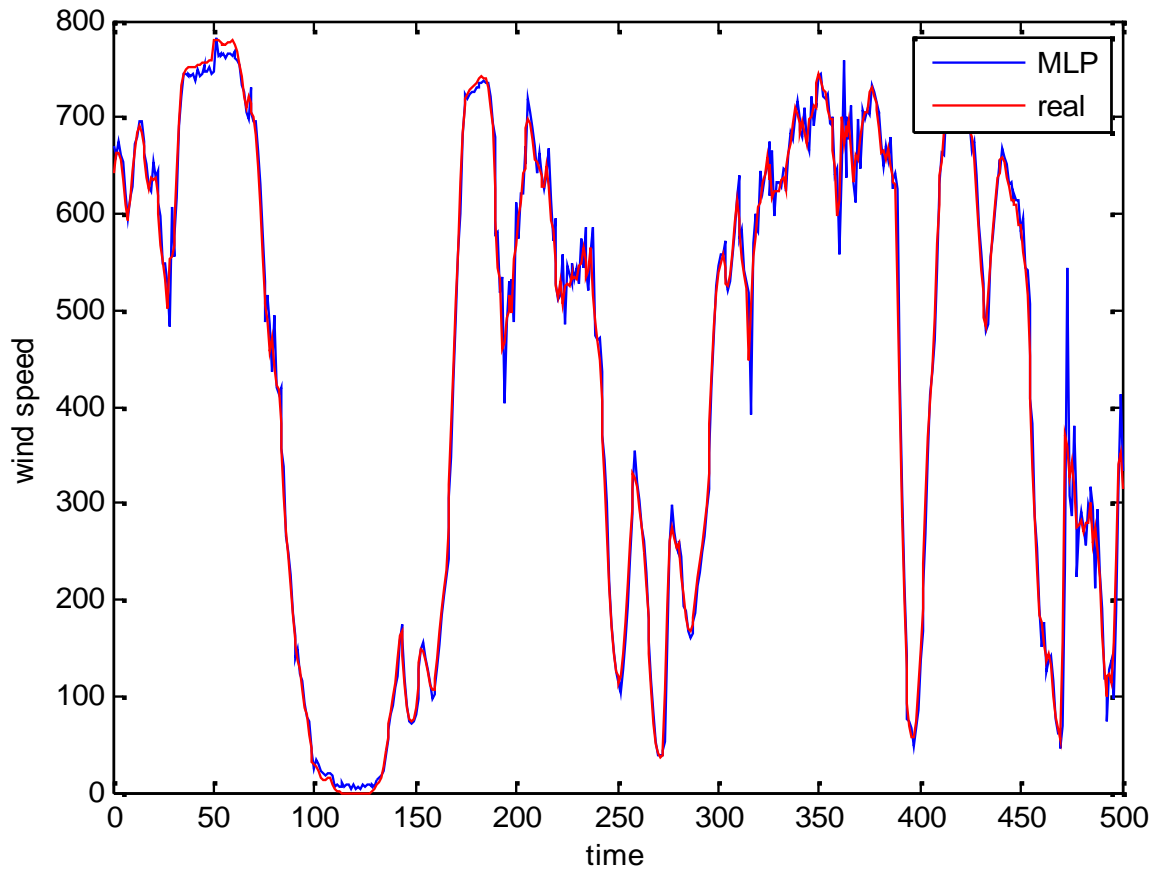


Figure 3.9 : Courbe comparaison valeurs réels et prédites de MLP pour 2 couche cachée pour le site Colorado.

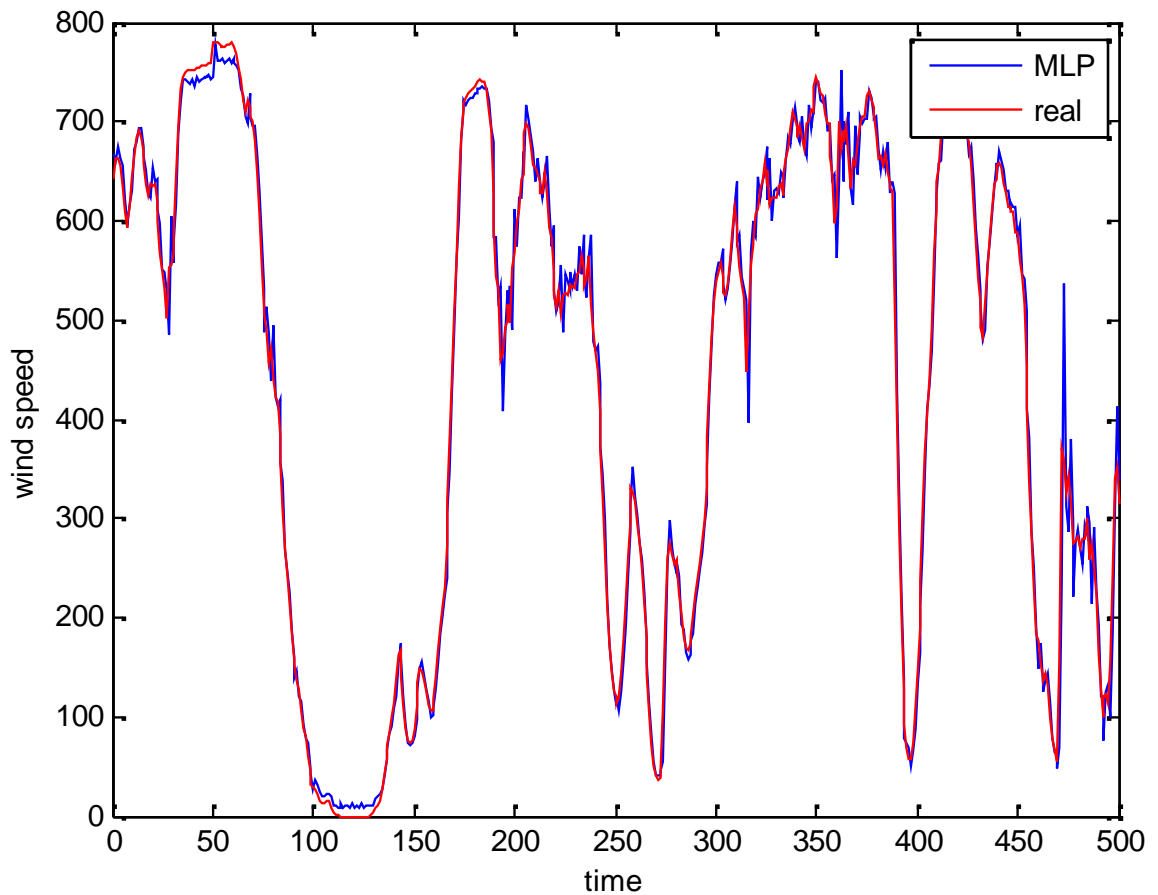


Figure 3.10 : Courbe comparaison valeurs réels et prédites de MLP pour 3 couches cachées pour le site Colorado.

	RMSE	MAPE	MABE	R
1 couche cachée	3.3352	8.2075	2.3815	0.9814
2 couche cachée	3.1005	7.3263	2.1258	0.9839
3 couche cachée	7.1445	19.2000	5.5711	0.9146

Tableau 5. Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site connecticut.

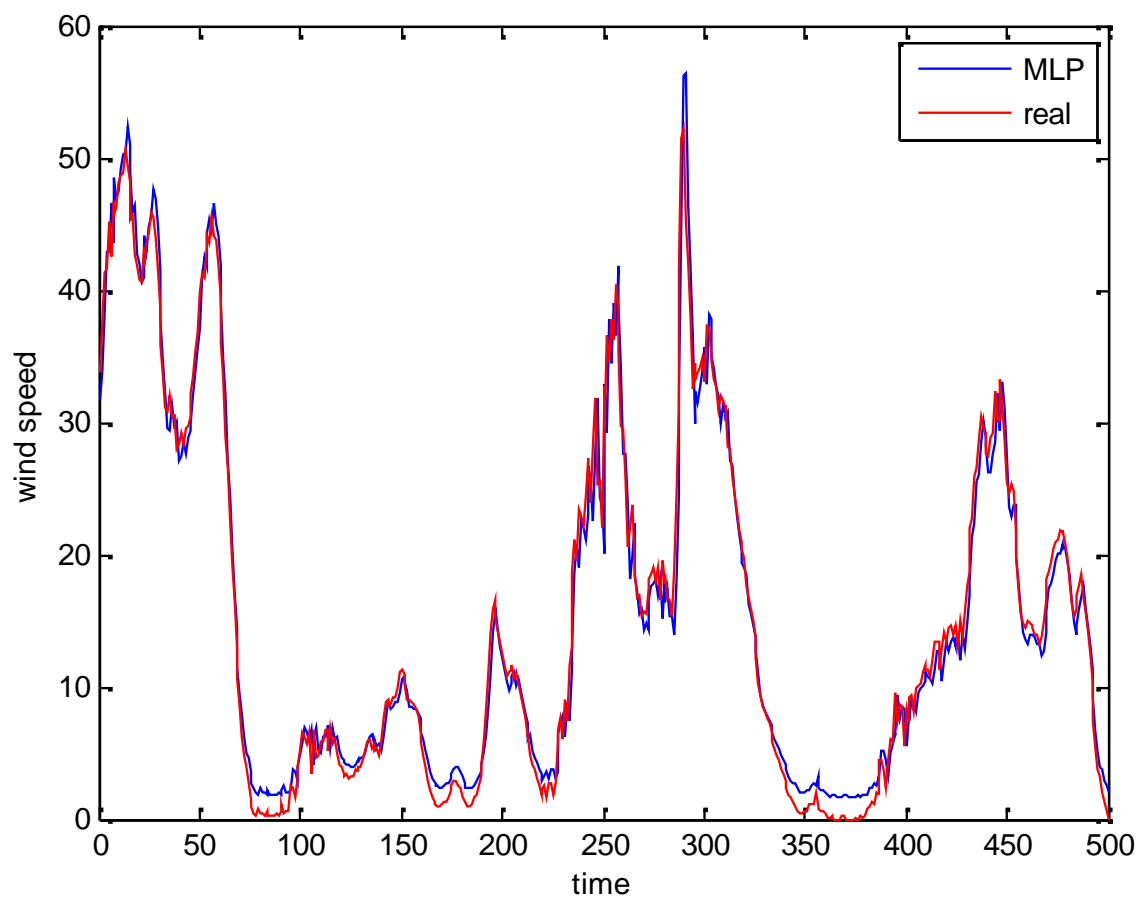


Figure 3.11 : Courbe comparaison valeurs réels et prédites de MLP pour 1 couche cachée pour le site Connecticut.

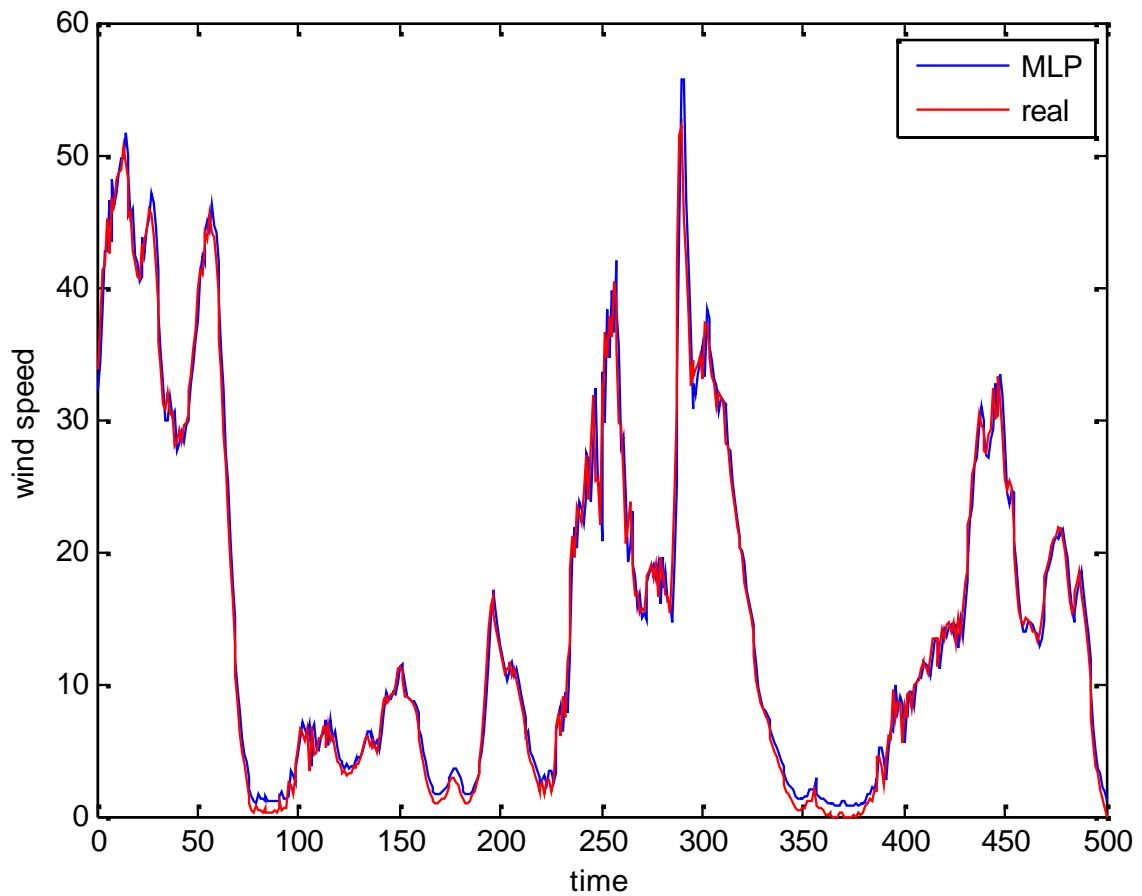


Figure 3.12 : Courbe comparaison valeurs réels et prédites de MLP pour 2 couches cachées pour le site Connecticut .

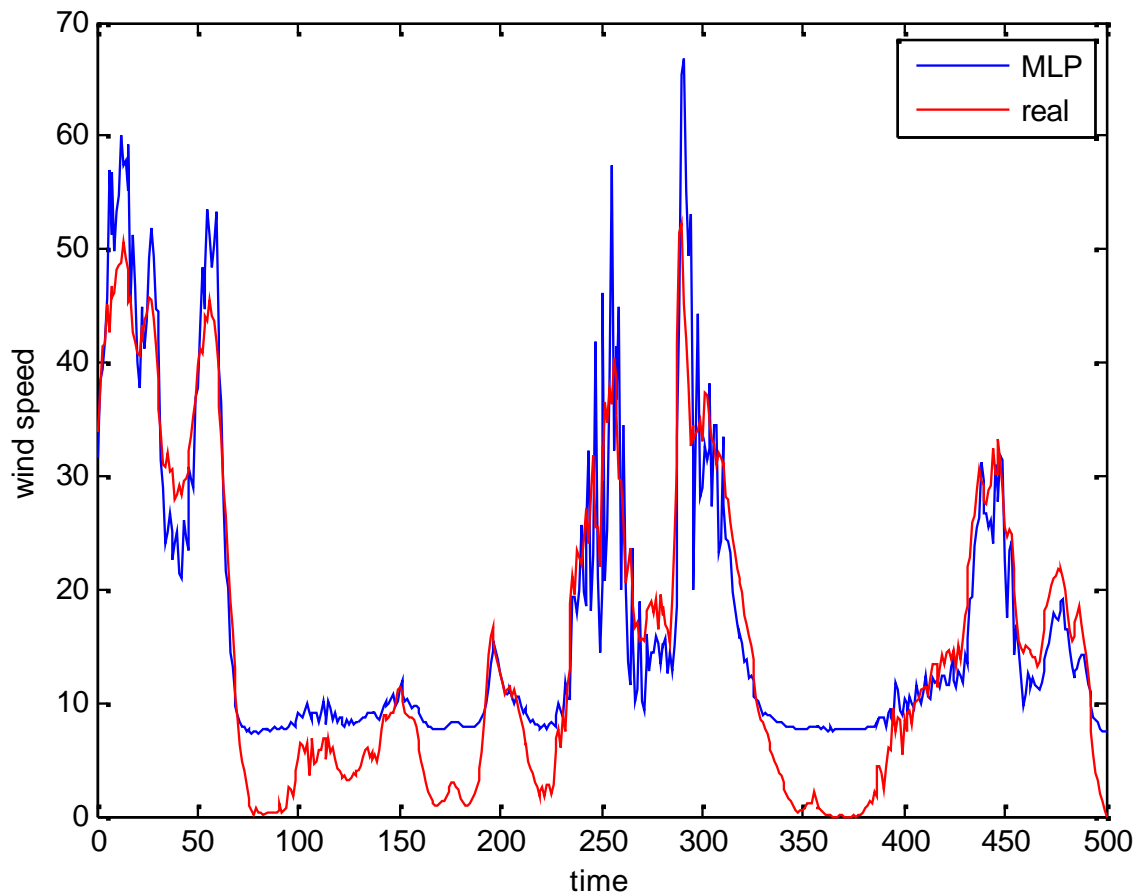


Figure 3.13 : Courbe comparaison valeurs réels et prédites de MLP pour 3 couches cachées pour le site Connecticut .

➤ La prédiction par la méthode de RBF:

	RMSE	MAPE	MABE	R
30 neurons	13.7405	4.7237	8.5262	0.9907
20 neurons	13.7399	4.7248	8.5281	0.9907
10 neurons	13.7373	4.7306	8.5386	0.9907

Tableau 6. Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site Colorado

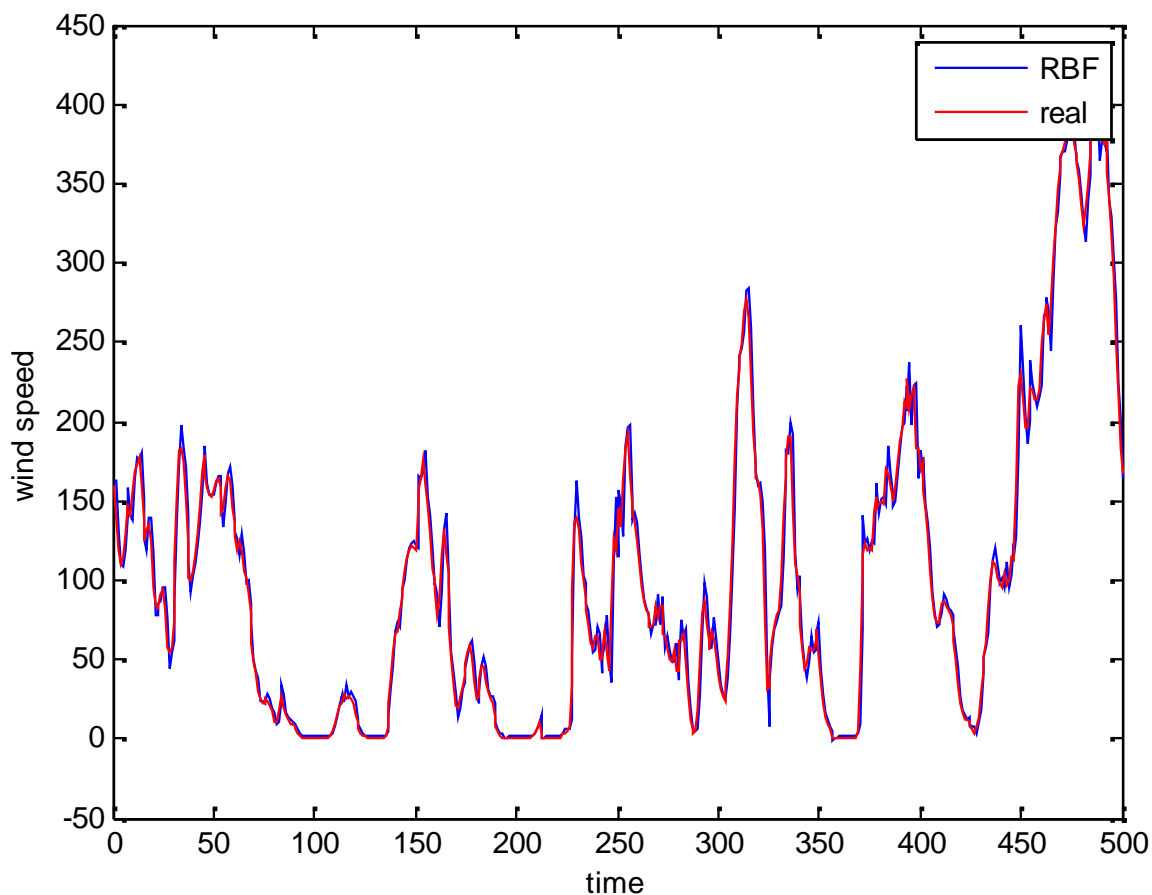


Figure 3.14 : Courbe comparaison valeurs réels et prédites de RBF pour 30 neurones pour le site Colorado.

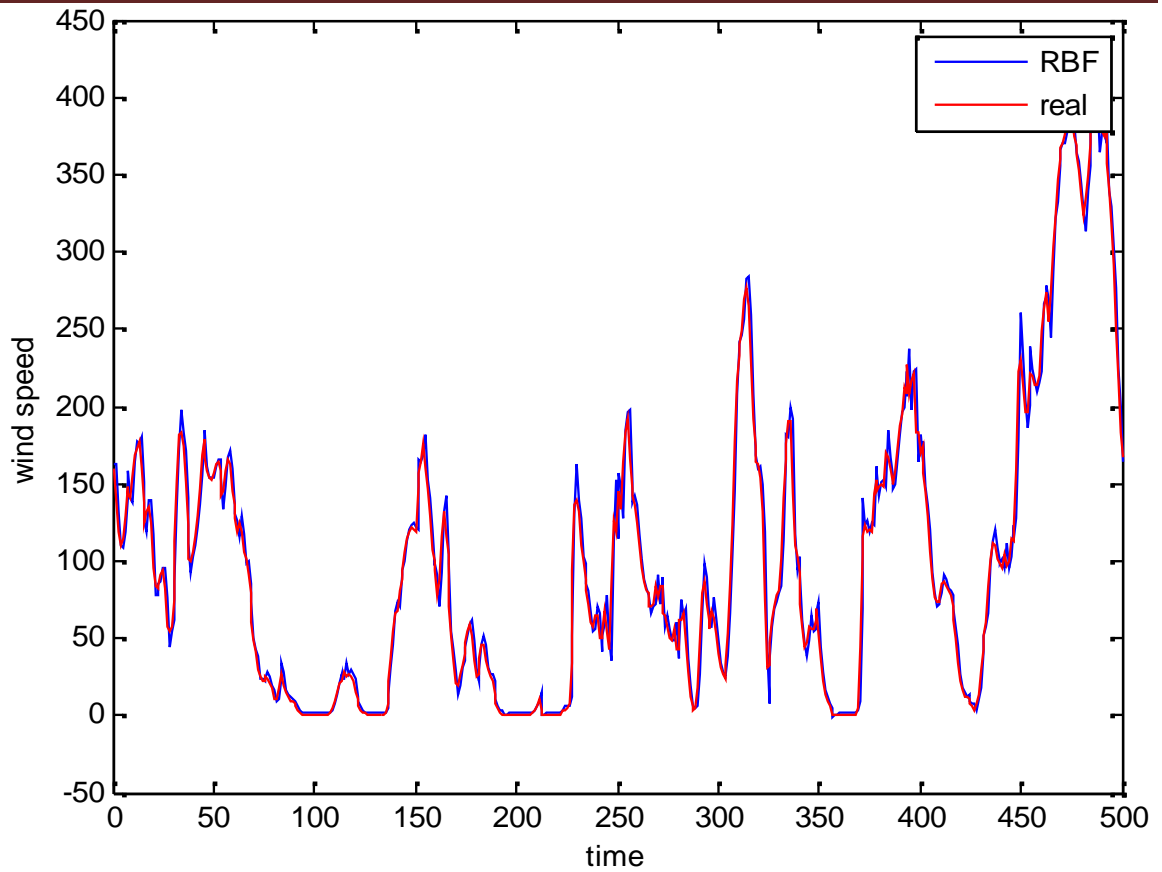


Figure 3.15 : Courbe comparaison valeurs réels et prédites de RBF pour 20 neurones pour le site Colorado.

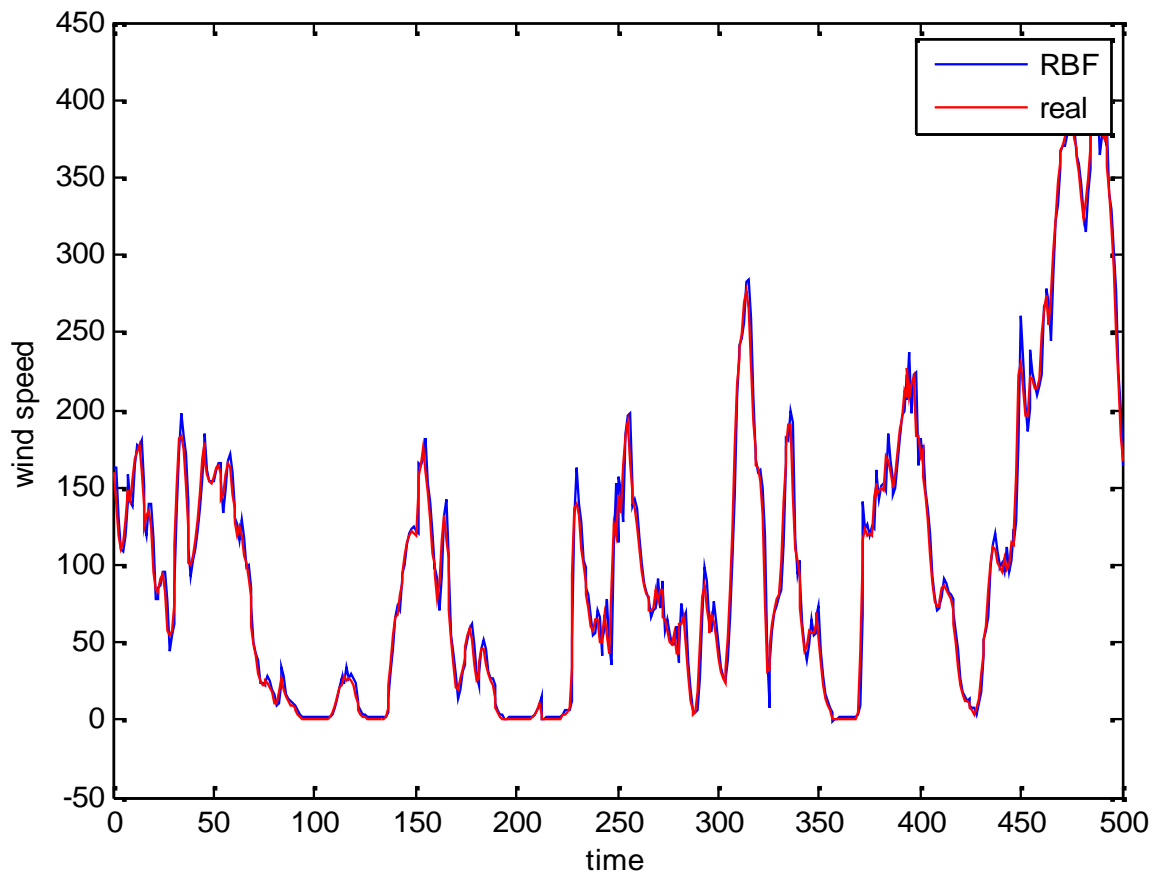


Figure 3.16 : Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Colorado.

	RMSE	MAPE	MABE	R
10 neurons	2.9973	6.9249	2.0093	0.9850
20 neurons	2.9972	6.9220	2.0085	0.9850
30 neurons	2.9971	6.9216	2.0084	0.9850

Tableau 7. Résultats des valeurs de critères RMSE, MAPE ,MABE et R pour le site Connecticut

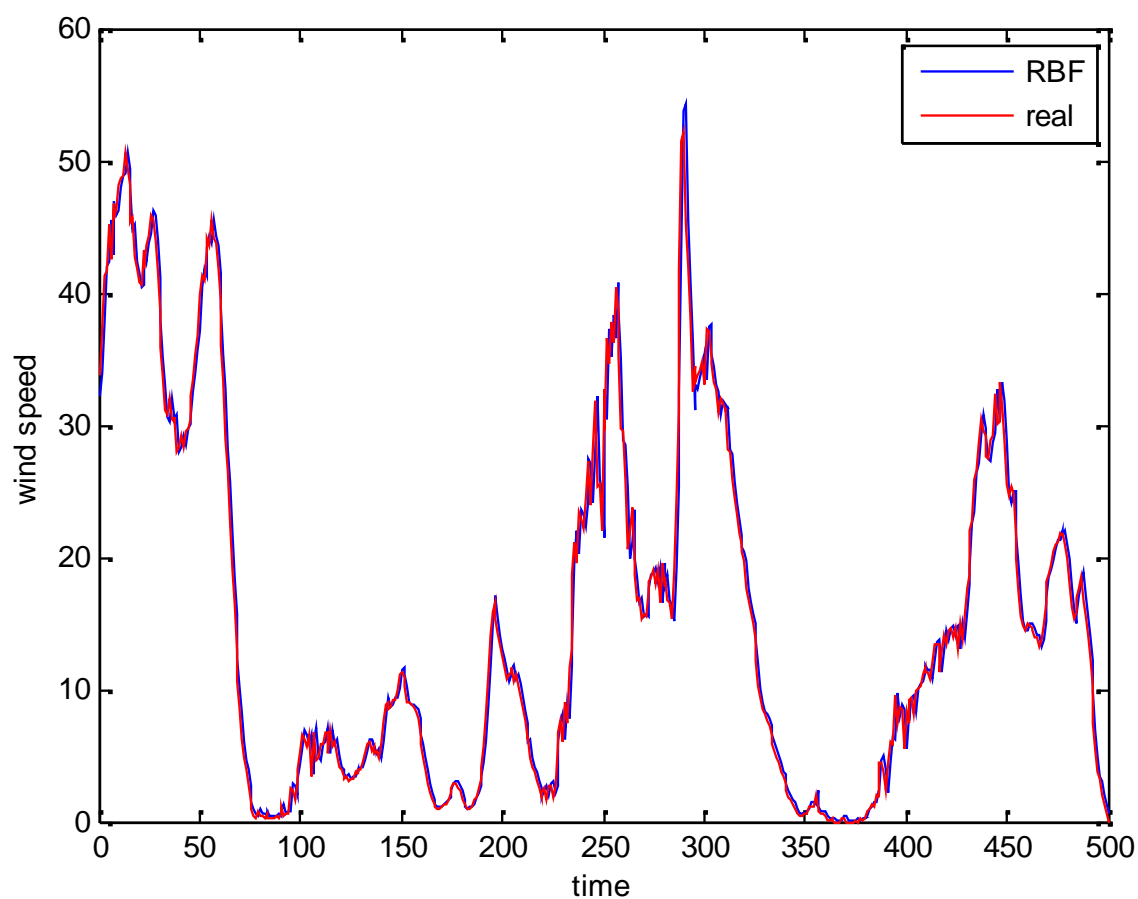


Figure 3.17 : Courbe comparaison valeurs réels et prédites de RBF pour 10 neurones pour le site Connecticut.

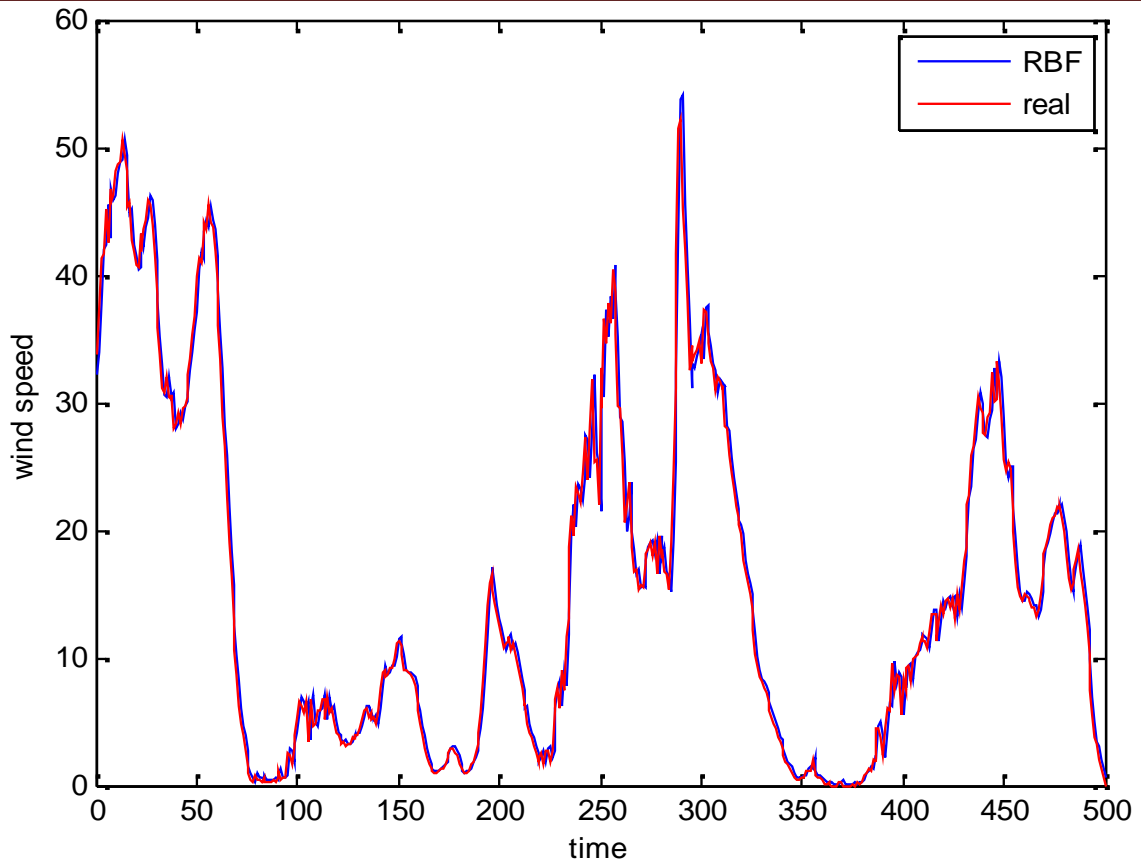


Figure 3.18 : Courbe comparaison valeurs réels et prédites de RBF pour 20 neurones pour le site Connecticut.

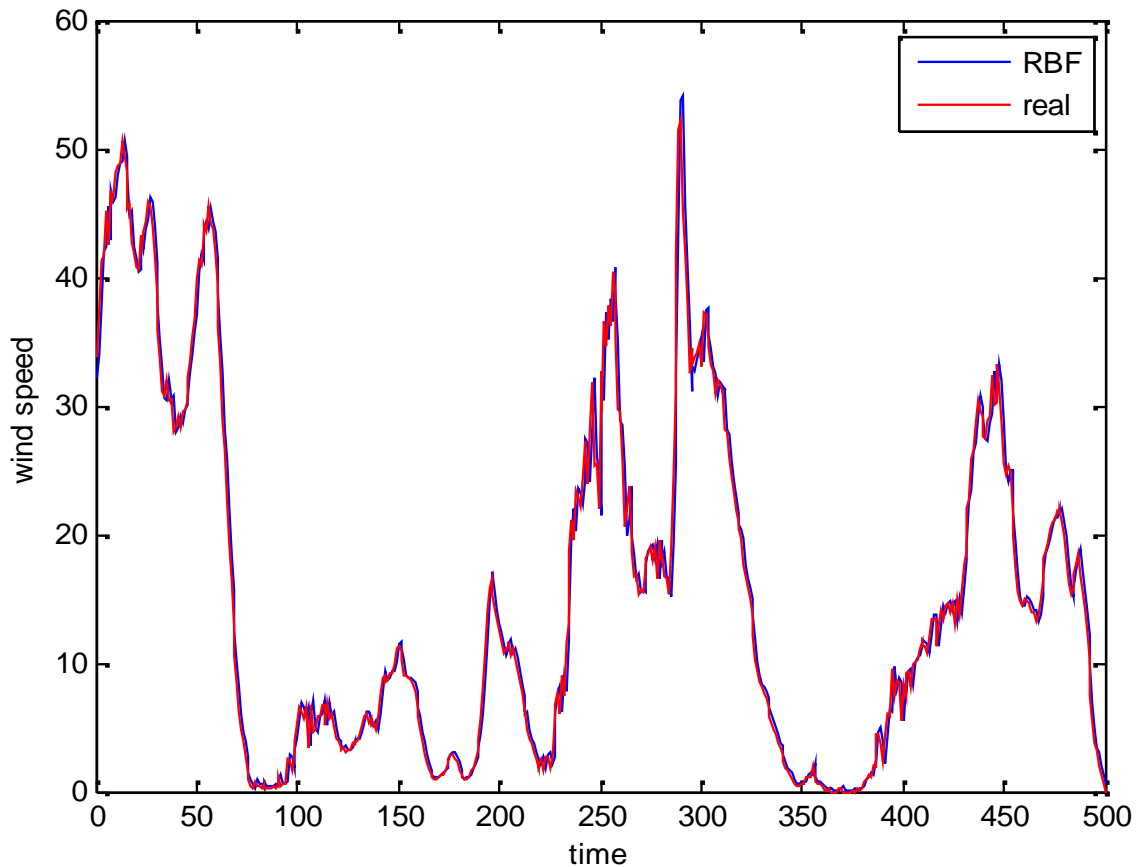


Figure 3.19 : Courbe comparaison valeurs réels et prédites de RBF pour 30 neurones pour le site Connecticut.

CONCLUSION

Dans ce chapitre, nous avons tenter 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 avec LDA et aussi deux types de réseaux de neurones MLP 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.

Dans la première partie, nous avons utilisé la technique LDA avec la méthode MLP pour l'obtention des valeurs des critères d'erreurs « RMSE , MAPE,MABE et R » sur la base complète avec une comparaison entre le courbe réel et prédite pour les deux sites étudiés.

En deuxième partie, on utilise la technique LDA avec la méthode RBF pour la base des données complète et avec les critères d'erreurs RMSE et « MAPE,MABE et R » et on fait une comparaison entre la courbe réelle et la courbe prédite pour les deux sites étudiés.

CONCLUSION GENERALE

CONCLUSION GENERALE

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 technique de LDA à deux modèles neuronal en « MATLAB », à savoir : MLP (Perceptron Multi couche) et RBF pour prédire la vitesse du vent pour deux sites.

Nous avons montré lors de cette étude, les performances de la méthode de prédiction par les réseaux de neurones en utilisant les réseaux PMC et RBF pour prédire la vitesse du vent.

Les résultats obtenus par l'ensemble des modèles sont satisfaisants ; les taux d'apprentissage fournis par le PMC sont les plus meilleurs mais ses performances dépendent de l'initialisation des poids. Les taux de test du réseau PMC sont les plus meilleurs mais que ceux des autres réseau neuronal RBF, d'où le réseau PMC a bien généralisé les nouvelles données. La phase d'apprentissage du réseau PMC est plus rapide que celle des réseaux neuronaux à fonction de base radiale et il converge au un nombre réduit d'itérations par rapport RBF.

D'une manière générale, l'utilisation de ces méthodes nous permet d'obtenir des valeurs futures bien précises concernant la vitesse du vent et par conséquent 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

References bibliographies

- [1] Site web : <http://www.planetoscope.com>
- [2] M. Said and A. Ibrahim, «Energy in the Arab World’, Energy» Vol. 9, N°3, pp. 217 - 238, 1984.
- [3] H. Bensaid, « The Algerian Programme on Wind Energy ». Proceeding of WEAC, pp. 21-27, Oxford, 1985.
- [4] R. Hammouche , « Atlas Vent de l’Algérie/ONM » Office des Publications Universitaires (OPU), Alger 1990.
- [5] N. Kasbadji merzouk , « An Evaluation of Wind Energy Potential in Algeria ». Proceeding of EWEC'94 congress, Thessaloniki, Grèce, 1994.
- [6] F. Youcef Ettoumi, « Ressources Energétiques Renouvelables en Algérie », Thèse de doctorat d’état, USTHB, 2002
- [7] F. Poitiers «Etude 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.
- [8] J MARTIN, « Energies éoliennes », Techniques de l’Ingénieur, Traités Energétique B8585, 1-, 22p,1997
- [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] 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.
- [11] site de constructeur d’éoliennes : <http://www.enercon.de/>;
- [12] T Grignoux, R Gibert, P Neau, and C Buthion “Eoliennes en milieu urbain -etat de l’art,” Agence Régionale de l’Environnement et des Nouvelles Energies Ile-de-France ARENE 2006., pp. 1-129
- [13] N. Cherfia , «Conversion d'énergie produite par des générateurs éoliens»,thèse de magister , Université Mentouri Constantine,2010
- [14] JM.Benmedjahed «Choix du site et optimisation du dimensionnement d'une installation éolienne dans le nord Algérien et son impact sur l'environnement»,thèse de doctorat, Universite abou-bakr belkaid tlemcen,2013.
- [15] <http://thesis.univbiskra.dz/Méthodes de Réduction et Classification>.
- [16] <https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-l-des-multi>
-

References bibliographies

- [17] Factorisation matricielle non négative sous contraintes. Application 'a l'identification de sources industrielles. Gilles Delmaire¹ , Gilles Roussel¹ , Dany Hleis² , Dominique Courcot²
¹Laboratoire d'Informatique Signal Image de la C^ote d'Opale, EA 4491 – Université du Littoral C^ote d'Opale Maison de la Recherche Blaise Pascal 50, rue Ferdinand Buisson, 62228 Calais Cedex, France. gilles.delmaire@lisic.univ-littoral.fr, gilles.roussel@lisic.univ-littoral.fr
- [18] Analyse en composantes indépendantes Jean-François Cardoso CNRS/ENST Paris, France cardoso@tsi.enst.fr <http://tsi.enst.fr/~cardoso>
- [19] Reconnaissance de visages par Analyse Discriminante Linéaire(LDA) par Amira Chiheb Université du 8 mai 45 de Guelma Algérie - Licence en informatique 2003
- [20] Norbet Wiener. Cybernetics; or, control and communication in the animal and the machine. Wiley, Published by the M.I.T.press cambridge,massachusetts. 1948
- [21] G-O Cimuca « système inertiel de stockage d'énergie associé à des générateurs éoliens » thèse de doctorat ,Ecole Nationale Supérieure d'Arts et Métiers Centre de Lille,2005
- [22] Franck Rosenblatt. The Perceptron : probabilistic model for information storage and organization in the brain. Psychological Review, 65:386-408, 1958
- [23] Donald O. Hebb. The organisation of behavior. Wiley, 1949.
- [26] J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79:2554-2558, 1982
- [24] 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
- [25] Marc Parizeau réseaux de neurones gif-21140 et gif-6432 par 2004 université laval –
- [26]Borgi, A., Apprentissage supervisé par génération de règles:le système SUCRAGE, Thèse de l'université de Paris 6 ,1999
- [30] Salima Nebti ; <<Thème Reconnaissance de Caractères Manuscrits par Intelligence Collective>> ; thèse Présenté à la Faculté des Sciences Département d'Informatique sétif.

Résumé

Le développement et exploitation des énergies renouvelables ont connu une forte augmentation ces dernières années .D'ici 20-30 ans , tout système énergétique 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 sûreté 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 **LDA** appliqué à un modèle intelligent à savoir: **Réseau de Neurone Artificiel** Pour prédire la vitesse du vent.

Mots clé: Vitesse du Vent, LDA, Réseau de Neurone Artificiel .

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 **LDA** technique applied to a neural model **Artificial Neural Network** to predict the wind speed with **MATLAB** software.

Keywords: LDA;wind energy ; Neural Networks Artificial .

لقد عرفت الطاقات المتجددة تطورا ملحوظا بعد عديد الاكتشافات في السنوات الاخيرة اذ يقدر انه في 20-30 الطاقة المستدامة ستعتمد على الاستخدام المنطقي للطاقة الاحفورية وعودة الطاقات المتجددة. المثير للاهتمام في الطاقات المتجددة هي اللامركزية فالاهتمام منصب على التشغيل في مكان الاستهلاك عن طريق التحويل المباشر إلى طاقة حرارية أو كهربائية حسب الحاجة. إن إنتاج الكهرباء اللامركزية عن طريق مصادر الطاقة المستدامة يوفر امن اكبر للمستهلكين و ذلك باحترام البيئة ولكن الطبيعة العشوائية للمصادر يفرض علينا تأسيس قواعد لاستعمال والتحكم بهذه الطاقة بشكل افضل لذا فالهدف من هذه المذكرة هو تطوير نموذج حسابي يعتمد على تقنية تحليل التمييز الخطي .
الكلمات الرئيسية تحليل التمييز الخطي , طاقة الرياح , الشبكات العصبية .
للتنبؤ بسرعة الرياح ببرنامج ما .
ية.