

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN GENIE CONCENTRATION ÉNERGIES RENOUVELABLES ET
EFFICACITÉ ÉNERGÉTIQUE
M. Ing.

PAR
Joseph Ismael TOUANI TCHANKO

DÉVELOPPEMENT D'UNE CAPACITÉ DE PRÉVISION DE LA PRODUCTION
ÉOLIENNE

MONTRÉAL, LE 21 SEPTEMBRE 2011

©Tous droits réservés, Joseph Ismael Touani Tchanko, 2011

©Tous droits réservés

Cette licence signifie qu'il est interdit de reproduire, d'enregistrer ou de diffuser en tout ou en partie, le présent document. Le lecteur qui désire imprimer ou conserver sur un autre media une partie importante de ce document, doit obligatoirement en demander l'autorisation à l'auteur.

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

Dr. Christian Masson, directeur de mémoire
Département du génie mécanique à l'École de technologie supérieure

Dr. Yves Gagnon, codirecteur de mémoire
Faculté des études supérieures et de la recherche à l'Université de Moncton

Dr. Wei Yu, collaborateur au projet
Division de la recherche météorologique à Environnement Canada

Dr. Robert Benoit, président du jury
Département du génie mécanique à l'École de technologie supérieure

Dr. Julien Choisnard, examinateur externe
HYDRO-QUÉBEC

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 14 SEPTEMBRE 2011

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Cette étude a été réalisée grâce aux données fournies par PEI Energy Corporation et Environnement Canada, deux institutions pour lesquelles je suis entièrement reconnaissant. Je remercie le Réseau stratégique du CRSNG sur l'énergie éolienne (WESNet), qui a initié et financé ce projet, et m'a permis d'échanger avec plusieurs étudiants et chercheurs en éolien de partout au Canada à travers des rencontres scientifiques.

Je remercie mon directeur de recherche, le professeur Christian Masson, qui m'a accueilli au sein de son groupe de recherche, la chaire de recherche du Canada en aérodynamique des éoliennes en milieu nordique, et m'a donné l'opportunité de travailler sur ce projet. Il a su être disponible chaque fois que j'ai eu besoin de le rencontrer, et m'a fait profiter de son expérience par ses commentaires éclairants. Je tiens à remercier mon codirecteur, le professeur Yves Gagnon, qui a toujours été disponible et m'a offert de collaborer avec les membres de la chaire K.C. Irving en développement durable de l'Université de Moncton. Son grand intérêt pour le détail et son talent de motivateur m'ont grandement assisté tout au long de ce projet. Je suis également reconnaissant envers Wei Yu de la division de la recherche météorologique à Environnement Canada, collaborateur au projet, qui a su m'encourager et a mis à contribution son expérience en prévision météorologique et sa rigueur scientifique, en portant des remarques et des critiques qui ont favorisé le bon déroulement de ce projet. Je remercie grandement les membres du jury qui ont accepté de revoir ce travail et d'apporter des suggestions afin de l'améliorer.

Mes sincères remerciements s'adressent à tous mes collègues de la chaire de recherche du Canada en aérodynamique des éoliennes en milieu nordique avec qui j'ai eu des échanges constructifs et partagé des moments inoubliables.

Je tiens enfin à remercier ma famille pour son soutien et son affection inconditionnels qui m'aident à rester déterminé à l'atteinte de mes objectifs. Je n'oublierai pas tous mes amis qui ont toujours su être présents. Merci à vous tous...

DÉVELOPPEMENT D'UNE CAPACITÉ DE PRÉVISION DE LA PRODUCTION ÉOLIENNE

Joseph Ismael TOUANI TCHANKO

RÉSUMÉ

L'énergie éolienne a connu une croissance importante au cours de la dernière décennie dans beaucoup de régions dans le monde. Elle constitue par conséquent la source d'énergie renouvelable qui présente le plus grand taux d'intégration au sein des réseaux électriques. Pour soutenir son évolution et augmenter sa valeur économique, l'industrie éolienne et la recherche doivent être capables de trouver les solutions des problèmes qui peuvent freiner son évolution et qui sont liés à la nature de cette source d'énergie. L'un des principaux défis étant de trouver des solutions face à l'intermittence du vent qui rend sa production variable et incontrôlable. Pour ce faire, le développement d'outils capables de prédire la puissance produite des centrales éoliennes à court terme est une solution qui aidera les opérateurs des réseaux électriques à intégrer davantage l'énergie éolienne dans leur portefeuille énergétique tout en maintenant le réseau électrique équilibré en tout temps.

Le développement d'un modèle de prévision de la production éolienne en utilisant une approche physique se fait en deux étapes. La première étape consiste à prédire la vitesse du vent et d'autres variables météorologiques (température, pression, humidité, etc.) au niveau de la centrale éolienne et la seconde consiste à utiliser un modèle de puissance pour convertir ces prévisions du vent en prévisions de la puissance. Cette étude se concentre sur la deuxième dont le but est de proposer une méthodologie permettant d'évaluer la courbe de puissance d'une centrale éolienne et par la suite, d'utiliser cette courbe de puissance pour convertir des prévisions du vent en prévisions de la puissance produite d'une centrale éolienne.

Le développement de la courbe de puissance est effectué en menant une série d'expériences qui ont permis de converger vers une méthodologie offrant une bonne précision, en utilisant les données mesurées du vent et de la puissance produite d'une centrale éolienne. En

comparant la méthode des Bins avec les réseaux de neurones artificiels (RNA), les résultats ont montré que les RNA procurent les meilleurs résultats, principalement à cause de leur facilité à prendre en compte plusieurs variables, même avec un nombre limité de données mesurées. Une expérience similaire est effectuée en comparant deux différentes approches tout en utilisant la même méthode. La première consiste à modéliser la courbe de puissance de la centrale éolienne en utilisant la vitesse du vent lue à un mât de mesures; tandis que la deuxième utilise la moyenne des vitesses du vent lues aux éoliennes. Les réseaux de neurones sont exploités pour tester l'influence de plusieurs paramètres météorologiques sur la courbe de puissance de la centrale éolienne. La courbe de puissance de la centrale éolienne est donc proposée en utilisant la meilleure méthode identifiée, l'approche offrant le meilleur résultat et les paramètres s'étant avérés plus pertinents. La courbe de puissance de la centrale éolienne est par la suite utilisée pour convertir les prévisions météorologiques calculées par GEM-LAM 2.5km en prévisions de la puissance produite. L'évaluation du modèle, par le calcul des paramètres statistiques, a permis d'enregistrer une erreur absolue moyenne normalisée qui varie entre 9.6 et 14.1 % de la capacité totale de la centrale sur les 48 horizons de prévision disponibles. Cette étude a été menée en utilisant les données mesurées de deux centrales éoliennes situées au Canada.

Mots-Clés : Courbe de puissance, prévisions météorologiques, prévisions de la puissance éolienne, réseaux de neurones, méthode des Bins.

DÉVELOPPEMENT D'UNE CAPACITÉ DE PRÉVISION DE LA PRODUCTION ÉOLIENNE

Joseph Ismael TOUANI TCHANKO

ABSTRACT

During the last decade, the total installed capacity of wind energy has grown all over the world. In most countries, governments and industry actors of wind energy plan to sustain this growth in order to be able to reach at least 20 percent of the total generated electricity provided by wind energy in 2030. To achieve that goal, it will be necessary to face the problems related to the nature of that source of energy, especially the problem of the wind intermittency. In fact, wind power output depends on weather conditions that cannot be directly controlled by humans. The wind intermittency induces fluctuations of the power outputs of wind farms and this is a special challenge for utility systems operations. To face that problem, one solution consists in developing wind power forecasting systems that will make wind power output more predictable and then, will facilitate its integration toward electricity grids.

Wind farm power output forecasting systems using physical approach are generally made in two steps. The first one consists in predicting the wind (wind speed, temperature, pressure, humidity, etc.) at the level of the wind farm and the second consists in translating the predicted wind into the predicted power output by using a wind to power model. This study is focused on the second step with the main objective to propose a methodology able to calculate the wind farm power curve with the best accuracy possible. This power curve will then be used to translate predicted winds into predicted power outputs of the wind farm.

In this study, several experiments have been conducted to identify an accurate way to model the wind farm power curve using historical wind farm power production and wind measurements. Comparing the Bin Method and Artificial Neural Networks (ANN), results show that ANN perform better, mainly due to the fact that they can easily add more variables, even with a limited quantity of data. Also, the influence of several parameters has

been tested using ANN. Two additional methods for the wind farm power curve modeling have been examined. The first method consists in modeling the wind farm power curve by comparing the wind speed at a reference wind mast to the total power output; while the second method consists in modeling the wind farm power curve by comparing the average wind speed at the hub height of each wind turbine to the total power output. The wind farm power curve developed is then used to predict power outputs using GEM-LAM 2.5 km predicted meteorological variables as inputs. The methods were applied to two operating wind farms in Canada. The analysis of the model forecasts revealed a normalized mean absolute error in the ranges of 9.6 to 14.1% for a 48-hour forecast horizon.

Keywords: Power curve, wind forecast, wind power forecast, artificial neural networks, Bins method.

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 MISE EN CONTEXTE.....	5
1.1 Concepts de base.....	5
1.2 Les familles de modèles.....	8
1.3 La performance des modèles	10
1.4 Avantages de la prévision régionale	13
1.5 Quelques exemples de modèle de prévision de la production éolienne.....	15
CHAPITRE 2 CONTRÔLE DE LA QUALITÉ DES DONNÉES ET DESCRIPTION DES VARIABLES IMPORTANTES.....	21
2.1 Les données mesurées.....	22
2.2 La procédure de contrôle de la qualité des données	23
2.3 Bilan du contrôle de la qualité des données.....	26
2.4 Description des variables importantes	34
CHAPITRE 3 MODÉLISATION DE LA COURBE DE PUISSANCE	37
3.1 L'impact du sillage sur la courbe de puissance de la centrale éolienne.....	40
3.1.1 Procédure d'évaluation des secteurs obstrués par le sillage	41
3.1.2 Secteurs obstrués du mât W2 et impact sur la courbe de puissance	42
3.2 Comparaison de deux différentes méthodes	44
3.2.1 La méthode des Bins.....	46
3.2.2 Les réseaux de neurones	50
3.2.2.1 Les entrées du réseau	52
3.2.2.2 Les sorties du réseau de neurones.....	52
3.2.2.3 Le choix des fonctions d'activation.....	52
3.2.2.4 Le nombre de couches cachées et le nombre de neurones par couche cachée	53
3.2.2.5 La technique d'apprentissage.....	55
3.2.2.6 Les résultats obtenus avec les réseaux de neurones.....	55
3.2.3 Comparaison de la méthode des Bins et des réseaux de neurones	57
3.3 Comparaison de deux différentes méthodologies.....	59
3.4 Choix des paramètres.....	62
3.5 Courbe de puissance proposée et analyse de l'incertitude.....	66
CHAPITRE 4 DÉVELOPPEMENT DU MODÈLE DE PRÉVISION.....	69
4.1 Description et architecture du modèle	69
4.2 Les arrêts des turbines.....	71
4.3 Prévision des variables météorologiques.....	75
4.3.1 Prévision de la moyenne des vitesses aux éoliennes	76
4.3.2 Prévision de la masse volumique de l'air	85

4.3.3	Prévision de la direction du vent.....	88
4.4	La conversion des prévisions météorologiques en prévisions de puissance.....	89
CHAPITRE 5 ÉVALUATION DU MODÈLE		92
5.1	Introduction.....	92
5.2	Procédure d'évaluation du modèle.....	95
5.3	Application de la procédure d'évaluation du modèle	97
5.4	Quelques résultats de la littérature.....	103
CONCLUSION.....		109
ANNEXE I	LISTE DE QUELQUES MODÈLES GLOBAUX DE PRÉVISIONS MÉTÉOROLOGIQUES OPÉRATIONNELS DANS LE MONDE	115
ANNEXE II	ÉVALUATION DE LA COURBE DE PUISSANCE DE LA CENTRALE ÉOLIENNE D'EAST POINT (IPE)	117
BIBLIOGRAPHIE.....		121

LISTE DES TABLEAUX

	Page
Tableau 2.1	Description du test d'intervalle25
Tableau 2.2	Description du test de relation25
Tableau 2.3	Description du test de tendance26
Tableau 2.4	Bilan des données suspectes27
Tableau 2.5	Bilan des données supprimées31
Tableau 3.1	Performances de la courbe de puissance (Méthode des Bins)49
Tableau 3.2	Erreurs du modèle (RNA).....56
Tableau 3.3	Comparaison des performances de deux méthodes : Bins vs RNA.....58
Tableau 3.4	Comparaison des performances de deux méthodologies : vitesse au mât vs moyenne des vitesses aux éoliennes61
Tableau 3.5	Performances de la courbe de puissance proposée68
Tableau 4.1	Variables météorologiques prédites par GEM-LAM et utilisées pour prédire la puissance de la centrale (niveau de hauteur : I=50m, II=120m)76
Tableau 5.1	Moyenne des erreurs du modèle proposé et de Persistance pour les horizons de prévision de une à quarante-huit heures102
Tableau 5.2	Moyenne des erreurs du modèle proposé et du NMR pour les horizons de prévision de une à quarante-huit heures103

LISTE DES FIGURES

		Page
Figure 1.1	Les différentes familles de modèles Tirée de Giebel <i>et al.</i> (2011).....	8
Figure 2.1	Procédure de contrôle de la qualité des données	23
Figure 2.2	Exemple 1 de données suspectes	28
Figure 2.3	Exemple 2 de données suspectes	29
Figure 2.4	Exemple 3 de données suspectes	30
Figure 2.5	Caractéristiques de vent du site.....	32
Figure 2.6	Courbe de puissance de la centrale avec arrêts des turbines.....	33
Figure 2.7	Courbe de puissance de la centrale sans arrêts des turbines	33
Figure 3.1	Exemple de courbe de puissance d'une éolienne	37
Figure 3.2	Comparaison entre la courbe de puissance d'une éolienne et la courbe de puissance d'une centrale.....	38
Figure 3.3	Position des turbines voisines au mât W2.....	40
Figure 3.4	Limite du secteur obstrué par le sillage en fonction de la distance relative.....	42
Figure 3.5	Secteurs obstrués par les turbines voisines au mât W2.....	43
Figure 3.6	Effets de sillage sur la courbe de puissance de la centrale éolienne	43
Figure 3.7	Données utilisées pour modéliser et tester la courbe de puissance de la centrale. L'ensemble des données sélectionnées a été divisé en deux groupes afin de s'assurer que les données utilisées pour modéliser la courbe de puissance soient différentes de celles utilisées pour la tester.....	46
Figure 3.8	Courbe de puissance calculée par la méthode des Bins	48

XVIII

Figure 3.9	Comparaison entre la puissance mesurée et la puissance modélisée avec la méthode des Bins (courbe de puissance et droite de régression).....	49
Figure 3.10	Exemple de PMC à 3 couches cachées.....	51
Figure 3.11	Le neurone artificiel.....	51
Figure 3.12	Performance du PMC en fonction du nombre de couches cachées et en fonction du nombre de neurones sur la couche cachée	54
Figure 3.13	Architecture du réseau de neurones implanté	56
Figure 3.14	Comparaison entre la puissance mesurée et la puissance modélisée avec les RNA (courbe de puissance et droite de régression)	56
Figure 3.15	Comparaison des résultats (méthode des Bins vs RNA)	57
Figure 3.16	Courbe de puissance de la centrale avec la moyenne des vitesses aux éoliennes (données modélisation et données test)	59
Figure 3.17	Comparaison des courbes de puissance de la centrale (vitesse lue au mât vs moyenne des vitesses aux éoliennes)	60
Figure 3.18	Comparaison des résultats de deux différentes méthodologies	61
Figure 3.19	Performances du PMC avec deux variables (vitesse du vent et masse volumique de l'air).....	64
Figure 3.20	Performances du PMC en fonction de plusieurs variables (WS=vitesse du vent; AirD=masse volumique de l'air; WD=direction du vent; Sh=cisaillement; Ti=intensité de turbulence)	65
Figure 3.21	Comparaison des performances du PMC en fonction de plusieurs variables météorologiques.....	66
Figure 3.22	Comparaison entre la puissance mesurée et la puissance modélisée par la courbe de puissance proposée.....	67
Figure 4.1	Les différentes étapes du modèle de prévision	69
Figure 4.2	Fréquence des arrêts des turbines sur une année complète.....	72
Figure 4.3	Comparaison entre les arrêts des turbines et la vitesse du vent	73
Figure 4.4	Courbe de puissance de la centrale sans arrêts des turbines	74

Figure 4.5 Courbe de puissance de la centrale incluant les arrêts des turbines.....75

Figure 4.6 Résultats de l'interpolation verticale de la vitesse du vent.....78

Figure 4.7 Comparaison de la position des points de la grille de prévision avec la position des éoliennes. Les points verts (dispersés) représentent la position des éoliennes et le point encerclé représente la position du mât79

Figure 4.8 Comparaison entre la moyenne des vitesses du vent prédites par GEM-LAM et la moyenne des vitesses du vent aux éoliennes80

Figure 4.9 Comparaison entre la moyenne des vitesses du vent prédites et la moyenne des vitesses aux éoliennes82

Figure 4.10 Comparaison entre les prévisions de la vitesse du vent et les mesures83

Figure 4.11 Comparaison entre les prévisions de la vitesse du vent au mât et les vitesses du vent mesurées. Les prévisions sont corrigées par une régression linéaire par secteur de direction de vent Tirée de Bédard (2010).....83

Figure 4.12 Comparaison entre les prévisions de la vitesse du vent au mât et les vitesses du vent mesurées. Les prévisions ne sont pas corrigées.....84

Figure 4.13 Comparaison des prévisions des variables météorologiques avec les valeurs mesurées: a- Température b- Pression atmosphérique.....87

Figure 4.14 Comparaison des prévisions de la direction du vent avec les valeurs mesurées.....88

Figure 4.15 Comparaison des séries temporelles des prévisions de la puissance avec les valeurs mesurées.....90

Figure 4.16 Performances du modèle de prévision en fonction de l'horizon de prévision.....90

Figure 5.1 Distribution de l'erreur de la prévision de puissance normalisée pour l'ensemble des horizons de prévision allant de une à douze heures98

Figure 5.2 Erreur absolue moyenne normalisée (NMAE) du modèle en fonction de l'horizon de prévision.....99

Figure 5.3	Erreur quadratique moyenne normalisée (NRMSE) du modèle en fonction de l'horizon de prévision	99
Figure 5.4	Comparaison du modèle avec Persistance en fonction de l'horizon de prévision	101
Figure 5.5	Pourcentage d'amélioration du modèle proposé en comparaison à la Persistance en fonction de l'horizon de prévision.....	102
Figure 5.6	Comparaison du modèle proposé avec le NMR en fonction de l'horizon de prévision	102
Figure 5.7	Pourcentage d'amélioration du modèle proposé en comparaison au NMR en fonction de l'horizon de prévision	103
Figure 5.8	Distribution de l'erreur de la prévision de puissance d'un modèle pour les horizons de prévision d'une et 24 heures Tirée de Pinson <i>et al.</i> (2004).....	104
Figure 5.9	Analyse des performances du modèle Prediktor et comparaison avec les modèles de référence Tirée de Giebel <i>et al.</i> (2011)	105
Figure 5.10	Comparaison de cinq modèles différents avec Persistance Tirée de Pinson (2006).....	106

LISTE DES ABRÉVIATIONS

ANEMOS	A next generation wind resource forecasting system for the large-scale integration of onshore and offshore wind farms
ANN	Artificial neural networks
AWS	Associated Weather Services
CanWEA	Canadian Wind Energy Association
CENER	Centre national espagnol des énergies renouvelables
CFD	Computational fluid dynamics
CIEMAT	Centre de recherche pour l'énergie, l'environnement et la technologie
CRSNG	Conseil de recherches en sciences naturelles et en génie du Canada
DISP	Dispersion
DTU	Université technique du Danemark
EPRI	Electric Power Research Institute
GEM	Global environment multiscale model
HIRLAM	High resolution limited area model
IEA	International Energy Agency
IEC	International Electrotechnical Commission
LAM	Limited area model
MAE	Mean absolute error
MOS	Model output statistics
NMAE	Normalized mean absolute error
NMR	Nouveau modèle de référence
NRMSE	Normalized root mean square error

NWP	Numerical Weather prediction
PMC	Perceptron multi-couches
RMSE	Root mean square error
RNA	Réseau de neurones artificiels
SCADA	Supervisory control and data acquisition
SDBias	Difference between standard deviations
USA	United States of America
WAsP	Wind Atlas Analysis and Application Program
WESNet	Wind Energy Strategic Network
WPPT	Wind Power Prediction Tool

INTRODUCTION

Contexte

Au cours de la dernière décennie, l'énergie éolienne a connu une croissance spectaculaire de la puissance totale installée dans le monde, partant de 18GW à la fin de l'année 2000 à 195GW à la fin de 2010 (IEA, 2011). Cette croissance se justifie par l'augmentation du prix du pétrole, par les actions mises en place par les gouvernements pour lutter contre les changements climatiques (Abreu, 2008) tout en garantissant les approvisionnements énergétiques, et par la réduction considérable des coûts de production d'énergie d'origine éolienne. D'après l'Association canadienne de l'énergie éolienne (CanWEA), le Québec devrait se donner comme objectif d'ajouter, après 2015, un bloc additionnel de 8000MW de sorte que l'éolien puisse contribuer environ 20 pour cent de la consommation totale d'énergie de la province d'ici 2025 (CanWEA, 2010). Dans un contexte de marché énergétique libéralisé, l'énergie éolienne doit être économiquement compétitive et techniquement fiable pour pouvoir atteindre les objectifs fixés. Il est donc important de comprendre les problèmes reliés à cette nouvelle source d'énergie et d'apporter les solutions pour faciliter davantage son insertion dans les réseaux électriques, et rendre sa valeur économique plus concurrentielle face aux autres sources d'énergie conventionnelles dont la technologie est maîtrisée.

Problématique

L'un des principaux éléments qui contribuent à ralentir l'intégration de l'énergie éolienne au sein des réseaux de distribution d'énergie est la variabilité de la puissance produite. En effet, la production d'énergie d'une centrale éolienne dépend des conditions météorologiques qui varient de manière intermittente. Il est donc impossible de contrôler la puissance produite des centrales éoliennes. La charge par contre, qui représente la demande en énergie des consommateurs, varie en général suivant un profil régulier et prévisible avec une très bonne précision (Bunn, 2000; Bunn et *al.*, 1985). Les opérateurs du réseau électrique qui doivent répondre à la demande en énergie de cette charge ont besoin de faire des prévisions de distribution de l'énergie électrique plusieurs jours à l'avance pour maintenir le réseau

électrique équilibré en tout temps. Le recours aux sources d'énergie intermittentes rend donc les opérations énergétiques plus difficiles à réaliser. Dans ce contexte, développer les outils capable de prédire la puissance de sortie des centrales éoliennes plusieurs jours à l'avance se présente comme une solution capitale, qui permettra à cette source d'énergie de s'intégrer plus facilement dans les réseaux de distribution d'énergie. Cette solution pourra alors aussi contribuer à hausser sa valeur économique qui est affectée par sa variabilité.

But de la recherche

Cette étude a pour principal but de proposer une méthodologie permettant de prédire la production des centrales éoliennes plusieurs heures à l'avance. Dans une première approche, le modèle sera développé pour les horizons de prévision allant d'une heure à 48 heures. L'atteinte de cet objectif se fera en deux étapes. La première consistera à proposer une méthodologie capable d'évaluer la courbe de puissance de la centrale éolienne. La deuxième étape consistera à utiliser cette courbe de puissance proposée pour convertir les prévisions météorologiques en prévisions de puissance produite d'une centrale éolienne. La courbe de puissance de la centrale éolienne sera développée en supposant qu'aucune éolienne n'est arrêtée. Il sera proposé une approche qui pourra tenir compte des arrêts des éoliennes dans le modèle de prévision de la puissance de la centrale éolienne, lorsque ces arrêts seront prévisibles avec une bonne précision.

Approche méthodique

Le développement de la courbe de puissance de la centrale éolienne se fait en analysant plusieurs méthodes, en comparant des méthodologies et en sélectionnant les paramètres météorologiques qui influencent la puissance produite des centrales éoliennes. Plusieurs similitudes sont observées entre la courbe de puissance des éoliennes individuelles et celle des centrales éoliennes au complet. Sur la base des études retrouvées dans la littérature, différentes méthodes peuvent être implantées et comparées afin de saisir l'impact du choix de la méthode sur la précision de la modélisation. Une comparaison entre différentes approches est aussi effectuée, en utilisant la même méthode, ce qui permet de comprendre l'influence du choix de la méthodologie pour définir la courbe de puissance de la centrale éolienne. En

choisissant des variables qui influencent la puissance produite des éoliennes individuelles selon des études précédentes, des expériences peuvent être effectuées afin d'évaluer la contribution de ces différentes variables sur la modélisation de la courbe de puissance de la centrale éolienne. L'ensemble de ces expériences pourra alors conduire à proposer une courbe de puissance de la centrale éolienne en choisissant la meilleure méthode, en adoptant une bonne approche et en sélectionnant les variables les plus influentes selon les expériences effectuées. Environnement Canada fournit les prévisions météorologiques à l'aide du modèle de prévision numérique appelé GEM-LAM. L'ensemble du domaine de prévision couvre l'Est et l'Ouest du Canada. Bédard et *al.* (2011) ont proposé un module de sortie statistique (MOS), qui permet d'améliorer la précision des sorties de GEM-LAM en prenant en compte l'influence de la topographie sur la précision du modèle. Le MOS développé par Bédard et *al.* (2011) utilise les sorties de GEM-LAM à 2.5 km de résolution pour prédire la vitesse du vent à un point précis de la centrale éolienne, avec la meilleure précision possible. Les prévisions météorologiques sur la centrale éolienne seront obtenues en utilisant les sorties de GEM-LAM et le MOS développé par Bédard et *al.* (2011); et la courbe de puissance servira à convertir ces prévisions météorologiques en prévisions de la puissance de la centrale.

La structure du mémoire

Une revue de la littérature est faite au premier chapitre. Elle permet de comprendre de manière générale le fonctionnement des modèles de prévisions de la production éolienne. La revue de la littérature effectuée donne aussi une connaissance des modèles qui ont déjà été développés, ainsi que de la précision de ceux-ci. Le développement du module de puissance, qui est l'objectif principal de cette étude, étant effectué à l'aide des données mesurées de vent et de production, le chapitre 2 décrit la centrale éolienne qui est utilisée pour l'étude, ainsi que la procédure qui a été adoptée pour le contrôle de la qualité des données. Au chapitre 3, il est principalement évoqué le développement de la courbe de puissance de la centrale éolienne. Ce chapitre décrit en détail les expériences effectuées qui ont permis d'analyser les différentes méthodes trouvées dans la littérature, de les comparer et enfin de proposer une méthodologie pour construire la courbe de puissance de la centrale. Le chapitre 4 évoque l'approche utilisée pour coupler la courbe de puissance développée aux sorties du modèle de

prévisions météorologiques (GEM-LAM) afin d'obtenir les prévisions de puissance de la centrale éolienne. Au chapitre 5, les résultats obtenus sont discutés et comparés avec ceux rencontrés dans la littérature.

CHAPITRE 1

MISE EN CONTEXTE

Les recherches en prévision de la production éolienne ont débuté au cours des années 1985-1987, alors que l'un des premiers modèles capable de prédire la puissance produite d'une centrale éolienne en se servant des prévisions météorologiques a été mis sur pied en Californie (Landberg *et al.*, 2003). Depuis lors, les modèles n'ont pas cessé de se développer tout en devenant de plus en plus performants au fil du temps. Ce chapitre a pour but de présenter de manière générale les concepts utilisés en prévision de la production éolienne. Il sera présenté, entre autres, les familles des modèles de prévision, les composantes des modèles, les méthodes d'évaluation des modèles, les différents modèles de références utilisés et les avantages de la prévision régionale. Le chapitre sera conclu en décrivant quelques modèles opérationnels de prévision de la production éolienne tout en évoquant leurs performances.

1.1 Concepts de base

Les modèles de prévision de la production éolienne sont des outils qui permettent de prédire à un temps t la puissance qui sera produite par une centrale éolienne à un temps $t + k$, où k se définit comme l'horizon de prévision. La plupart des modèles de prévision de la production éolienne sont composés des modules suivants : les modèles numériques de prévisions météorologiques (NWP); un système d'acquisition des données mesurées de la centrale éolienne; et un modèle de puissance qui convertit les prévisions météorologiques en prévisions de puissance (Landberg *et al.*, 2003). Cependant, la prévision étant la projection dans le futur des événements observés dans le présent et basée sur des suppositions, prédire un phénomène est toujours associé à un degré d'incertitude. Une formulation générique du problème est présentée par Pinson (2006) dans laquelle il note les variables de la manière suivante : la prévision de la puissance P_t produite par la centrale éolienne effectuée à un

temps t pour un horizon de temps $t + k$ se note $\hat{P}_{t+k/t}$. Cette notation sera utilisée tout au long de ce chapitre pour présenter les concepts utilisés en prévision éolienne.

Horizon de prévision

Dans le domaine de l'éolien, plusieurs horizons de prévision sont requis selon différentes applications (Giebel *et al.*, 2003). La puissance prédite pour des horizons allant des millisecondes à quelques minutes peut être utilisée pour le contrôle des éoliennes. En effet certaines conditions météorologiques peuvent s'avérer dommageables pour les éoliennes. La prévision à très court terme pourrait alors permettre à une éolienne de détecter ces conditions météorologiques, et déclencher un système automatique pour l'arrêter afin de s'en préserver. On peut aussi citer la prévision à court terme, pour des horizons allant de quelques heures à 72 heures. Ces horizons de prévision sont souhaités pour les opérations d'exploitation du réseau électrique et permettent de planifier les moyens de répondre à la demande d'énergie des consommateurs tout en équilibrant les réseaux électriques. Enfin, les prévisions de la production éolienne sont aussi souhaitées pour des horizons de temps plus long, pouvant aller jusqu'à 5 à 7 jours. Celles-ci peuvent être particulièrement importantes pour planifier les opérations d'entretien sur le réseau électrique qui nécessiteraient l'arrêt complet de la centrale éolienne. L'objectif de cette étude est de proposer une méthodologie capable de prédire la production éolienne qui facilitera son intégration dans les réseaux électriques. Par conséquent, c'est la prévision de la puissance de la centrale éolienne à court terme (l'horizon de prévision allant de quelques heures jusqu'à 72 heures) qui nous intéresse.

Les modèles numériques de prévisions météorologiques (NWP)

Les modèles numériques de prévisions météorologiques sont utilisés dans la plupart des modèles de prévision de la production éolienne à court terme. À partir des observations météorologiques mesurées de partout sur le globe, ils sont capables de calculer l'état futur de l'atmosphère à l'aide des lois de la physique qui gouvernent la météorologie. Les modèles globaux, qui évaluent l'état futur de l'atmosphère sur tout le globe, sont opérationnels dans à

peu près 15 services nationaux de météorologie. Ces modèles opérationnels fournissent les prévisions météorologiques sur des maillages pouvant aller de 25km à 80km de résolution et à différents niveaux de hauteur. Le tableau en annexe 1 présente une liste des modèles globaux de prévisions météorologiques opérationnels dans les services météorologiques de plusieurs pays, ainsi que leurs différentes caractéristiques. Les prévisions météorologiques sur un maillage ayant des dizaines de kilomètres de résolution ne peuvent pas être suffisantes pour définir l'état de l'atmosphère au niveau d'une centrale éolienne. Il est donc utilisé les modèles à aire limitée ou LAM (en anglais Limited Area Model) qui se servent des sorties des modèles globaux comme conditions initiales et conditions aux limites pour évaluer les prévisions météorologiques pour une région du globe avec une résolution beaucoup plus fine pouvant aller de 1km jusqu'à plusieurs kilomètres (Monteiro *et al.*, 2009). L'opération qui consiste à partir des prévisions évaluées par le modèle global au LAM est généralement appelée *downscaling* en prévision éolienne (Giebel *et al.*, 2003).

La courbe de puissance

La courbe de puissance d'une éolienne représente la relation qui existe entre les conditions météorologiques et la puissance produite par l'éolienne. Elle peut être modélisée pour une éolienne individuelle ou pour une centrale au complet. Dans certains modèles de prévision de la production éolienne, la courbe de puissance est utilisée pour convertir les prévisions météorologiques en prévisions de la production de la centrale ou de la turbine. Le chapitre 3 de cette étude est consacré à la modélisation de la courbe de puissance de la centrale éolienne. Dans ce chapitre, une revue de la littérature spécifique à la courbe de puissance, ainsi que les défis de la modélisation seront discutés en détail.

Les modules statistiques de traitement des sorties du modèle (MOS)

Les modules statistiques de traitement des sorties des modèles (MOS) sont souvent utilisés pour corriger les erreurs systématiques des modèles de prévision. Ils sont discutés en détail dans les travaux de Bédard *et al.* (2011) dont l'objet a consisté à implanter un MOS pour

améliorer les prévisions de la vitesse du vent calculées par GEM-LAM, qui est un modèle numérique de prévisions météorologiques utilisé par Environnement Canada. Le MOS est souvent appliqué sur les sorties du NWP (Monteiro *et al.*, 2009), pour améliorer les prévisions de la vitesse du vent au niveau de la centrale éolienne. Bédard *et al.* (2011) ont implanté un MOS en utilisant une méthode de régression linéaire et les réseaux de neurones. Les détails de la méthode et les différentes équations utilisées pour les implanter peuvent être consultés dans Bédard (2010).

1.2 Les familles de modèles

Les modèles de prévision de la production éolienne sont implantés suivant deux principales approches qui constituent les familles de modèles. On distingue ainsi les modèles physiques et les modèles statistiques. Les familles de modèles diffèrent par les étapes nécessaires pour calculer les prévisions de production de la centrale. La figure 1.1 présente de manière claire la différence entre les deux familles de modèles.

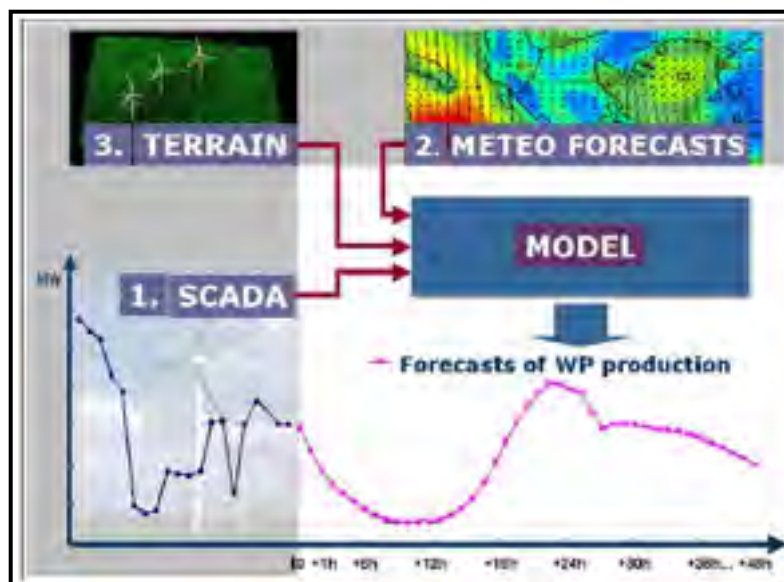


Figure 1.1 Les différentes familles de modèles
Tirée de Giebel *et al.* (2011)

Le SCADA fournit en ligne les données mesurées de la centrale éolienne.

- (1) : L'approche statistique utilisant les données mesurées.
- (2) : L'approche statistique utilisant les prévisions météorologiques.
- (2)+(3) : L'approche physique.
- (1)+(2) +(3) : L'approche combinée.

Les modèles physiques de la production sont décrits de manière détaillée par Lange *et al.* (2005). Ils utilisent une combinaison des informations données par deux modules: le NWP et les caractéristiques physiques de la centrale éolienne. Les sorties du modèle NWP sont extrapolées au niveau de la turbine ou de la centrale en utilisant un MOS. Pour cette étape, un modèle de CFD peut aussi être utilisé pour prendre en compte les influences de la topographie et de la rugosité sur le vent. La courbe de puissance est requise après cette étape pour passer des prévisions météorologiques aux prévisions de puissance. Pour ce faire, l'utilisation de la courbe de puissance certifiée du constructeur des éoliennes peut être une solution. Mais il est de plus en plus préféré de la modéliser. La famille de modèles statistiques quant à elle utilise moins d'information pour évaluer les prévisions de la production. D'après la figure 1.1, les modèles statistiques utilisent les données mesurées de la puissance produite de la centrale ou les sorties du modèle NWP pour évaluer directement les prévisions de la production éolienne. Ils n'ont donc pas besoin d'information sur les caractéristiques physiques de la centrale éolienne, une relation statistique étant directement évaluée entre les données mesurées et la puissance produite de la centrale éolienne; ou entre les sorties du NWP et la puissance produite sans tenir compte de la topographie ou de la rugosité du terrain. Par conséquent, dans cette famille de modèle, il n'y a pas usage d'un MOS, ou d'un module de CFD pour évaluer les prévisions de vent au niveau de la centrale éolienne. Les deux familles de modèles se diffèrent aussi par leurs performances. Les modèles statistiques, bien qu'ils soient simples, offrent une meilleure précision que les modèles physiques pour les horizons de prévisions plus courts (allant jusqu'à 2 ou 3 heures). Mais, cette précision se dégrade rapidement pour des horizons de prévision plus longs et les modèles physiques deviennent meilleurs dans ce cas. Certains modèles de prévision utilisent

une approche combinée des deux familles, soit en combinant l'approche physique à celle statistique qui procure une meilleure précision.

1.3 La performance des modèles

L'évaluation des modèles de prévision de la production éolienne se fait à l'aide des critères statistiques qui permettent de calculer leurs incertitudes. Les critères statistiques les plus utilisés sont le biais (*Bias*), l'erreur absolue moyenne (*MAE*), l'erreur quadratique moyenne (*RMSE*) et le coefficient de régression (R^2). Cette section a pour but de montrer comment sont calculés ces différents critères et comment les valeurs obtenues sont interprétées en prévision de la production éolienne. Dans sa thèse de doctorat, (Lange, 2003) décrit en détail les équations utilisées en expliquant leurs significations. En notant P_{t+k} la valeur de la puissance mesurée et $\hat{P}_{t+k/t}$ la valeur de la puissance prédite, l'erreur de la prévision peut s'interpréter comme la différence entre la valeur de la puissance prédite et celle mesurée, ce qui est décrit par l'équation 1.1.

$$\epsilon = \hat{P}_{t+k/t} - P_{t+k} \quad (1.1)$$

L'erreur quadratique moyenne peut donc être évaluée en utilisant l'équation 1.2:

$$RMSE = \sqrt{\bar{\epsilon}^2} \quad (1.2)$$

où $\bar{\epsilon}$ représente la moyenne temporelle de ϵ obtenue en évaluant le modèle à l'aide d'un ensemble de données tests. Cette moyenne est par définition le biais du modèle. L'équation 1.3 peut donc s'écrire :

$$Bias = \bar{\epsilon} \quad (1.3)$$

Le *RMSE* peut être décomposé en utilisant les équations suivantes :

$$RMSE^2 = Bias^2 + SDE^2 \quad (1.4)$$

où $SDE = \sigma(\epsilon)$ représente l'écart-type de l'erreur sur la prévision de la puissance. Dans sa thèse de doctorat, Lange (2003) souligne que le *RMSE* peut encore se décomposer suivant l'équation 1.5 :

$$RMSE^2 = Bias^2 + SDBias^2 + DISP^2 \quad (1.5)$$

dans laquelle :

$$SDBias = \sigma\left(\hat{P}_{t+k/t}\right) - \sigma(P_{t+k}) \quad (1.6)$$

et

$$DISP = \sqrt{2\sigma\left(\hat{P}_{t+k/t}\right)\sigma(P_{t+k})\left(1 - r\left(\hat{P}_{t+k/t}, P_{t+k}\right)\right)} \quad (1.7)$$

où r est le coefficient de corrélation croisée entre $\hat{P}_{t+k/t}$ et P_{t+k} .

Tambke *et al.* (2005) ont fait une étude de la contribution de chacun de ces paramètres sur la nature de l'erreur : le *Bias* représente l'erreur systématique du modèle, *SDBias* représente l'erreur d'amplitude et *DISP* représente l'erreur de phase du modèle. En faisant une erreur d'amplitude, le modèle prédit la puissance en se trompant sur son intensité. Tandis qu'une erreur de phase indique un décalage dans le temps entre la valeur mesurée et la valeur prédite.

Les critères statistiques précédents permettent de comprendre et d'interpréter les erreurs des modèles de prévision. Mais le plus souvent, les modèles sont évalués en les comparant avec les modèles de référence. Le modèle de référence Persistance est le plus utilisé en prévision

éolienne. C'est un modèle bien simple, qui suppose que la valeur de la puissance prédite soit égale à celle mesurée à l'instant où la prévision est calculée. Ce modèle se décrit donc par l'équation 1.8 :

$$\hat{P}_{t+k/t} = P_t \quad (1.8)$$

Le modèle Climatologie est le deuxième modèle de référence souvent utilisé pour évaluer et comparer les modèles. Climatologie suppose que la puissance prédite à un temps t correspond à la moyenne des puissances mesurées disponibles.

$$\hat{P}_{t+k/t} = \bar{P} \quad (1.9)$$

Nielsen *et al.* (1998) ont estimé qu'il n'est pas raisonnable d'évaluer les modèles de prévision de la production éolienne en les comparant avec Persistance particulièrement pour des horizons de prévision plus grand que quelques heures. Ils ont alors proposé le nouveau modèle de référence qui combine les prévisions effectuées par Persistance et Climatologie. Le nouveau modèle de référence évalue les prévisions de la production en utilisant 1.10, soit :

$$\hat{P}_{t+k/t} = a_k P_t + (1 - a_k) \bar{P} \quad (1.10)$$

Lorsque k est petit, a_k prend une valeur sensiblement égale à 1 et la prévision correspond à celle évaluée par Persistance. Si k est grand, a_k sera sensiblement égale à zéro et la prévision correspond à celle calculée par Climatologie. Alors, a_k , qui est le coefficient de corrélation entre P_t et P_{t+k} , s'évalue de la manière suivante :

$$a_k = \frac{\frac{1}{N} \sum_{t=1}^{N-k} \tilde{P}_t \tilde{P}_{t+k}}{\frac{1}{N} \tilde{P}_t^2} \quad (1.11)$$

où

N est le nombre de données de puissance mesurée disponible; et,

$$\tilde{P}_t = P_t - \bar{P} \quad (1.12)$$

Pour comparer les modèles développés avec les modèles de référence, il est recommandé d'utiliser le coefficient d'amélioration (Bédard, 2010) généralement appelé 'Skill Score'. Le Skill Score fait référence à la différence entre l'erreur du modèle développé et celle du modèle de référence et est normalisée par rapport à l'erreur du modèle de référence. L'équation permettant d'évaluer le Skill Score est la suivante :

$$Skill\ Score = \frac{E_{ref} - E_{mod}}{E_{ref}} \quad (1.13)$$

où : E_{ref} est l'erreur du modèle de référence (*RMSE* ou *MAE*); et,
 E_{mod} est l'erreur du modèle développé.

1.4 Avantages de la prévision régionale

Focken *et al.* (2001) ont démontré que la performance des modèles de prévision de la production éolienne peut être améliorée en prédisant la puissance totale produite par un ensemble de centrales éoliennes d'une région, comparée à la prévision de la puissance produite par une centrale éolienne individuelle. Le processus qui consiste à partir des prévisions de la production d'une centrale éolienne individuelle pour la prévision de la production de l'ensemble des centrales d'une région est appelé *upscaling*. L'idée consiste à évaluer les prévisions d'un ensemble de centrales éoliennes d'une région ou raccordées à un même réseau électrique pour augmenter la précision du modèle et réduire les coûts computationnels qui pourraient être reliés à plusieurs calculs de prévisions de la production des différentes centrales éoliennes individuelles. Monteiro *et al.* (2009) présentent trois

différentes approches qui peuvent être utilisées pour évaluer les prévisions régionales de la production éolienne :

- L'approche directe

Elle consiste à établir un lien direct entre la production totale de l'ensemble des centrales éoliennes de la région et les prévisions météorologiques au niveau d'un ou de plusieurs centrales éoliennes de cette région. C'est donc une approche bien simple mais qui requiert les informations détaillées de la production totale de l'ensemble des centrales éoliennes. Le principal inconvénient de cette approche est que la relation statistique qui évalue la production éolienne régionale doit être mise à jour chaque fois qu'une centrale éolienne est ajoutée dans la région.

- L'approche en cascade

Elle se fait en deux étapes. La première consiste à évaluer les prévisions de la production des centrales éoliennes individuelles. En effectuant la somme des puissances prédites par les centrales éoliennes individuelles, il est possible d'estimer les prévisions de la production de l'ensemble des centrales éoliennes de la région. Cette approche est l'une des plus utilisée pour évaluer les prévisions de la production éolienne régionale (Monteiro *et al.*, 2009).

- L'approche en cluster

Elle consiste à regrouper les centrales éoliennes en plusieurs groupes d'une même région géographique et utiliser une approche directe pour évaluer les prévisions de la production de chaque groupe. L'approche en cascade est donc utilisée pour partir des prévisions de la production de chaque groupe aux prévisions de l'ensemble de la région.

1.5 Quelques exemples de modèle de prévision de la production éolienne

Plusieurs modèles de prévision de la production éolienne ont déjà été développés et sont opérationnels dans tous les pays exploitant l'énergie d'origine éolienne. Dans cette section, la méthodologie utilisée par quelques modèles est décrite, ainsi que les performances observées.

Prediktor

Prediktor est un modèle physique de prévision de la production éolienne développé par Landberg (Monteiro *et al.*, 2009). Il utilise les sorties du modèle de prévision météorologique HIRLAM à 57 km de résolution et est capable de fournir les prévisions de la production d'une centrale éolienne jusqu'à 48 heures à l'avance. Prediktor exploite les prévisions de la vitesse du vent et son orientation évaluées par HIRLAM. Ces variables sont introduites dans le module WASP qui prend en compte la topographie et la rugosité pour prédire le vent sur la centrale éolienne avec une résolution de quelques mètres. Le programme PARK, qui est introduit actuellement dans WASP, se charge de convertir les prévisions météorologiques au sein de la centrale en prévisions de puissance produite. La particularité de PARK est qu'il est capable d'évaluer les pertes de production de la centrale éolienne en calculant l'influence du sillage sur chaque turbine individuelle. Un module de traitement statistique des sorties est utilisé à deux étapes du processus. L'un est introduit pour corriger les prévisions du vent sur la centrale éolienne et l'autre introduit après la prévision de puissance pour corriger les erreurs systématiques. Le modèle Prediktor n'a pas un algorithme pour évaluer la prévision éolienne régionale, c'est donc un modèle qui évalue les prévisions de la production des centrales éoliennes individuelles. Les performances enregistrées par Prediktor montrent qu'il est capable de battre le modèle de référence Persistence à partir de la quatrième heure de prévision. Des tests ont été élaborés par Electric Power Research Institute (EPRI) aux USA en prédisant la puissance de deux centrales éoliennes avec Prediktor. Pour la première centrale éolienne de 900 turbines d'une capacité totale de 90 MW, un MAE de 14.2% a été enregistré. Concernant la deuxième centrale éolienne de 111 éoliennes, dont la capacité totale est de 66.6 MW, une erreur de 22.3% a été enregistrée (Monteiro *et al.*, 2009). Ces performances ont été évaluées avec des données couvrant une période de temps d'une année.

LocalPred

LocalPred est un modèle physique de la production développé par le Centre national espagnol des énergies renouvelables (CENER), l'ancien Centre de recherche pour l'énergie, l'environnement et la technologie (CIEMAT), et prédit la puissance des centrales éoliennes en Espagne depuis l'année 2003 (Monteiro *et al.*, 2009). Il utilise les sorties d'un modèle numérique de prévisions météorologiques et les introduit dans un module MM5 qui est capable de calculer les prévisions du vent avec une résolution spatiale de 1km et peut aller jusqu'à un horizon de prévision de 72 heures. Un module CFD utilise les prévisions du vent fournies par MM5, de même que les informations sur la rugosité et la topographie pour évaluer les prévisions météorologiques avec une résolution spatiale de quelques mètres. Un MOS est appliqué aux sorties du modèle de prévisions météorologiques pour corriger les erreurs systématiques de prévision. Pour convertir les prévisions de vent en prévisions de la puissance, le module de puissance W2P utilise la courbe de puissance de la centrale développée pour chaque secteur de direction de vent et pour différentes valeurs de la masse volumique de l'air. RegioPred est le module chargé de calculer les prévisions de la production éolienne régionale à partir des prévisions de la puissance des centrales individuelles évaluées par LocalPred.

WPPT (Wind Power Prediction Tool)

WPPT a été développé par l'Institut de l'informatique et de la modélisation mathématique (IMM) de l'Université technique du Danemark (DTU). C'est un modèle statistique de prévision de la production éolienne qui est capable d'évaluer les prévisions de la production d'une centrale individuelle aussi bien que celle d'une région entière (Madsen *et al.*, 2005). Une particularité de ce modèle de prévision est qu'il est capable de fournir une prévision de l'erreur associée à la valeur de la puissance prédite. Il combine deux approches différentes d'évaluation de la prévision de la puissance. La première consiste à utiliser une courbe de puissance définie par secteur de direction de vent pour convertir les prévisions météorologiques à grande échelle en prévisions de la puissance produite de la centrale. La deuxième consiste à trouver une relation statistique entre les puissances antérieures mesurées de la centrale éolienne et la puissance prédite. Pour effectuer la prévision régionale, WPPT

utilise une combinaison de deux différentes approches. L'une consiste à utiliser l'approche directe, décrite au paragraphe 1.4, qui établit une relation entre le NWP et la production totale d'énergie de la région, en utilisant une approche similaire à celle utilisée pour évaluer les prévisions de la production d'une centrale. La seconde approche consiste à sommer les puissances prédites par les centrales éoliennes individuelles, ce qui permet d'obtenir les prévisions de la production de la région. Le modèle combine donc les deux dernières approches en faisant une moyenne pondérée des puissances prédites par chacune d'entre elles. Les poids sont mis à jours continuellement et s'évaluent en fonction de la valeur du *RMSE*. Ce modèle a été testé en Australie par la compagnie Tasmania Hydro pour la prévision de la production d'une centrale éolienne d'une capacité totale de 64.75MW. WPPT a enregistré une erreur quadratique moyenne qui varie entre 9% et 27% pour les prévisions de la puissance à plus de 36 heures à l'avance (Monteiro *et al.*, 2009).

GH Forecaster

Garrad Hassan and Partners ont développé un modèle capable de prédire la puissance produite des centrales éoliennes (Monteiro *et al.*, 2009). GH Forecaster utilise les prévisions météorologiques d'un modèle numérique à 12 km de résolution. Une technique de régression linéaire est ensuite utilisée pour partir des sorties du modèle numérique à la prévision de la vitesse du vent, sa direction, la température et la pression au niveau du mât de mesures météorologiques. Pour permettre aux prévisions météorologiques de s'auto-corriger, GH Forecaster utilise les données enregistrées de la centrale éolienne comme entrées du modèle de prévision du vent. La conversion des prévisions météorologiques en prévisions de la puissance est effectuée par une courbe de puissance définie par secteur de direction du vent. Cette courbe est aussi évaluée par Bins de vitesses du vent et est normalisée par rapport à la densité de référence. GH Forecaster a été testé en prédisant la puissance d'une centrale éolienne situé en terrain complexe, où le *MAE* varie entre 10% et 20% tandis que le *RMSE* varie entre 15% et 26% pour des horizons de prévision variant d'une heure à 60 heures. Une comparaison avec persistance a permis d'observer une amélioration autour de 40% (Monteiro *et al.*, 2009).

eWind

eWind est le modèle de prévision de la production éolienne développé par AWS Truewind Inc. (Monteiro *et al.*, 2009). Les sorties d'un modèle régional de prévisions météorologiques sont introduites dans le module ForeWind qui utilise les caractéristiques physiques du terrain pour calculer les prévisions du vent au niveau de la centrale éolienne. Il utilise aussi d'autres modules tels que MM5 qui exploitent aussi les sorties du modèle régionale pour calculer les prévisions du vent au niveau de la centrale éolienne. Le MOS utilise les données mesurées de la centrale pour corriger les erreurs systématiques des différents modèles météorologiques et produire les prévisions du vent au niveau des différents mâts de mesures de la centrale éolienne. Pour convertir les prévisions météorologiques en prévisions de puissance, deux approches sont possibles. La première consiste à établir une relation statistique entre la moyenne des vitesses à la hauteur des éoliennes et la puissance totale de la centrale éolienne, tandis que la deuxième consiste à utiliser un réseau de neurones qui utilise comme entrées la vitesse du vent, son orientation, la variabilité de la vitesse et la stabilité atmosphérique à un point de référence et fournit à la sortie la puissance totale de la centrale éolienne. Une étape finale du modèle consiste à faire une vérification de la qualité de prédiction évaluée par le modèle. La méthodologie utilisée par eWind est décrite par (Monteiro *et al.*, 2009).

Previento

Previento est un modèle physique de prévision de la production éolienne opéré par Energy and Meteo Systems GmbH (Focken *et al.*, 2007). Il utilise les prévisions météorologiques de plusieurs modèles numériques pour capturer les forces et les faiblesses de chaque modèle selon les situations météorologiques. Une modélisation physique de la basse atmosphère permet alors d'évaluer les prévisions du vent au niveau de la centrale et à la hauteur des éoliennes en tenant compte de la stratification thermique. Deux étapes sont requises pour convertir les prévisions météorologiques au niveau de la centrale éolienne en prévisions de la puissance produite. La première consiste à utiliser les courbes de puissances certifiées des éoliennes et d'évaluer les pertes par effets de sillage dues à la géométrie de la centrale. La deuxième étape utilise les données mesurées de la puissance totale de la centrale pour capturer l'influence de la direction du vent et des caractéristiques de la centrale (topographie

et rugosité) sur la production de puissance. Cette étape permet alors d'optimiser la prévision et d'adapter le modèle en fonction des caractéristiques de chaque centrale éolienne. La deuxième étape n'étant pas obligatoire, Previento présente la capacité de prédire la puissance produite par une centrale éolienne sans nécessairement utiliser les données mesurées. Cela pourrait être particulièrement bénéfique pour une centrale éolienne qui vient d'être opérationnelle et n'a pas encore une banque de données mesurées. Ce modèle fournit aussi une marge d'incertitude sur la prévision de la puissance délivrée. Cette incertitude est obtenue à l'aide de la plus petite et la plus grande valeur de la puissance prédite obtenues en exploitant les différents modèles de prévisions météorologiques disponibles.

Dans ce chapitre, il a principalement été question d'introduire les concepts fondamentaux utilisés en prévision de la production éolienne. La description des familles des modèles, des techniques d'évaluation des modèles et des différentes étapes requises pour effectuer les prévisions de la production éolienne peuvent permettre de comprendre le fonctionnement des différents modèles opérationnels qui ont été décrits. Un fait essentiel qu'il est important de mentionner est que ces modèles opérationnels sont pour la plupart commercialisés et par conséquent les descriptions de leurs fonctionnement telles que retrouvées dans la littérature sont plutôt d'ordre général. Les chapitres 3 et 4, qui constituent le développement de cette étude, seront introduits par une revue de la littérature plus spécifique à chacune des parties, ce qui permettra de mieux exposer les choix qui auront été faits.

CHAPITRE 2

CONTRÔLE DE LA QUALITÉ DES DONNÉES ET DESCRIPTION DES VARIABLES IMPORTANTES

Le développement de la courbe de puissance de la centrale éolienne qui sera présenté au chapitre 3 et son application pour convertir les prévisions météorologiques en prévisions de puissance présenté au chapitre 4 se font avec des données mesurées d'une centrale éolienne. La qualité de la courbe de puissance et du système de prévision de la production de la centrale éolienne dépendent donc de la quantité et de la qualité des données disponibles. Il est par conséquent nécessaire de s'assurer que les données brutes mesurées de la centrale éolienne suivent une procédure de contrôle de la qualité pour identifier les plages de données absentes, et écarter les données aberrantes. Une contrainte importante à laquelle on doit faire face en modélisant la puissance d'une centrale éolienne ayant plusieurs éoliennes est énoncée clairement par Negnevitsky *et al.* (2008) : les arrêts des turbines peuvent refléter des données absentes ou aberrantes; il est donc nécessaire de les identifier et de normaliser la puissance totale de la centrale éolienne dans ces conditions en prenant en compte le nombre de turbines disponibles.

Comme première approximation, la courbe de puissance de la centrale qui sera présentée au chapitre 3 reflètera donc la centrale éolienne dans les conditions de fonctionnement normal, avec toutes les turbines disponibles. Par la suite, l'étude des arrêts des turbines au chapitre 4 permettra de prendre en compte les plages de données pour lesquelles un certain nombre d'éoliennes sont arrêtées. Ce nombre sera déterminé de manière expérimentale, en étudiant la fréquence des arrêts des turbines de la centrale éolienne. Dans ces conditions, la procédure de contrôle de la qualité des données présentée dans ce chapitre s'effectue en deux étapes. La première consiste à identifier les données aberrantes et les données manquantes. La deuxième consiste à identifier les plages de données pour lesquelles des turbines sont arrêtées. Le bilan de la procédure et la courbe de puissance de la centrale seront présentés, ce qui permettra de s'assurer que le nombre de données conservées, et leur qualité, sont suffisants pour modéliser

la courbe de puissance de la centrale éolienne. L'analyse de la ressource éolienne de la centrale (distribution de Weibull et rose des vents de la centrale éolienne) permettra de cerner les caractéristiques de vent sur le site.

Les quatre parties de ce chapitre seront présentées de la manière suivante : la première partie permettra de décrire les données mesurées disponibles de la centrale éolienne; par la suite, il sera décrit la procédure utilisée pour identifier les données aberrantes; la troisième partie permettra de montrer le bilan de la procédure du contrôle de la qualité des données; et enfin la dernière partie de ce chapitre montrera les équations utilisées pour calculer les variables nécessaires pour mener cette étude. Ces variables sont donc calculées à partir des données sélectionnées.

2.1 Les données mesurées

Les données disponibles pour mener l'étude proviennent d'une centrale éolienne située au Canada. La centrale éolienne expérimentale est composée de plusieurs éoliennes et plusieurs mâts de mesures météorologiques et a une capacité totale supérieure à 100 MW. Elle est située sur un terrain complexe. Les données de 66 éoliennes et d'un seul mât de mesures météorologiques ont été disponibles pour cette étude. Ces données ont été mesurées sur une plage de temps couvrant une année au complet; soit du 1^{er} janvier au 31 décembre de l'année 2009. Ces données sont mesurées et moyennées sur des intervalles de temps de 10min. Les données de chaque éolienne et celles du mât de mesures sont fournies sur des feuilles Excel. Les fichiers des données mesurées aux éoliennes contiennent les valeurs des variables suivantes : la vitesse du vent à la hauteur de l'éolienne, la puissance produite par l'éolienne et la puissance réactive de l'éolienne. Le mât permet de mesurer plus de variables météorologiques dont : la vitesse du vent à plusieurs hauteurs, la température, la pression, l'humidité relative de l'air et la direction du vent. Pour chacune des variables mesurées, il est aussi disponible les informations telles que l'écart-type, et le minimum et le maximum de chaque variable pour chaque intervalle de 10min. Parmi ces données, certains intervalles de

10min n'ont pas pu être relevés. La procédure suivante pourra permettre de les identifier et de supprimer tous les intervalles de temps correspondants.

2.2 La procédure de contrôle de la qualité des données

La procédure qui permet d'identifier les données absentes et les données aberrantes se fait en trois étapes. La première étape consiste à identifier les données suspectes ou « flags » qui reflètent des situations météorologiques improbables, ou des situations météorologiques qui peuvent conduire à un mauvais fonctionnement des instruments de mesures, tel que le givrage. Cette étape se fait automatiquement en définissant des critères physiques à l'aide d'un algorithme et en répertoriant les dates et les heures auxquelles ces événements sont observés. La deuxième étape consiste à tracer des graphiques de l'évolution temporelle des valeurs mesurées par les instruments de mesures afin de visualiser et de confirmer les anomalies détectées dans les données. Cette analyse permet d'observer les incohérences des instruments de mesures, ou d'interpréter les « flags » qui ont été détectés. Elle se fait donc visuellement. Après avoir visualisé les données suspectes, la troisième étape consiste à décider d'écarter les données suspectes, ou de les conserver. La figure 2.1 résume la procédure qui permet de filtrer les données.

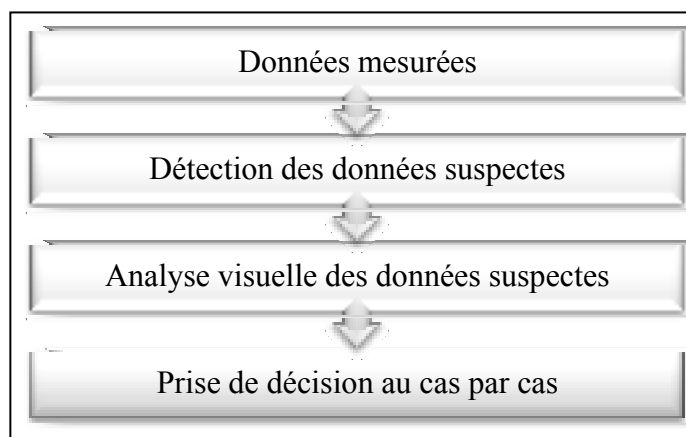


Figure 2.1 Procédure de contrôle de la qualité des données

La procédure utilisée a été rédigée par AWS Scientific Inc.(1997), et permet d'identifier les données suspectes qui révèlent des phénomènes physiques improbables, ou pouvant conduire à une lecture erronée des instruments de mesures. Pour ce faire, l'ensemble des données doit subir trois principaux tests à savoir :

- Le test d'intervalle : Ce test permet de comparer les données mesurées aux valeurs minimales et maximales physiquement possibles. Par exemple pour un site donné, une vitesse de vent négative présentée par l'anémomètre indique clairement un problème. Par contre, une vitesse de vent affichée supérieure à 30m/s peut être très peu probable dépendamment des caractéristiques du vent du site. Cet évènement pourrait donc nécessiter une vérification manuelle. Ce test est appliqué sur l'ensemble des données suivantes : la vitesse du vent, la direction du vent, la température et la pression.
- Le test de relation : Ce test permet de comparer les valeurs de différents instruments de mesures en se basant sur les relations physiques escomptées de ces différents instruments. Par exemple, d'après le profil vertical de la vitesse du vent, l'on ne doit pas s'attendre à observer une vitesse du vent à 40m qui soit supérieure à la vitesse du vent présentée par un anémomètre à 80m. Cela est toute fois observable dans certaines conditions météorologiques ou sous l'influence du terrain. La comparaison entre les lectures de ces deux instruments de mesures pourra donc permettre de détecter des anomalies.
- Le test de tendance : Ce test évalue l'évolution de certains paramètres dans le temps afin de s'assurer que les variations temporelles de ces derniers soient physiquement probables.

L'ensemble de ces trois tests permet donc de détecter les données suspectes. Une vérification des données a été faite au préalable pour identifier et marquer les dates pour lesquelles des instruments de mesures n'ont pas enregistré les données. Les tableaux 2.1, 2.2 et 2.3 regroupent de manière plus précise l'ensemble des conditions qui ont été utilisées et les valeurs choisies pour identifier les données aberrantes. À ces conditions s'est ajouté un test pour identifier les conditions de givrage. La formation de givrage sur les instruments de

mesures a été identifiée par une vitesse nulle, une température en dessous de 2° Celsius et l'humidité relative au dessus de 80% ou l'écart-type de la direction du vent inférieur à 3, la température en dessous de 2° Celsius et l'humidité relative en dessus de 80%. Le contrôle de la qualité des données est appliqué essentiellement au mât de mesures. Les conditions de givrage qui seront observées permettront de supprimer toutes les dates de ces événements dans les données des éoliennes. La condition des arrêts des turbines qui sera appliquée aux turbines pourra permettre de sélectionner les données pour lesquelles la centrale éolienne fonctionne en condition normale d'opération, avec toutes les turbines disponibles. La procédure de contrôle de la qualité des données sera validée en observant les courbe de puissance de la centrale éolienne avant et après la procédure et en comparant le nombre de données mesurées au nombre de données sélectionnées.

Tableau 2.1 Description du test d'intervalle

Paramètres	Critères de validation
Vitesse du vent	
Moyenne sur 10min (V10min)	$3 < V_{10min} < 25$
Écart-type sur 10min (StdV10min)	$0 < StdV_{10min} < 3$
Maximum sur 10min (MaxV10min)	$3 < MaxV_{10min} < 30$
Direction du vent	
Moyenne sur 10min (D10min)	$0 < D_{10min} < 360$
Température	
Moyenne sur 10min (T10min)	$-40 < T_{10min} < 40$

Tableau 2.2 Description du test de relation

Paramètres	Critères de validation
Vitesse du vent	
MaxV10min & V10min	$MaxV_{10min} < 2.5 \times V_{10min}$
V80m & V50m	$V_{80m} - V_{50m} < 2 \text{ m/s}$

Tableau 2.3 Description du test de tendance

Paramètres	Critères de validation
Vitesse du vent	
Variation de V10min en 1 heure	< 5m/s
Température	
Variation de T10min en 1 heure	< 5° C

Dans les tableaux 2.1, 2.2 et 2.3, les critères de validation sont les exigences que les données doivent satisfaire pour être considérées comme a priori valides. Les dates et les heures pour lesquelles les données ne satisfont pas ces critères sont identifiées et marquées. Une visualisation des données permet alors de prendre une décision au cas par cas sur les données à conserver et à supprimer.

2.3 Bilan du contrôle de la qualité des données

La procédure décrite au paragraphe précédent a donc été utilisée pour identifier les données suspectes. Le tableau 2.4 présente le bilan de l'ensemble des données qui ont été identifiées comme suspectes pour chacun des critères.

Tableau 2.4 Bilan des données suspectes

Type de test	Critères	Nombre de cas observés	Pourcentage
Manquantes	-	10214	19.43
Test d'intervalle	$3 < V_{10min} < 25$	4973	9.46
	$0 < StdV_{10min} < 3$	178	0.34
	$3 < MaxV_{10min} < 30$	7200	13.70
	$0 < D_{10min} < 360$	0	0
	$-40 < T_{10min} < 40$	0	0
Test de relation	$MaxV_{10min} < 2.5 \times V_{10min}$	708	1.35
	$V_{80m} - V_{50m} < 2 \text{ m/s}$	720	1.37
Test de tendance	$\Delta V_{10min 1h} < 5 \text{ m/s}$	282	0.54
	$\Delta T_{10min 1h} < 5^\circ \text{ C}$	36	0.06
Givrage*	$V_{10min}=0 \ \& \ T_{10min}<2^\circ \ \& \ P_{10min}>80$ ou $StdD_{10min}<3 \ \& \ T_{10min}<2^\circ \ \& \ P_{10min}>80$	972	1.85
Total		21356	40.63

* Les critères du givrage sont ceux correspondants aux données suspectes contrairement à tous les autres critères qui correspondent aux données valides ne nécessitant pas de vérification.

La vérification visuelle des données suspectes a permis de trier l'ensemble des données en deux groupes : les données valides et celles non valides. La décision de garder les données suspectes ou de les supprimer a été prise sur la base des graphiques journaliers des données relevées par les instruments de mesures. Les figures 2.2, 2.3 et 2.4 montrent trois exemples de décisions qui ont été prises après vérification des données suspectes.

À la figure 2.2, la procédure a identifié des données suspectes, principalement à cause de la faible variation de l'écart-type de l'orientation du vent après 20h, une température en dessous de 2° C et une humidité relative au-dessus de 80%. La procédure a donc suspecté l'état de givrage pendant cette plage de temps. La visualisation des graphiques montre que la vitesse du vent varie de manière régulière pendant cette plage de temps. Un comportement normal est aussi observé en observant l'évolution de la direction du vent durant toute la journée. La

faible variation de l'écart-type de l'orientation du vent s'explique donc par le fait que la direction du vent a pratiquement été stable durant toute la journée. Bien que la température soit en dessous de -20°C , les graphiques ne montrent pas de manière claire la formation de givrage sur les instruments de mesures. Toutes les données de cette journée ont donc été conservées.

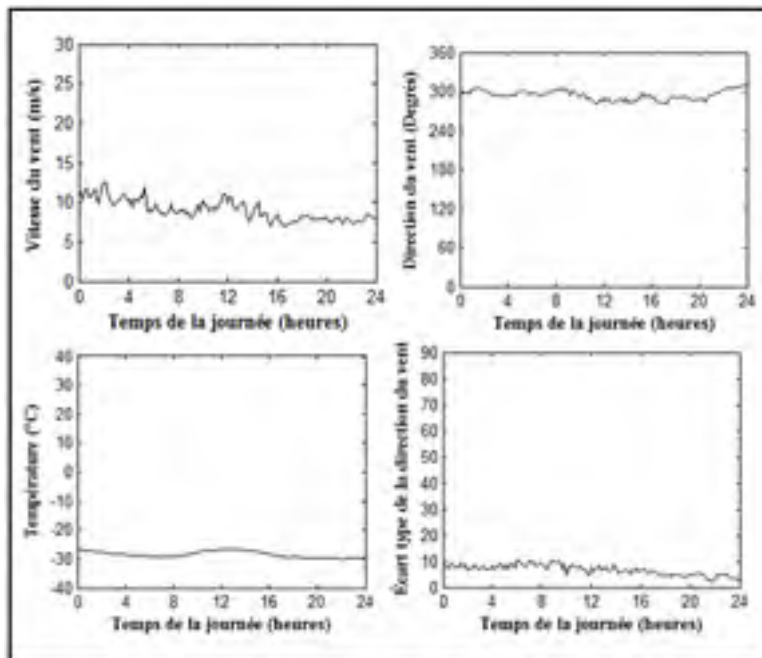


Figure 2.2 Exemple 1 de données suspectes

La figure 2.3 montre un exemple d'arrêt de fonctionnement des instruments de mesures. La vitesse du vent n'a pas été relevée dans les douze premières heures de la journée. Quant à la direction du vent, aucune information n'est disponible pour la journée au complet. Il est important d'avoir toutes les variables disponibles dans l'ensemble des données sélectionnées pour éviter les erreurs lors des calculs et maintenir les séries temporelles cohérentes. Pour ces raisons, cette journée a été supprimée de l'ensemble des données.

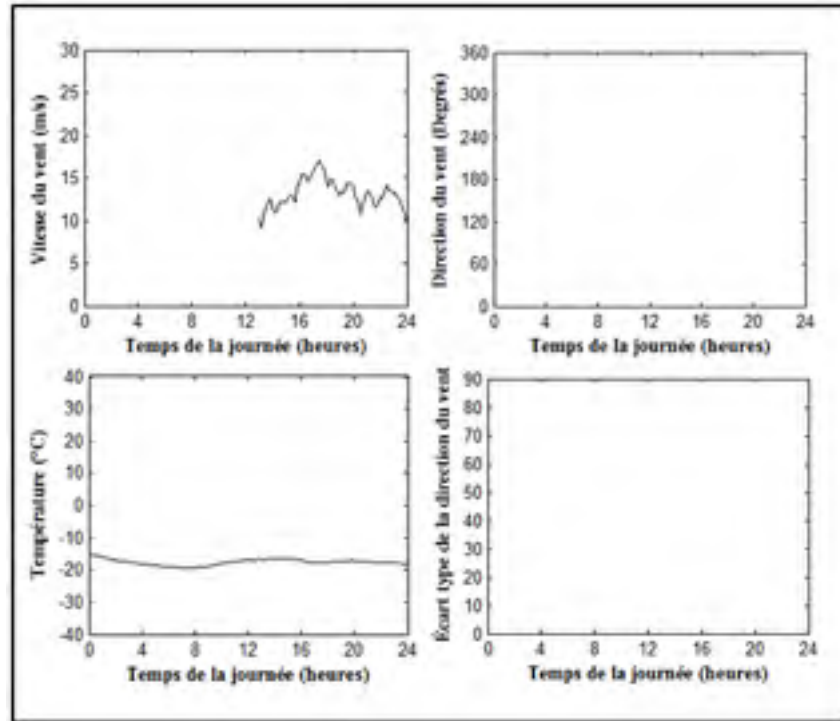


Figure 2.3 Exemple 2 de données suspectes

La figure 2.4 montre un exemple de cas pour lequel des données ont été supprimées. La vitesse du vent chute à zéro après 12h, ce qui porte à croire qu'il n'y a pas de vent sur la centrale éolienne entre 12h et 15h. L'observation de la direction du vent et de l'écart-type de l'orientation montre que la direction du vent change quand même dans cette plage de temps. De plus, l'écart-type de l'orientation montre aussi une variation de la direction du vent durant cette plage de temps, même si elle paraît faible. Cette information laisse croire qu'il vente quand même sur la centrale éolienne. Le givrage, qui a été identifié par la procédure, pourrait expliquer pourquoi les instruments de mesures sont moins sensibles dans cette plage de temps, la température étant en-dessous de $-10\text{ }^{\circ}\text{C}$. Durant cette journée, les données relevées de 12h à 20h ont alors été retirées de l'ensemble des données.

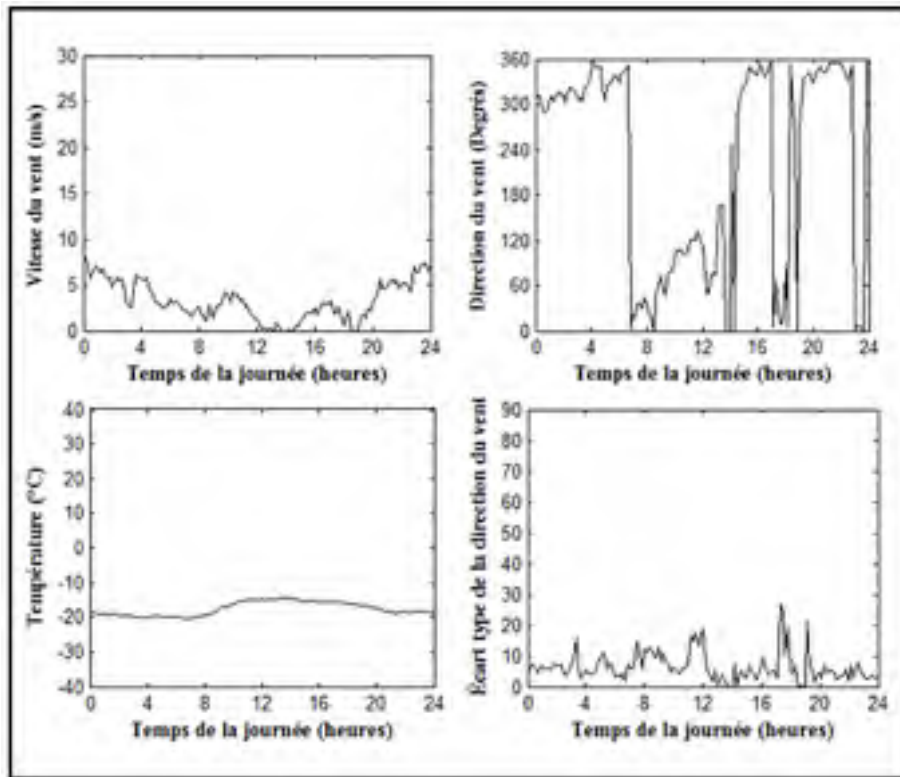


Figure 2.4 Exemple 3 de données suspectes

La procédure a été utilisée pour contrôler la qualité des données du mât de mesures. Les plages de temps supprimées après le contrôle de la qualité sont complètement retirées de l'ensemble des données du mât et de toutes les turbines. Le tableau suivant présente le bilan de la procédure de contrôle de la qualité de données du mât de mesures.

Tableau 2.5 Bilan des données supprimées

Mois	Nombre de données détectées	Nombre de données supprimées	Pourcentage de données supprimées
Janvier	1854	347	0.66
Février	2863	2711	5.16
Mars	2252	1380	2.63
Avril	1549	841	1.60
Mai	2210	1404	2.67
Juin	898	134	0.25
Juillet	1032	203	0.39
Août	1024	192	0.37
Septembre	868	416	0.79
Octobre	1687	398	0.76
Novembre	1376	757	1.44
Décembre	3739	3357	6.39
Total	21356	12144	23.11

L'ensemble des données supprimées couvre les plages de temps pendant lesquelles les données n'ont pas été relevées et les plages de temps pendant lesquelles la qualité des données a été identifiée comme de mauvaise qualité par la procédure qui a été utilisée. D'après le tableau 2.4, l'on peut s'apercevoir que 19% de données sont manquantes. Par ailleurs, la procédure de contrôle de la qualité des données a conduit à écarter 23% des données, les données manquantes étant incluses (Tableau 2.5). On peut donc conclure qu'environ 4 points de pourcentage des données ont été écartées à cause de leur mauvaise qualité. Ce nombre étant inférieur à 5 points de pourcentage, les données utilisées sont donc de bonne qualité. Le coefficient de recouvrement tel que présenté par AWS Scientific Inc. (1997) à l'équation 2.1 vaut donc 77%.

$$\%Recouvrement = \frac{\text{Nombre de données retenues}}{\text{Nombre total de données}} \times 100 \quad (2.1)$$

La figure 2.5 montre les caractéristiques moyennes de vent du site, soit la rose des vents et la distribution de Weibull pour la période complète des mesures. Ces caractéristiques ont été calculées avec les données sélectionnées.

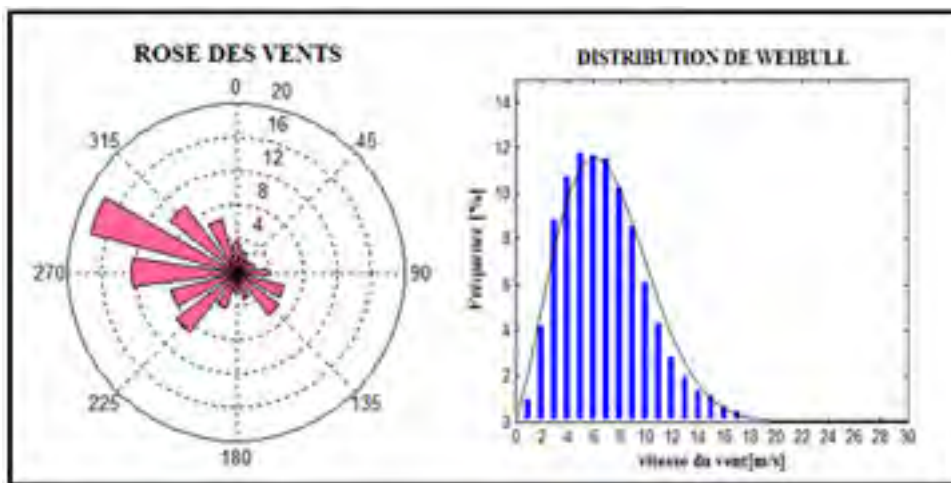


Figure 2.5 Caractéristiques de vent du site

Le deuxième volet du contrôle de la qualité des données permet de garantir que la centrale éolienne fonctionne en mode d'opération normale avec toutes les turbines disponibles. Cette hypothèse permet de maintenir les données qui reflètent le mieux possible la relation entre les variables météorologiques et la production de puissance sans prendre en compte d'autres facteurs externes. La figure 2.6 montre la relation entre la vitesse du vent lue au mât et la puissance totale de la centrale (obtenue en sommant la puissance des turbines individuelles). Les données présentées sont celles retenues après le contrôle de la qualité présenté précédemment. On peut apercevoir à la figure 2.6 l'apparition de plusieurs points pour lesquelles la puissance totale de la centrale a une valeur négative. Ces données correspondent aux arrêts des turbines. Plus le nombre de turbines arrêtées est élevé, plus on observe une chute de la puissance totale de la centrale éolienne dans les mêmes conditions météorologiques.

Dans cette étude, les plages de temps où les arrêts des turbines sont observés seront d'abord écartées afin de bien étudier la courbe de puissance. Dans la suite, ces données seront prises en compte lors de l'utilisation de la courbe de puissance pour prédire la puissance de sortie de la centrale éolienne. La figures 2.7 montre la courbe de puissance de la centrale éolienne en ayant enlevé les données correspondant aux arrêts des turbines (environ 50% du nombre total de données). Cet ensemble de données est donc de bonne qualité, et garantit un fonctionnement normal de la centrale éolienne.

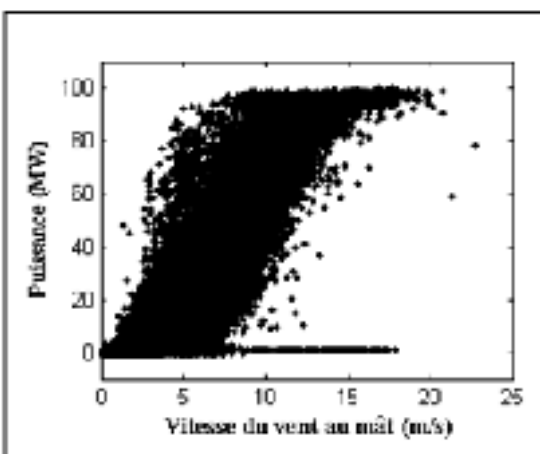


Figure 2.6 Courbe de puissance de la centrale avec arrêts des turbines

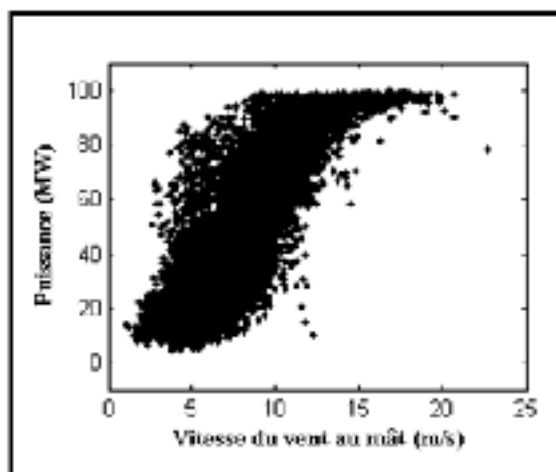


Figure 2.7 Courbe de puissance de la centrale sans arrêts des turbines

2.4 Description des variables importantes

Le contrôle de la qualité des données étant effectué, l'on a été capable de repérer les plages de données manquantes, d'écartier les données de mauvaise qualité et de conserver uniquement les données de bonne qualité. Tout au long de cette étude, plusieurs variables sont utilisées et sont par conséquent évaluées avec les données sélectionnées de la centrale éolienne expérimentale. L'objectif de cette section est de montrer comment ont été calculées les principales variables utilisées tout au long de cette étude.

La vitesse du vent

La vitesse du vent est mesurée au mât de mesures météorologiques, à la hauteur des nacelles des éoliennes pour chaque intervalle de temps de 10min. Elle est aussi disponible au niveau de la nacelle de chaque éolienne. Dans les chapitres 3 et 4, l'on fera référence à deux vitesses du vent : celle mesurée au mât de mesures et la moyenne des vitesses du vent à la nacelle des éoliennes. La moyenne des vitesses aux éoliennes est donc évaluée de la manière suivante :

$$V_{moy} = \frac{\sum_{i=1}^n V_i}{n} \quad (2.2)$$

où V_i représente la vitesse du vent au niveau de la nacelle de l'éolienne i et n est le nombre total d'éoliennes de la centrale.

La direction du vent

La direction du vent fera référence à la direction du vent lue au niveau du mât de mesures météorologiques, à la hauteur de 80m. Il n'est donc pas nécessaire de la calculer.

La puissance

Les données mesurées des éoliennes présentent la puissance P_i produite par chaque éolienne et pour chaque intervalle de temps de 10min. La puissance totale produite par la centrale éolienne peut donc être évaluée suivant l'équation 2.3 :

$$P_{totale} = \sum_{i=1}^n P_i \quad (2.3)$$

L'intensité de turbulence

L'intensité de turbulence fait référence aux irrégularités ou variations de la vitesse du vent ou sa direction (AWS Scientific Inc., 1997). Elle est évaluée au mât de mesures, avec la vitesse du vent pour chaque intervalle de 10min et l'écart-type de la vitesse du vent dans cet intervalle de temps. Elle a été évaluée suivant l'équation 2.4 :

$$IT = \frac{\sigma_V}{V} \quad (2.4)$$

où V représente la vitesse du vent lue à 80m de hauteur, au mât de mesures et moyennée sur un intervalle de temps de 10min et σ_V est l'écart-type de la vitesse du vent sur chaque intervalle de 10min.

Le cisaillement

Le cisaillement se définit comme la variation de la vitesse du vent en fonction de la hauteur. L'équation qui a été utilisée pour évaluer le cisaillement a été décrite par AWS Scientific Inc. (1997), et s'écrit de la manière suivante :

$$Sh = \frac{\text{Log}_{10} \left(\frac{V_2}{V_1} \right)}{\text{Log}_{10} \left(\frac{z_2}{z_1} \right)} \quad (2.5)$$

où :

V_2 est la vitesse à la hauteur z_2 ; et,

V_1 est la vitesse à la hauteur z_1 .

CHAPITRE 3

MODÉLISATION DE LA COURBE DE PUISSANCE

La courbe de puissance d'une éolienne peut se définir comme une relation entre les variables météorologiques et la production de puissance de l'éolienne. La figure 3.1 montre un exemple de courbe de puissance d'une éolienne. Cette courbe représente la relation entre la vitesse du vent et l'énergie produite par la turbine. Elle peut être utilisée pour mesurer les performances de l'éolienne. Dans ce cas, elle permet de détecter les comportements anormaux de la machine en mode d'opération afin d'y apporter des corrections. Tous les modèles physiques de prévision de la production éolienne utilisent la courbe de puissance pour convertir les prévisions météorologiques en prévisions de puissance (Siebert, 2008). La courbe de puissance du fabricant représente une relation simple entre la vitesse du vent et la puissance produite par l'éolienne; toutefois, les données mesurées de la vitesse du vent et de la puissance produite présentent une relation ayant beaucoup plus de dispersion (Collins *et al.*, 2009). La première n'est donc pas suffisante pour être utilisée à des fins de prévision de la production éolienne.

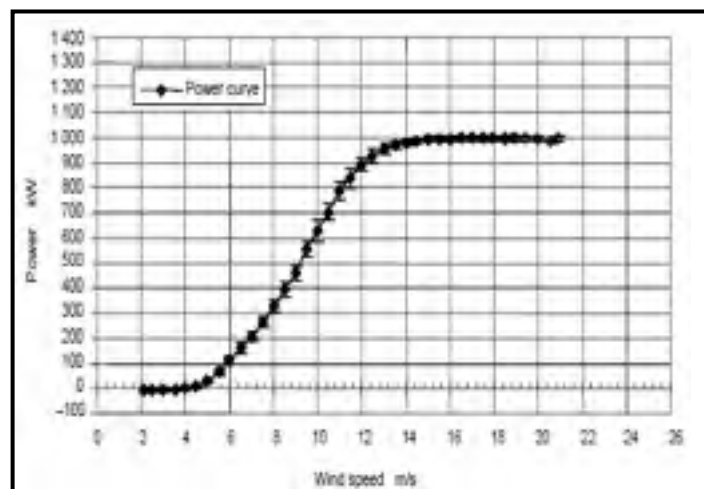


Figure 3.1 Exemple de courbe de puissance d'une éolienne
Tirée de IEC-61400-12-1 (2005)

Bien que la courbe de puissance d'une éolienne soit similaire à celle d'une centrale, elles présentent tout au moins des différences. La figure 3.2 montre une comparaison entre la courbe de puissance d'une centrale éolienne et la courbe de puissance d'une éolienne individuelle.

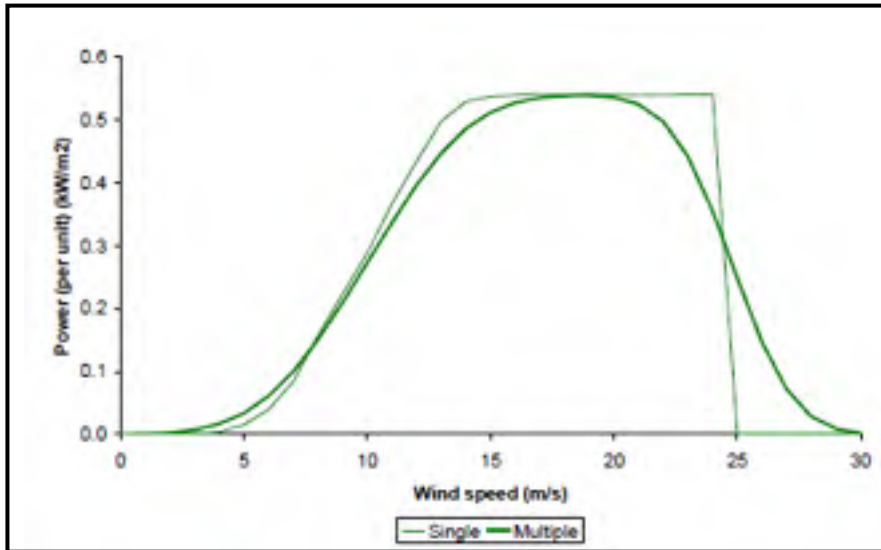


Figure 3.2 Comparaison entre la courbe de puissance d'une éolienne et la courbe de puissance d'une centrale
Tirée de Nørgård *et al.* (2004)

Les différences sont surtout remarquables aux niveaux des vitesses de démarrage et aux niveaux des vitesses de saturation. La courbe de puissance d'une centrale éolienne présente une pente plus faible pour les vitesses de vent faibles, comparativement à la courbe de puissance d'une éolienne. De plus, la puissance totale de la centrale éolienne décroît plus rapidement que celle de l'éolienne pour de fortes vitesses de vent. Ces différences sont dues au fait que toutes les éoliennes de la centrale ne perçoivent pas nécessairement la même vitesse du vent à un temps donné. L'une des causes étant les effets de sillage qui sont présentés au paragraphe 3.1.

Pour modéliser la fonction de transfert entre les variables météorologiques et la production de puissance, il est possible d'évaluer la courbe de puissance de chaque éolienne de la centrale

et ensuite, sommer la puissance totale obtenue. Cependant, cette approche ne procure pas nécessairement le meilleur résultat (Siebert, 2008). Touani *et al.* (2010) ont évalué la courbe de puissance de la centrale éolienne d'East Point situé à l'Île-du-Prince-Édouard. Les résultats obtenus ont montré que l'évaluation de la courbe de puissance de la centrale en calculant celle de chaque éolienne n'offre pas une amélioration significative comparée à une approche globale. Une étude sur un centrale éolienne située en Allemagne et composée de 17 éoliennes a permis de tirer des conclusions similaires (Saleck *et al.*, 2007). Il n'est donc pas nécessaire de modéliser la courbe de puissance de chaque éolienne afin de définir une fonction de transfert entre les variables météorologiques et la puissance totale produite de la centrale éolienne.

Les similitudes entre la courbe de puissance d'une éolienne et celle d'une centrale montrent que les méthodes utilisées pour modéliser la première peuvent être appliquées pour évaluer les performances de la centrale éolienne au complet. Cependant, il est possible que des conclusions différentes soient tirées. L'objectif de ce chapitre est de proposer une approche permettant de calculer la courbe de puissance d'une centrale éolienne, avec la meilleure précision possible. Afin d'évaluer l'influence du choix de la méthode, la méthode des Bins, servant de méthode de référence, est comparée à une méthode avancée d'apprentissage, telle que les réseaux de neurones artificiels (RNA), en gardant la même méthodologie. Par la suite, une influence de la méthodologie est effectuée en évaluant la courbe de puissance de la centrale avec la méthode des Bins et en utilisant deux approches différentes. La sélection des variables est effectuée en utilisant les réseaux de neurones. Les ajouts successifs de la direction du vent, de l'intensité de turbulence et du cisaillement à la courbe de puissance calculée précédemment avec la vitesse de vent et la masse volumique de l'air permettent d'apprécier l'influence de chaque paramètre sur la courbe de puissance de la centrale. Les études sont effectuées à l'aide des données d'une centrale éolienne canadienne pour laquelle une année complète de données est disponible. 80% des données sélectionnées par la procédure de contrôle de la qualité permettent de modéliser la courbe de puissance, 20% des données permettent de la tester. La courbe de puissance développée servira de fonction de

transfert entre les prévisions météorologiques et les prévisions de puissance de la centrale éolienne étudiée.

3.1 L'impact du sillage sur la courbe de puissance de la centrale éolienne

La norme IEC (IEC-61400-12-1, 2005) mentionne que la lecture de la vitesse du vent par un mât de mesures situé près des obstacles ou des turbines éoliennes peut être influencée. En effet, le vent qui traverse une éolienne ou un obstacle connaît une diminution de son énergie cinétique. Il lui faut parcourir une certaine distance avant de retrouver sa vitesse non perturbée. Cet effet peut influencer l'évaluation de la courbe de puissance de la centrale éolienne, calculée avec la vitesse du vent mesurée au mât. Afin d'évaluer l'impact du sillage sur la courbe de puissance de la centrale, le calcul des secteurs obstrués du mât de mesures météorologiques s'effectue en suivant la procédure détaillée dans la norme IEC (IEC-61400-12-1, 2005). Le mât de mesures météorologiques W2 de la centrale éolienne est représenté à la figure 3.3. Il est remarquable de la présence de quelques éoliennes dans son voisinage, ce qui constitue des obstacles potentiellement capables de créer des effets de sillage dans des directions de vent qui peuvent être évaluées suivant la norme IEC. L'éolienne la plus proche du mât W2 est située à une distance qui représente environ 2.8 fois le diamètre du rotor. Les effets de sillage du mât W2 par les éoliennes voisines sont étudiés en calculant les directions du vent influencées par chaque éolienne.



Figure 3.3 Position des turbines voisines au mât W2

3.1.1 Procédure d'évaluation des secteurs obstrués par le sillage

Les caractéristiques nécessaires pour évaluer les secteurs obstrués sont : la hauteur de l'éolienne, la largeur de l'obstacle en fonction du diamètre du rotor, et la distance entre l'éolienne et le mât de mesures météorologiques. Premièrement, le diamètre du rotor équivalent D_e est déterminé par l'équation 3.1 :

$$D_e = \frac{2l_h l_w}{l_h + l_w} \quad (3.1)$$

où l_h est la hauteur de l'obstacle; et,
 l_w est la largeur de l'obstacle.

Le calcul du diamètre du rotor équivalent D_e permet d'évaluer l'angle correspondant au secteur obstrué par l'obstacle situé à la distance L_e par l'équation 3.2.

$$\alpha = 1.3 \arctan\left(\frac{2.5D_e}{L_e} + 0.15\right) + 10 \quad (3.2)$$

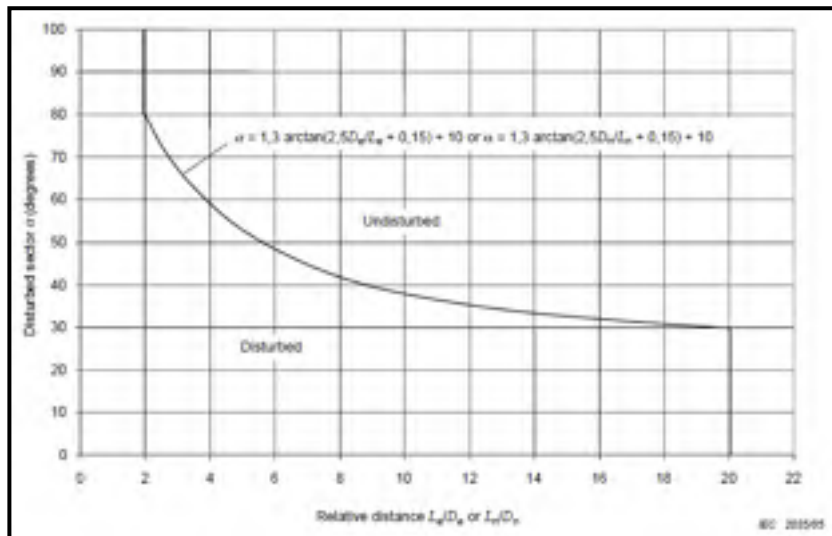


Figure 3.4 Limite du secteur obstrué par le sillage
en fonction de la distance relative
Tirée de IEC-61400-12-1 (2005)

La figure 3.4, tirée de la norme IEC-61400-12-1, présente la limite de la fonction qui représente le secteur obstrué en fonction de la distance relative (L_e/D_e). Ainsi, la valeur de la direction du vent incident au mât supérieure à cette limite est comprise dans le secteur non influencé; tandis que la valeur de la direction du vent inférieure à la limite évaluée est comprise dans le secteur obstrué.

3.1.2 Secteurs obstrués du mât W2 et impact sur la courbe de puissance

Le calcul des secteurs obstrués du mât W2 par les éoliennes voisines représentées à la figure 3.3 a été effectué en suivant la procédure décrite au paragraphe 3.3.1. Les secteurs de direction du vent affectés par le sillage sont présentés à la figure 3.5.

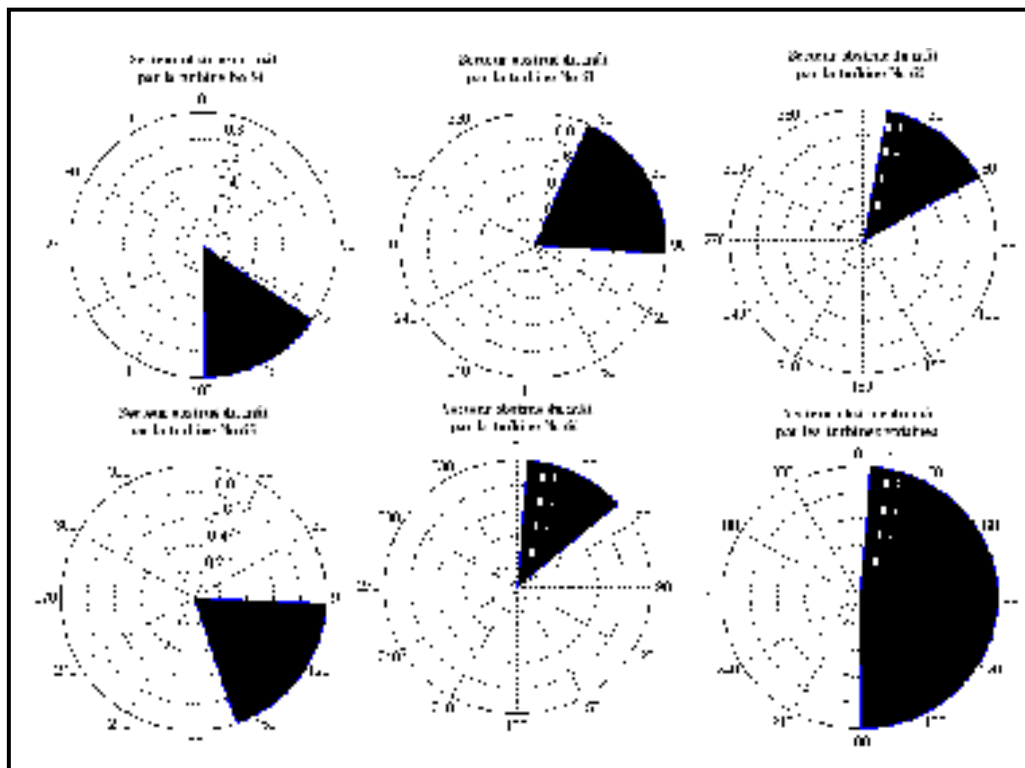


Figure 3.5 Secteurs obstrués par les turbines voisines au mât W2

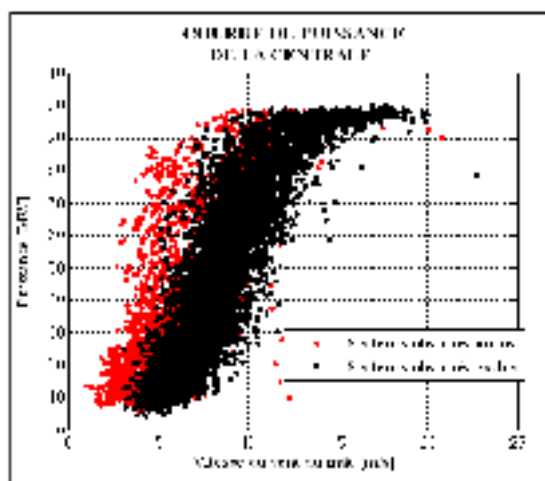


Figure 3.6 Effets de sillage sur la courbe de puissance de la centrale éolienne

La figure 3.6 montre les effets de sillage sur la courbe de puissance représentée en utilisant les vitesses du vent lues au mât de mesures W2 et la puissance totale de la centrale. En rouge (points pâles), sont représentés l'ensemble des données utilisées. Tandis qu'en noir (points foncés), il apparaît les données dont la direction du vent n'est pas comprise dans les secteurs obstrués. Le sillage crée une dispersion qui rend la courbe de puissance plus difficile à modéliser. Dans le but d'étudier l'impact de différentes méthodes et méthodologies servant à modéliser la courbe de puissance de la centrale, les données correspondant aux directions du vent comprises dans le sillage (entre 6 et 180 degrés) sont écartées. Ainsi, les effets de sillage ne pourront pas biaiser les conclusions. Il serait important de mentionner qu'après avoir exclu les données dans le sillage, il reste un nombre de données équivalent à une période de temps de 3 mois. Toutefois, les données issues des secteurs obstrués par le sillage seront reprises en compte lors de l'analyse des variables et l'évaluation de la courbe de puissance.

3.2 Comparaison de deux différentes méthodes

Plusieurs méthodes d'évaluation de la courbe de puissance ont été analysées et comparées par des études précédentes. Les performances de la méthode des Bins, qui est présentée par la norme IEC-61400-12-1, ont fait l'objet de comparaisons avec d'autres méthodes statistiques, telles que les méthodes de régression, les réseaux de neurones, les filtres de Kalman ou la logique floue pour ne citer que quelques-unes. La plupart du temps, les comparaisons sont effectuées en modifiant la méthodologie d'évaluation de la courbe de puissance. Ainsi l'influence du choix de la méthode est difficilement mise à l'évidence par les études précédentes.

La courbe de puissance d'une éolienne a été évaluée avec une méthode de régression polynomiale et les résultats obtenus ont été comparés avec la méthode des Bins (Lombart *et al.*, 2006). La méthode implantée a donné des résultats semblables à ceux obtenus avec la méthode des Bins, mais avec une légère diminution de la précision. Cependant, la méthode de régression polynomiale implantée utilise la vitesse du vent sans la normaliser au préalable, ce qui ne prend pas en compte l'influence de la masse volumique de l'air telle que prescrite

par la norme IEC. Li *et al.* (2010) ont évalué la courbe de puissance d'une éolienne avec une méthode semblable à la méthode des Bins : la méthode du maximum de vraisemblance (« maximum likelihood »). La méthode du maximum de vraisemblance consiste à regrouper les vitesses de vent par intervalles de 0.5 m/s, tout comme la méthode des Bins. Cependant, la puissance correspondante à chaque intervalle est évaluée différemment. Les détails de la méthode utilisée peuvent être consultés dans Li *et al.* (2010). Les résultats obtenus ne présentent pas d'amélioration en comparaison à la méthode des Bins, cependant, Li *et al.* (2010) estiment que la méthode peut produire de meilleurs résultats en utilisant beaucoup plus de données.

Siebert (2008) a comparé deux méthodes d'évaluation de la courbe de puissance d'une centrale éolienne (la méthode de régression linéaire et les réseaux de neurones combinés à la logique floue). La méthode de régression linéaire utilisée consiste à séparer les vitesses par intervalles réguliers (comme la méthode des Bins). Cependant, la relation entre la vitesse du vent et la puissance pour chacun des intervalles est approximée par une droite qui est évaluée par une régression linéaire. Le nombre d'intervalles de vitesse du vent et la distance entre les intervalles a fait l'objet d'une optimisation. Aussi, la méthode des réseaux de neurones a été combinée avec la méthode de la logique floue pour évaluer la courbe de puissance. Les deux modèles implantés ont été comparés sur trois centrales éoliennes différentes. Des résultats similaires ont été obtenus. Li (1999) a aussi modélisé la courbe de puissance de plusieurs éoliennes d'une centrale située à 20 miles à l'ouest de Ford Davis dans le Texas (USA); la méthode des régressions et les réseaux de neurones ont aussi fait l'objet d'une comparaison dans ces travaux. Li utilise la vitesse et la direction du vent lues par deux mâts météorologiques et les compare à la puissance produite par chaque éolienne. Les réseaux de neurones utilisés ont été entraînés par la méthode de rétro-propagation de l'erreur. Les résultats obtenus lui ont permis de conclure que les réseaux de neurones donnent de meilleurs résultats en comparaison à la méthode de régression linéaire. Plus loin, il a utilisé les filtres de Kalman pour accélérer l'entraînement de son réseau de neurones. L'utilisation des filtres de Kalman lui a alors permis de réduire le temps d'entraînement du réseau de neurones. Il en ressort de son étude que les réseaux de neurones procurent des résultats plus intéressants que

la méthode de régression linéaire et que l'apprentissage des réseaux peut être accéléré en utilisant les filtres de Kalman.

Bien que plusieurs études précédentes aient modélisé la courbe de puissance d'une turbine éolienne ou d'une centrale avec plusieurs approches différentes, la meilleure méthode à utiliser n'est pas claire. L'objectif de cette partie consiste à effectuer une étude sur l'impact du choix de la méthode sur la précision de la modélisation de la courbe de puissance d'une centrale éolienne. La norme IEC-61400-12-1 décrit de manière détaillée la méthodologie d'évaluation de la courbe de puissance d'une éolienne avec la méthode des Bins. Elle peut donc servir de méthode de référence. Les réseaux de neurones, qui sont considérés comme une méthode « avancée », seront aussi utilisés. Par conséquent, un réseau de neurones artificiel (RNA) de type perceptron multicouches (PMC) est entraîné ; les résultats des deux méthodes sont comparés. La figure 3.7 présente les données utilisées (hors sillage) pour bâtir et tester chacune des méthodes. Ces données ont été divisées de manière aléatoire en 80% pour modéliser la courbe de puissance et 20% pour la tester.

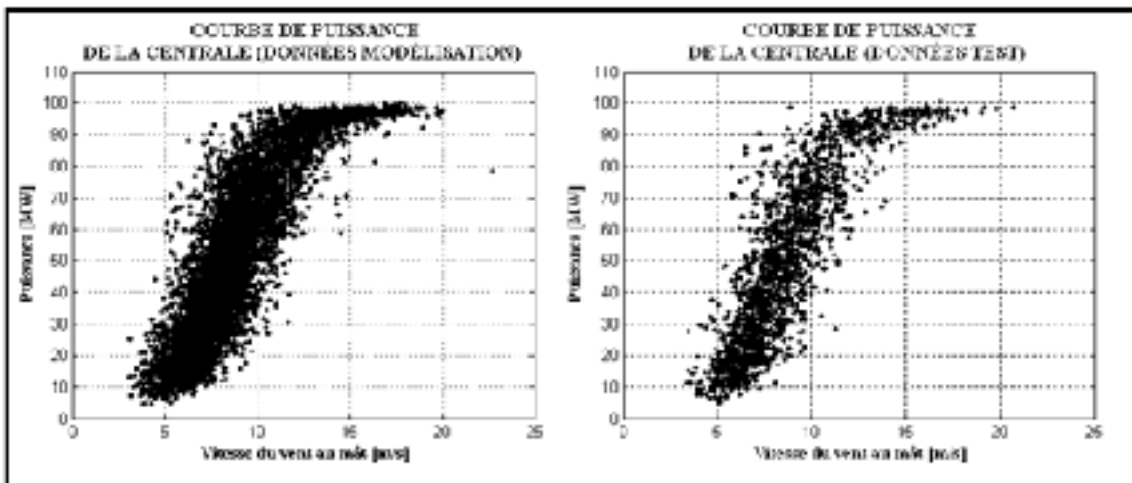


Figure 3.7 Données utilisées pour modéliser et tester la courbe de puissance de la centrale. L'ensemble des données sélectionnées a été divisé en deux groupes afin de s'assurer que les données utilisées pour modéliser la courbe de puissance soient différentes de celles utilisées pour la tester

La norme IEC-61400-12-1 présente la procédure d'évaluation de la courbe de puissance d'une éolienne avec la méthode des Bins. La procédure utilise les vitesses du vent et les puissances produites moyennées sur des intervalles de 10min. La première étape consiste à normaliser les vitesses du vent ou les puissances produites à l'aide de la masse volumique de l'air prise au niveau de la mer et servant de masse volumique de référence. Si les turbines sont à contrôle actif, les vitesses de vent sont normalisées. Si elles sont à contrôle passif, les puissances sont normalisées. Les formules utilisées pour normaliser les vitesses et les puissances sont présentées par les équations 3.3 et 3.4 :

$$V_n = V_{10min} \left(\frac{\rho_{10min}}{\rho_0} \right)^{1/3} \quad (3.3)$$

$$P_n = P_{10min} \frac{\rho_0}{\rho_{10min}} \quad (3.4)$$

où V_n est la vitesse du vent normalisée;
 V_{10min} est la vitesse du vent mesurée et moyennée sur une période de temps de 10 min;
 P_n est la puissance normalisée;
 P_{10min} est la puissance mesurée et moyennée sur une période de temps de 10 min;
 ρ_{10min} est la moyenne de la masse volumique de l'air sur une période de 10 min; et,
 ρ_0 est la masse volumique de référence (masse volumique de l'air au niveau de la mer).

La centrale éolienne étudiée est constituée des éoliennes à contrôle actif. Par conséquent, les vitesses de vent sont normalisées suivant l'équation 3.3 ; tandis que les puissances produites ne sont pas normalisées. La seconde étape consiste à regrouper les vitesses de vent normalisées suivant des intervalles de 0.5 m/s. Les puissances équivalentes aux mêmes

intervalles de temps sont aussi regroupées. L'évaluation de la moyenne des vitesses (Équation 3.5) et la moyenne des puissances (Équation 3.6) de chaque intervalle permet d'associer une puissance à une vitesse de vent.

$$V_i = \frac{1}{N_i} \sum_{j=1}^{N_i} V_{n,i,j} \quad (3.5)$$

$$P_i = \frac{1}{N_i} \sum_{j=1}^{N_i} P_{i,j} \quad (3.6)$$

- où V_i est la vitesse du vent correspondant à l'intervalle i ;
 P_i est la puissance correspondante à l'intervalle i ;
 N_i est le nombre de points contenus dans l'intervalle i ;
 $V_{n,i,j}$ est la vitesse normalisée du point j dans l'intervalle i ; et,
 $P_{i,j}$ est la puissance du point j dans l'intervalle i .

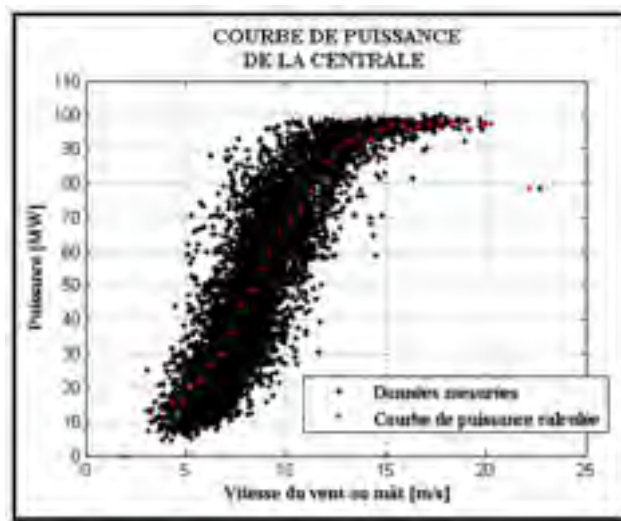


Figure 3.8 Courbe de puissance calculée par la méthode des Bins

La courbe de puissance calculée est testée en modélisant la puissance avec les données servant de « tests » présentées à la figure 3.7. Les performances sont évaluées en fonction de trois critères principaux qui sont : l'erreur absolue moyenne normalisée (NMAE), l'erreur quadratique moyenne normalisée (NRMSE) et le coefficient de régression (R^2). Les valeurs de ces critères sont présentées dans le tableau 3.1. La figure 3.9 permet aussi d'apprécier la précision de la modélisation.

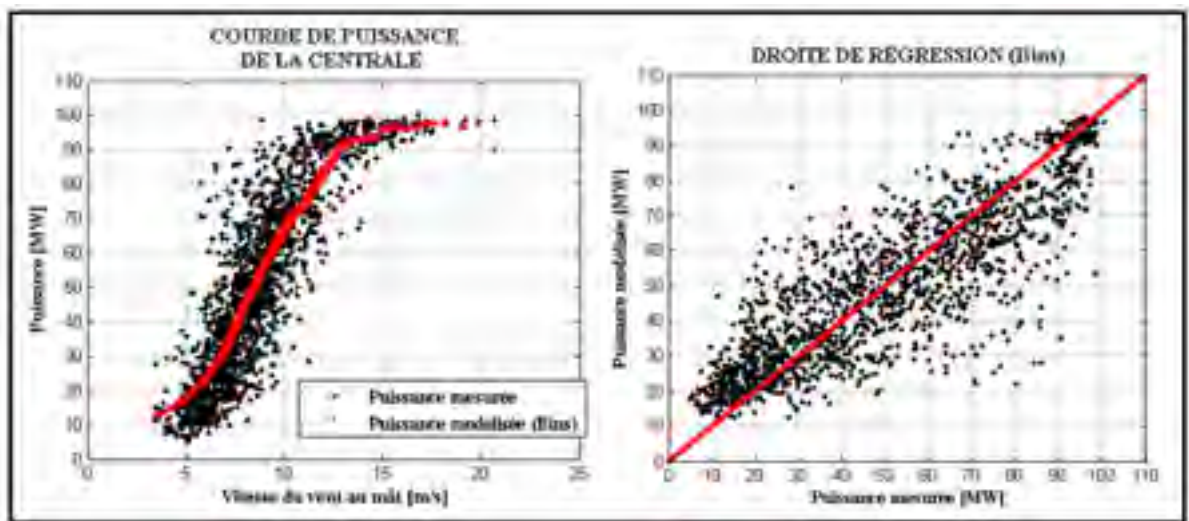


Figure 3.9 Comparaison entre la puissance mesurée et la puissance modélisée avec la méthode des Bins (courbe de puissance et droite de régression)

Tableau 3.1 Performances de la courbe de puissance (Méthode des Bins)

Critère d'évaluation	Score (%)
NMAE	9.8951
NRMSE	13.2213
R^2	0.9427

3.2.2 Les réseaux de neurones

Les réseaux de neurones artificiels (RNA) sont des outils d'apprentissage qui se basent sur des informations contenues dans un ensemble de données et qui sont capables d'extrapoler à un ensemble de données n'ayant pas servi à l'apprentissage. L'évaluation de la courbe de puissance de la centrale éolienne peut être identifiée comme un problème d'approximation de fonction. Pour ce type de problèmes, les réseaux de neurones peuvent être entraînés sur des données mesurées servant à l'apprentissage, afin de capturer la relation existant entre les variables météorologiques et la puissance produite de la centrale éolienne. Le réseau de neurones entraîné pourra donc être évalué en lui soumettant un ensemble de données test n'ayant pas servi à l'apprentissage, et en comparant la réponse du réseau à la puissance produite mesurée de la centrale éolienne.

Les réseaux de neurones sont utilisés dans plusieurs domaines. Les applications les plus utilisées couvrent la reconnaissance de forme, la classification et l'approximation des fonctions. Implanter un réseau de neurones pour la résolution d'un problème consiste à définir les paramètres du réseau qui faciliteront l'apprentissage, ce qui permettra de converger vers la solution optimale. Un RNA de type PMC à une couche cachée, dont les fonctions d'activation sont de type sigmoïde, et à une couche de sortie, dont les fonctions d'activation sont de type linéaire, peut être utilisé pour approximer une fonction ayant un nombre infini de discontinuité (Demuth *et al.*, 2008). Il peut donc être utilisé pour approximer la relation entre les variables météorologiques et la puissance produite de la centrale éolienne. Un exemple est montré à la figure 3.10.

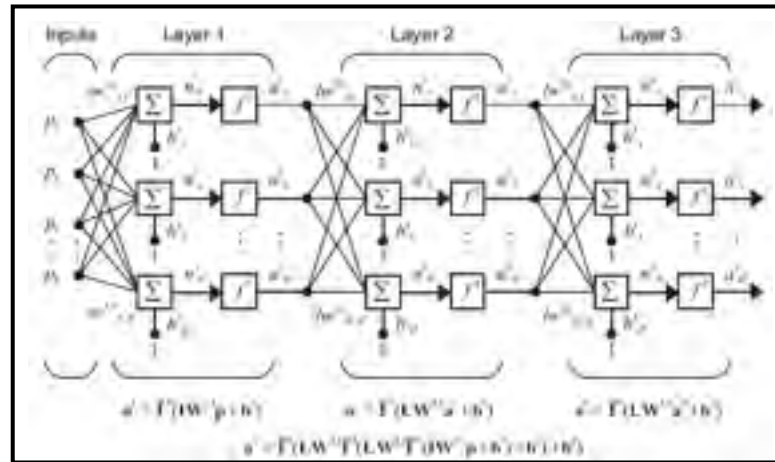


Figure 3.10 Exemple de PMC à 3 couches cachées
Tirée de Demuth *et al.* (2008)

Le RNA est constitué de plusieurs neurones étroitement liés qui traitent l'information. Chaque neurone qui reçoit un signal effectue une somme pondérée des signaux reçus et restitue l'information à travers une fonction d'activation tel que représenté à la figure 3.11.

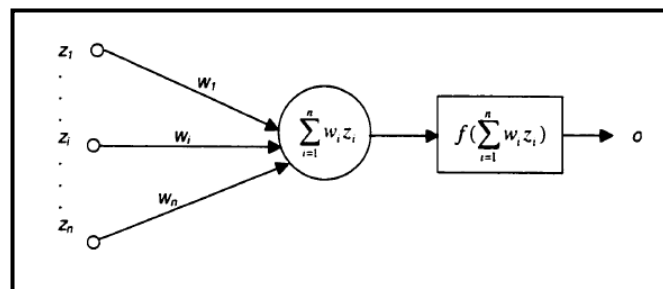


Figure 3.11 Le neurone artificiel
Tirée de Benahmed (2002)

$$o = f\left(\sum_{i=1}^n w_i z_i\right) \quad (3.7)$$

où Z_i est un signal incident au neurone;

- w_i est la valeur du poids de l'entrée du neurone;
- f est la fonction d'activation; et,
- o est valeur du signal de sortie du neurone.

Pour implanter le RNA, il est indispensable de choisir les paramètres tels que les entrées et les sorties du réseau, le nombre de couches cachées, le nombre de neurones par couche cachée, la constante d'apprentissage, les fonctions d'activation et la technique d'apprentissage.

3.2.2.1 Les entrées du réseau

Le PMC a pour but d'évaluer la relation existante entre les variables météorologiques et la puissance produite de la centrale éolienne. L'utilisation des réseaux de neurones permet de comparer la précision d'une méthode avancée (RNA) à celle de la méthode de base (méthode des Bins) en utilisant une même méthodologie. Les variables d'entrées du réseau sont donc les mêmes paramètres qui ont été utilisés pour évaluer la courbe de puissance avec la méthode des Bins. La vitesse du vent lue au mât W2 et la masse volumique de l'air constituent les entrées du RNA.

3.2.2.2 Les sorties du réseau de neurones

Le PMC permet d'évaluer les relations entre les entrées et les sorties du réseau de neurones. La sortie du PMC sera donc la puissance produite de la centrale. Il sera ainsi possible de modéliser la relation entre les variables météorologiques (vitesse du vent et masse volumique de l'air) et la puissance produite de la centrale éolienne.

3.2.2.3 Le choix des fonctions d'activation

Les RNA de type PMC transportent l'information par des signaux à partir de l'entrée vers la sortie du réseau, d'où le nom de « feedforward neural networks ». Les informations reçues à

l'entrée du réseau se propagent à travers les neurones interconnectés jusqu'à atteindre la sortie du réseau. Chaque neurone recevant un signal le transmet aux autres neurones avec qui il est lié suivant l'équation 3.7. Les poids w_i sont déterminés par un algorithme d'apprentissage. Le guide d'utilisation de la boîte d'outils de Matlab pour les réseaux de neurones (Demuth *et al.*, 2008) présente trois principales fonctions d'activation qui permettent aux neurones de convertir les signaux émis en réponse du système. La fonction d'activation de type sigmoïde est très souvent utilisée dans les algorithmes utilisant l'apprentissage par rétro-propagation des erreurs. L'avantage est qu'elle est facilement différentiable. La fonction d'activation de type sigmoïde est donc utilisée pour convertir les signaux sur chaque neurone des couches cachées. La fonction d'activation linéaire utilisée sur la couche de sortie permet d'effectuer une simple sommation des signaux reçus par le neurone de la couche de sortie afin de fournir la réponse du RNA.

3.2.2.4 Le nombre de couches cachées et le nombre de neurones par couche cachée

Le réseau de neurones implanté est à trois couches, soit composé d'une couche d'entrée qui reçoit les entrées du réseau, une couche cachée qui permet de traiter l'information et une couche de sortie qui fournit la réponse du réseau. Le nombre de neurones de la couche d'entrée est égale au nombre d'entrées de RNA qui est de deux pour le cas étudié (vitesse du vent et masse volumique de l'air). Le nombre de neurones sur la couche de sortie du réseau est égale au nombre de sorties du RNA et qui est de un (la puissance produite). Il existe des heuristiques permettant de déterminer le nombre de neurones sur une couche cachée (Benahmed, 2002). Un exemple tiré du mémoire de Benahmed (2002) est présenté à l'équation 3.8. Cependant, la détermination peut aussi se faire de manière expérimentale. La procédure consiste à fixer tous les autres paramètres du RNA et de réaliser plusieurs expériences en faisant varier le nombre de neurones sur la couche cachée. Pour chacun des cas, le réseau est appris 30 fois avec les données servant à modéliser, et la médiane de l'erreur du modèle est évaluée pour chaque configuration avec les données servant aux tests. La comparaison des scores obtenus pour chaque nombre de neurones sur la couche cachée permet de déterminer le nombre de neurones adéquat.

$$N_c = \left(\frac{N_e + N_s}{2} \right) + 1 \quad (3.8)$$

où : N_e est le nombre de neurones sur la couche d'entrée;
 N_c est le nombre de neurones sur la couche cachée; et,
 N_s est le nombre de neurones sur la couche de sortie.

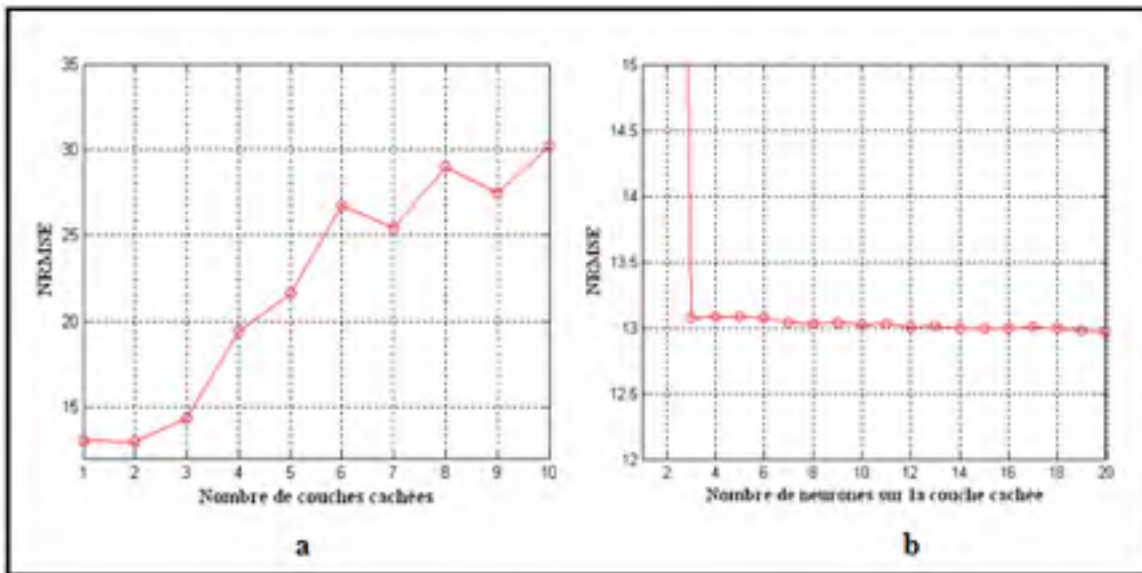


Figure 3.12 Performance du PMC en fonction du nombre de couches cachées et en fonction du nombre de neurones sur la couche cachée

La figure 3.12a montre l'erreur du réseau en fonction du nombre de couches cachées et la figure 3.12b montre l'évolution de l'erreur en fonction du nombre de neurones sur la couche cachée. Cette erreur est évaluée sur l'ensemble de données servant aux tests, en modélisant la puissance sur 30 répliques et pour chaque configuration, et en observant la médiane de l'erreur sur les 30 répliques. La figure 3.12 montre que plus le nombre de couches cachées est grand, plus le RNA a la capacité d'approximer les fonctions plus complexes. Cependant, l'ajout de plusieurs couches cachées avec un nombre de données limité peut conduire à la dégradation des performances du RNA, car dans ce cas, le RNA aura besoin de plus de

données pour réussir son apprentissage. Les résultats présentés à la figure 3.12a confirment bien le fait que l'« approximateur universel » composé d'une couche cachée est capable d'approximer les fonctions complexes. À la figure 3.12b, on remarque que l'erreur diminue en fonction du nombre de neurones sur la couche cachée. A partir de trois neurones sur la couche cachée, le RNA atteint une précision qui varie très peu en rajoutant des neurones additionnels. Il n'a pas été observé une différence de temps de calcul considérable entre les simulations à trois neurones sur la couche cachée et celles à huit neurones. Cependant, en choisissant huit neurones sur la couche cachée, l'on est capable de gagner environ 0.1% de NRMSE. N'ayant pas observé une contrainte de temps de calcul, il a donc été choisi huit neurones sur la couche cachée pour modéliser la courbe de puissance.

3.2.2.5 La technique d'apprentissage

Le nombre de couches cachées est fixé à un avec huit neurones sur la couche cachée, la technique d'apprentissage est celle par rétro-propagation des erreurs. Les algorithmes d'apprentissage permettent d'établir la valeur des poids synoptiques w_i au fur et à mesure que les exemples sont présentés au réseau pendant l'apprentissage. Ces algorithmes sont basés sur la réduction de l'erreur en ajustant les poids à chaque présentation des vecteurs d'entraînement. Benahmed (2002) présente de manière concise les différentes étapes de l'apprentissage par la technique de retro-propagation avec minimisation du gradient d'erreur qui a été utilisé.

3.2.2.6 Les résultats obtenus avec les réseaux de neurones

Le réseau de neurones implanté, schématisé à la figure 3.13, a réussi à simuler la relation entre les variables météorologiques (vitesse du vent lue au mât W2 et la masse volumique de l'air) et la production de puissance de la centrale éolienne. L'apprentissage a été effectué par la technique de rétro-propagation de l'erreur. Les résultats obtenus sont présentés à la figure 3.14 et le tableau 3.2.

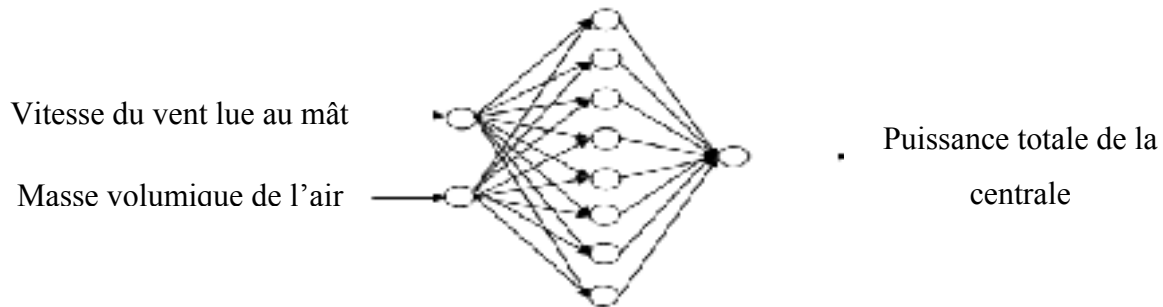


Figure 3.13 Architecture du réseau de neurones implémenté

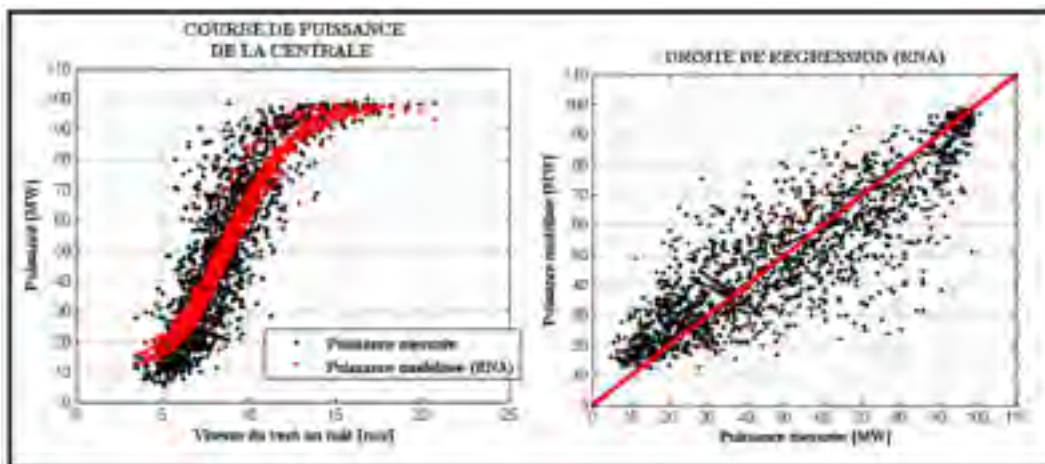


Figure 3.14 Comparaison entre la puissance mesurée et la puissance modélisée avec les RNA (courbe de puissance et droite de régression)

Tableau 3.2 Erreurs du modèle (RNA)

Critère d'évaluation	Score (%)
NMAE	9.6793
NRMSE	12.9955
R^2	0.9486

3.2.3 Comparaison de la méthode des Bins et des réseaux de neurones

La méthode des Bins et les réseaux de neurones ont été utilisés pour modéliser la courbe de puissance de la centrale éolienne étudiée. Pour les deux différentes méthodes, les mêmes données ont été utilisées pour les implanter ainsi que pour les tester. Les résultats obtenus, en modélisant la courbe de puissance de la centrale éolienne avec ces deux méthodes respectives, sont présentés à la figure 3.15 et au tableau 3.3. En comparant une méthode de base (la méthode des Bins) à une méthode avancée (RNA) et en conservant les mêmes paramètres d'entrée et de sortie, il est donc possible d'apprécier l'influence du choix de la méthode sur la modélisation de la courbe de puissance de la centrale éolienne.

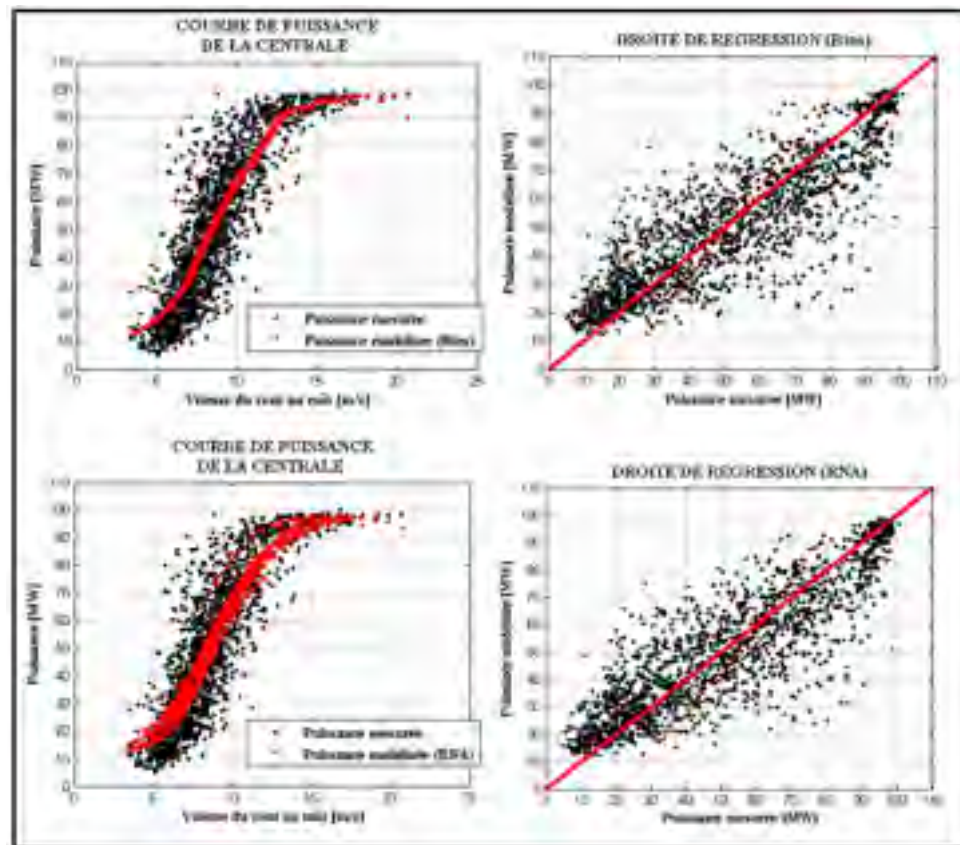


Figure 3.15 Comparaison des résultats (méthode des Bins vs RNA)

La comparaison entre la méthode des Bins et les RNA a eu pour principal but d'étudier l'impact du choix de la méthode sur la précision de la modélisation de la courbe de puissance de la centrale éolienne. Tel qu'il a été mentionné à l'introduction de la section 3.2, plusieurs études précédentes ont utilisé et comparé plusieurs méthodes différentes. Cependant, il reste très difficile d'identifier la meilleure approche en consultant la littérature; car la plupart du temps, les méthodes sont comparés soit en utilisant des variables différentes, soit en adoptant des méthodologies différentes. Dans un tel contexte, il peut être difficile de tirer de bonnes conclusions sur l'influence de la méthode, car il est possible que le gain en précision soit obtenu grâce à l'adoption d'une méthode différente, ou à l'ajout de nouveaux paramètres météorologiques. Dans cette section, il a été question de comparer deux méthodes dont l'une est identifiée comme simple par la littérature, et l'autre avancée. En comparant donc une méthode de base (la méthode des Bins) à une méthode avancée (RNA) tout en conservant les mêmes paramètres d'entrée et de sortie et en adoptant la même méthodologie, il est donc possible d'apprécier l'influence du choix de la méthode sur la modélisation de la courbe de puissance de la centrale éolienne.

Tableau 3.3 Comparaison des performances de deux méthodes : Bins vs RNA

Méthode	NMAE (%)	NRMSE (%)	R²
Méthode des Bins	9.8951	13.2213	0.9427
RNA	9.6793	12.9955	0.9486
Amélioration	<i>2.18</i>	<i>1.71</i>	<i>0.0059</i>

En ayant modélisé la courbe de puissance de la centrale avec un RNA ayant deux entrées (la vitesse du vent au mât et la masse volumique de l'air), il a été observé une légère amélioration de la précision, soit de 2.18% du NMAE, et 1.71% du RMSE. Les résultats obtenus montrent que les réseaux de neurones permettent de modéliser la courbe de puissance de la centrale avec une précision légèrement supérieure à celle obtenue avec la méthode des Bins en utilisant les mêmes variables. Cependant, les réseaux de neurones offrent un avantage incontestable : la possibilité de prendre en compte plusieurs variables en entrées, même dans un contexte où le nombre de données est limité. Ils pourront donc servir

à tester l'influence d'autres paramètres météorologiques sur la courbe de puissance de la centrale éolienne.

3.3 Comparaison de deux différentes méthodologies

La section précédente a permis de comparer la méthode des Bins et les réseaux de neurones pour l'évaluation de la courbe de puissance de la centrale éolienne. L'objectif de cette partie consiste à comparer deux différentes méthodologies pour la résolution du problème. Cabezon *et al.* (2004) ont montré qu'il est possible de modéliser la courbe de puissance d'une centrale en utilisant la moyenne des vitesses aux éoliennes. Cependant, le nombre de données qu'ils ont utilisé pour tester la méthode équivalait à deux semaines de données, ce qui pourrait ne pas être suffisant pour tirer des conclusions valables. L'objectif de cette section consiste à évaluer la courbe de puissance de la centrale éolienne étudiée en comparant la moyenne des vitesses aux éoliennes à la puissance totale de la centrale. La méthodologie proposée par Cabezon *et al.* (2004) est reprise, mais avec une quantité de données plus élevée. Une année complète de données est utilisée, soit 80% permettant de modéliser et 20% permettant de tester (parmi les données sélectionnées après la procédure de contrôle de la qualité, le retrait des données dans le sillage et le retrait des données correspondant aux arrêts des turbines). Ces données sont présentées à la figure 3.16.

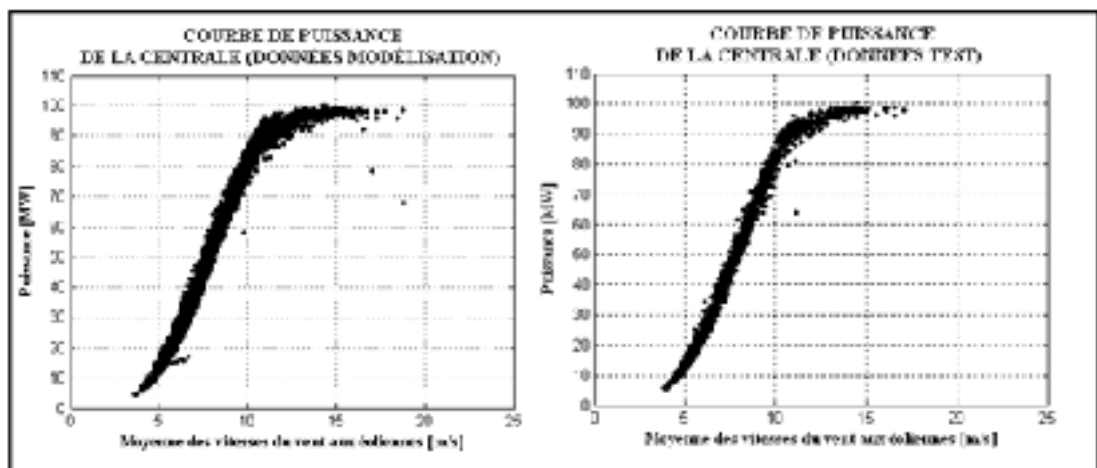


Figure 3.16 Courbe de puissance de la centrale avec la moyenne des vitesses aux éoliennes (données modélisation et données test)

La figure 3.17 compare deux courbes de puissance de la centrale éolienne étudiée. Les points en rouges (points pâles) représentent la courbe de puissance de la centrale tracée en comparant la moyenne des vitesses aux éoliennes et la puissance totale de la centrale éolienne. Pour leur part, les points noirs (points foncés) représentent la courbe de puissance de la centrale, en comparant la vitesse lue au mât et la puissance totale de la centrale éolienne. Il serait important de mentionner que les deux courbes ont exactement le même nombre de points, et sont mesurées dans les mêmes périodes de temps. La superposition de ces deux courbes de puissance montre que l'approche avec la moyenne des vitesses aux éoliennes présente moins de dispersion dans la relation entre la vitesse du vent et la puissance. Par conséquent, il serait beaucoup plus facile de modéliser la courbe de puissance de la centrale en évaluant la moyenne des vitesses du vent aux éoliennes. Cette hypothèse est vérifiée en calculant la courbe de puissance avec la moyenne des vitesses du vent aux nacelles, et en comparant les résultats obtenus avec la courbe de puissance précédemment calculée avec la méthode des Bins, et en considérant la vitesse du vent lue au mât. Les détails du calcul de la courbe de puissance de la centrale éolienne en utilisant la méthode des Bins sont présentés à la section 3.2.1. Les comparaisons entre la puissance modélisée et la puissance mesurée avec les deux méthodes respectives sont présentée à la figure 3.18, tandis que le tableau 3.4 montre les valeurs des incertitudes.

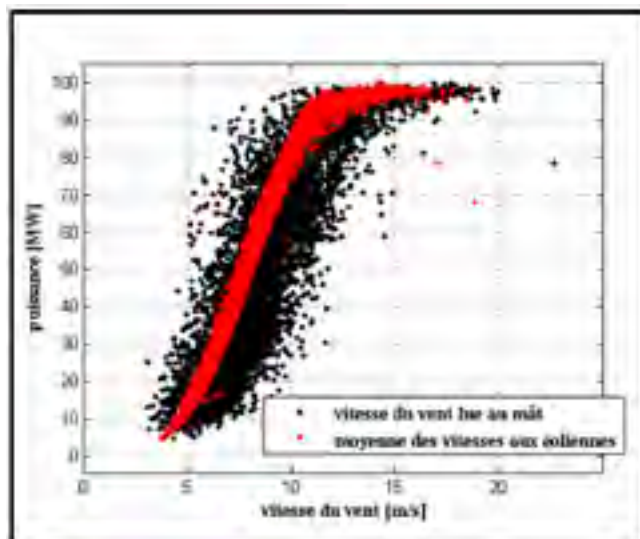


Figure 3.17 Comparaison des courbes de puissance de la centrale (vitesse lue au mât vs moyenne des vitesses aux éoliennes)

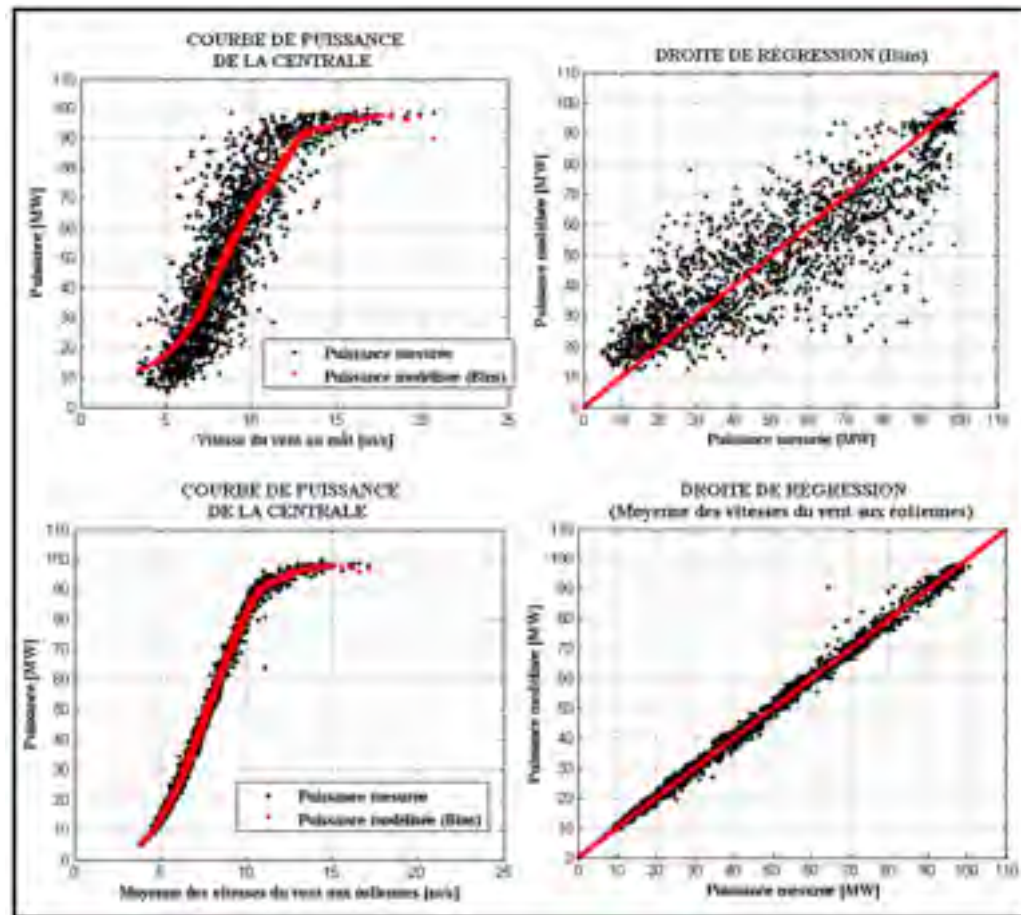


Figure 3.18 Comparaison des résultats de deux différentes méthodologies (vitesse lue au mât vs moyenne des vitesses aux éoliennes)

Tableau 3.4 Comparaison des performances de deux méthodologies : vitesse au mât vs moyenne des vitesses aux éoliennes

Méthodologie	NMAE (%)	NRMSE (%)	R ²
Vitesse lue au mât	9.8951	13.2213	0.9427
Moyenne des vitesses aux éoliennes	1.4110	1.9420	0.9984
Amélioration	85.74	85.31	0.0557

Les résultats obtenus confirment qu'il est beaucoup plus facile de modéliser la courbe de puissance de la centrale en utilisant la moyenne des vitesses aux éoliennes. En comparant le gain obtenu en modélisant la courbe de puissance avec les RNA au paragraphe 3.2.3, il s'avère que la méthode de base (Bins), en utilisant la moyenne des vitesses du vent lues aux

éoliennes, procure un résultat beaucoup plus précis que celui obtenu avec les RNA en utilisant la vitesse du vent lue au mât. Ces résultats mettent à l'évidence l'impact de la méthodologie adoptée pour modéliser la courbe de puissance. Bien que cette dernière courbe de puissance ait été modélisée avec la méthode des Bins, elle procure une amélioration de 85% sur la précision, ce qui est nettement supérieure à celle obtenue en modélisant la courbe de puissance avec les réseaux de neurones et en adoptant la première méthodologie. Cela pourrait s'expliquer par le fait que la vitesse du vent lue en un point ne représente pas toujours les conditions météorologiques sur l'ensemble de la centrale éolienne. Il peut arriver par exemple que la vitesse du vent en un point précis de la centrale soit sous l'influence du terrain. Il pourra donc arriver que la vitesse en ce point de la centrale soit plus forte, ou plus faible que celles en tous les autres points de la centrale. En calculant la courbe de puissance tout en utilisant la vitesse du vent en ce point, l'on pourrait donc sous-estimer ou surestimer la puissance de la centrale selon que la vitesse du vent en ce point soit en dessous, ou au-dessus de la vitesse du vent moyenne sur l'ensemble de la centrale éolienne.

3.4 Choix des paramètres

Cette section a pour but de déterminer les variables météorologiques qui doivent être utilisées pour calculer la courbe de puissance de la centrale éolienne avec une bonne précision. En se basant sur les connaissances physiques du processus de conversion du vent en puissance, et en se basant aussi sur des études précédentes, l'on peut être capable de sélectionner des variables météorologiques susceptibles d'influencer la puissance produite de la centrale éolienne et de les analyser en menant une série d'expériences avec les RNA. Les éoliennes produisent de l'électricité en convertissant l'énergie cinétique contenue dans le vent. L'équation 3.9 (Manwell *et al.*, 2009) montre que cette énergie cinétique dépend principalement de deux paramètres : la vitesse du vent et la masse volumique de l'air. Les effets de sillage présentés au paragraphe 3.1 varient en fonction de la direction du vent. Par conséquent, la vitesse du vent, la masse volumique de l'air et la direction du vent constituent trois variables à prendre absolument en compte lors de l'évaluation de la courbe de puissance de la centrale éolienne.

$$P = \frac{1}{2} \rho A U^3 \quad (3.9)$$

où P est l'énergie cinétique du vent par unité de temps;
 A est la surface balayée par les pâles éoliennes;
 ρ est la masse volumique de l'air; et,
 U est la vitesse du vent.

Plusieurs études ont aussi montré l'influence du cisaillement et de l'intensité de turbulence sur la courbe de puissance d'une éolienne individuelle (Hofman, 2007; Montes *et al.*, 2009; Wagner *et al.*, 2010). L'objectif de cette partie est de faire une analyse des variables à prendre en compte pour modéliser la courbe de puissance de la centrale éolienne. Les RNA offrent l'avantage d'ajouter facilement des variables à l'entrée du modèle; par conséquent, ils sont utilisés dans cette section. La première modélisation se fait en comparant la moyenne des vitesses aux éoliennes et la masse volumique de l'air à la puissance totale de la centrale. L'erreur obtenue en modélisant la courbe de puissance de la centrale dans ces conditions est présentée à la figure 3.19. L'intensité de turbulence et le cisaillement sont successivement ajoutés afin d'évaluer l'impact de chaque variable sur la précision du modèle. L'intensité de turbulence et le cisaillement ont été calculés au chapitre 2. Les études sont faites en conservant les données issues dans les secteurs de sillage identifiés au paragraphe 3.1 car l'ajout de la direction du vent pourrait permettre d'en tenir compte. Les résultats obtenus permettent de soutenir le choix des variables à utiliser pour modéliser la courbe de puissance de la centrale éolienne; ils sont présentés à la figure 3.20.

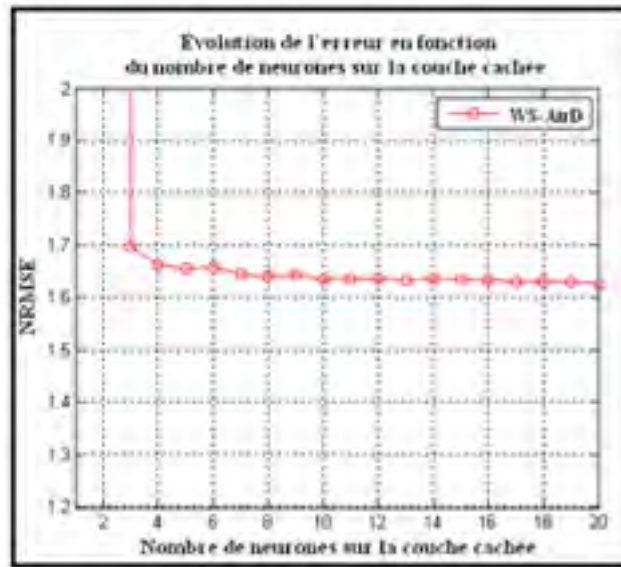


Figure 3.19 Performances du PMC avec deux variables (vitesse du vent et masse volumique de l'air)

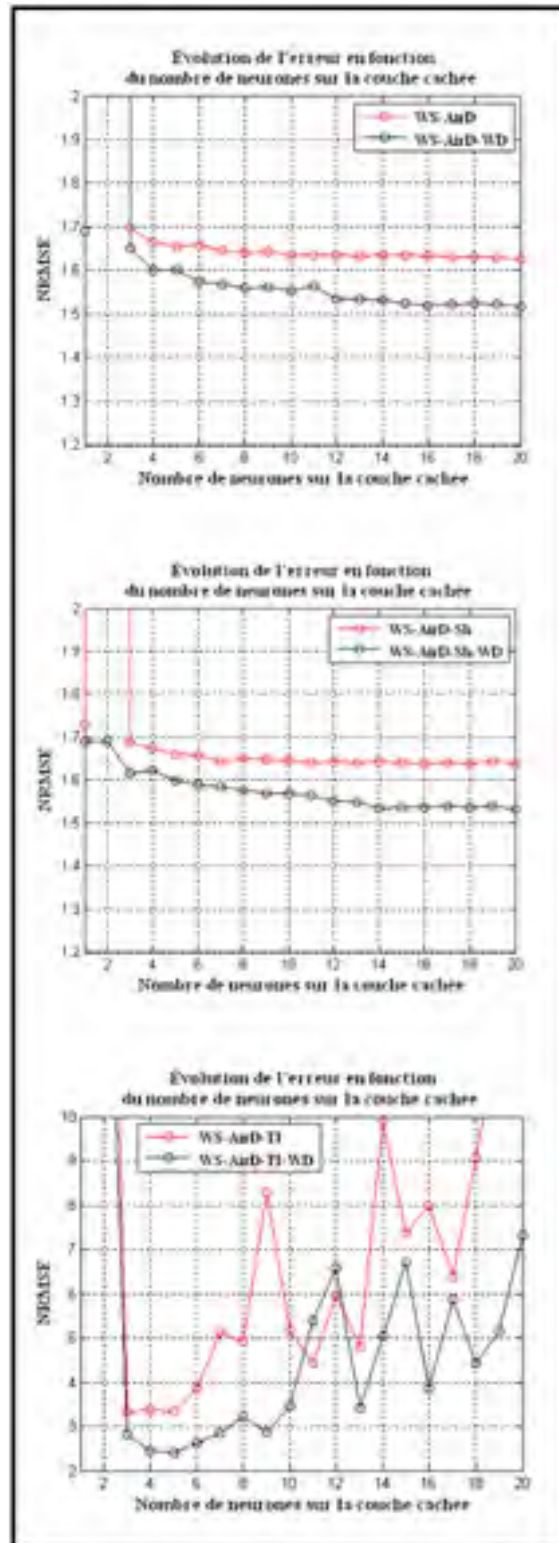


Figure 3.20 Performances du PMC en fonction de plusieurs variables (WS=vitesse du vent; AirD=masse volumique de l'air; WD=direction du vent; Sh=cisaillement; Ti=intensité de turbulence)

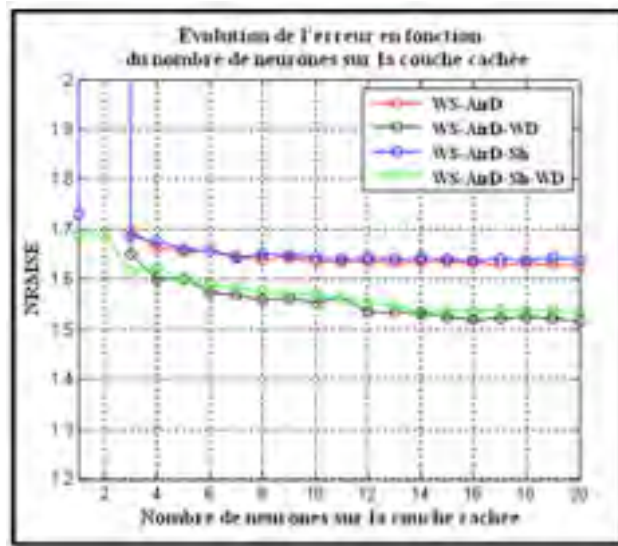


Figure 3.21 Comparaison des performances du PMC en fonction de plusieurs variables météorologiques

Les résultats de l'analyse des variables, résumés par la figure 3.21, montrent que la direction du vent est un paramètre important à utiliser pour modéliser la courbe de puissance de la centrale éolienne. Elle permet chaque fois de diminuer l'erreur. Par contre, la modélisation du réseau de neurones en introduisant l'intensité de turbulence a conduit à une mauvaise généralisation des données. Le cisaillement quant à lui ne montre pas une influence significative sur le phénomène étudié. Les résultats obtenus permettent donc de retenir trois variables : la vitesse du vent, la masse volumique de l'air et la direction du vent.

3.5 Courbe de puissance proposée et analyse de l'incertitude

Au paragraphe 3.2, il a été comparé deux méthodes pour évaluer de la courbe de puissance de la centrale éolienne. Les RNA ont permis d'obtenir un résultat légèrement supérieur à celui obtenu avec la méthode des Bins. La comparaison de deux différentes méthodologies au paragraphe 3.3 montre que l'évaluation de la courbe de puissance de la centrale en utilisant la moyenne des vitesses aux éoliennes permet d'obtenir une précision bien plus élevée. Les tests ont aussi été effectués sur différentes variables, à savoir l'intensité de turbulence, la direction du vent et le cisaillement. Les résultats ont conduit à sélectionner la vitesse de vent,

sa direction et la masse volumique de l'air comme paramètres météorologiques influençant la production de la puissance de la centrale éolienne au complet. Ces expériences peuvent mener à proposer une courbe de puissance pour la centrale éolienne étudiée. La courbe de puissance proposée est donc modélisée avec un RNA de type PMC. Trois paramètres météorologiques sont pris en compte : la moyenne des vitesses du vent aux éoliennes, la direction du vent au mât et la masse volumique de l'air au mât. Une étude de l'architecture similaire à celle présentée au paragraphe 3.2.2.4 a permis de choisir une topologie à 12 neurones sur la couche cachée. La courbe de puissance modélisée et la droite de régression sont présentées à la figure 3.22.

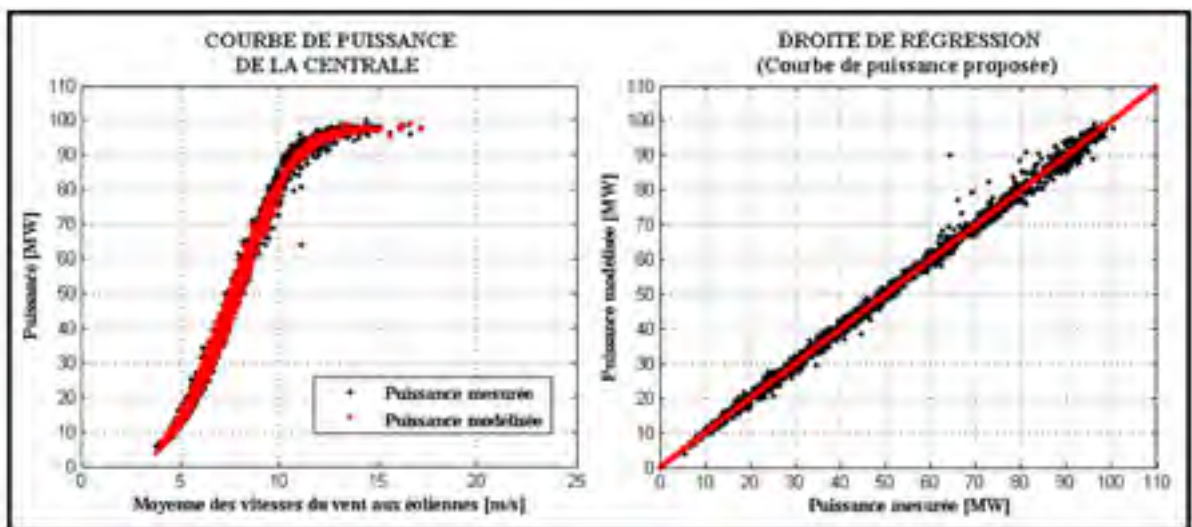


Figure 3.22 Comparaison entre la puissance mesurée et la puissance modélisée par la courbe de puissance proposée

Une analyse de sensibilité de la courbe de puissance proposée est effectuée afin de s'assurer que cette courbe de puissance peut être utilisée à des fins de prévision de la production éolienne. Pour ce faire, on injecte une perturbation de la vitesse du vent de 25% selon l'équation 3.10, et on évalue l'erreur en modélisation de production calculée par la courbe de puissance proposée. Les résultats obtenus sont présentés dans le tableau 3.5.

$$V_p = V + 2(0.5 - x) * 0.25V \quad (3.10)$$

où V est la vitesse du vent mesurée;
 V_p est la vitesse du vent perturbée; et,
 ϵ est un nombre aléatoire tel que $0 < \epsilon < 1$.

Tableau 3.5 Performances de la courbe de puissance proposée

Méthodologie	NMAE (%)	NRMSE (%)	R ²
Courbe de puissance proposée (CP)	1.0360	1.5402	0.9990
CP perturbée	10.7724	13.3554	0.9739

En apportant une perturbation de 25% à la vitesse de vent, la courbe de puissance proposée est capable de modéliser la puissance produite de la centrale éolienne avec une erreur quadratique (NRMSE) moyenne de 13%. Cela signifie que la courbe de puissance qui a été proposée est capable de convertir les prévisions météorologiques ayant une erreur de 25% sur les prévisions de la vitesse du vent; toutefois, l'on devrait s'attendre dans ces conditions de prédire la puissance avec une incertitude équivalente à peu près à 13% de NRMSE. Cette erreur est de l'ordre des erreurs des modèles actuels de prévision de la production éolienne. La courbe de puissance développée peut donc être utilisée à des fins de prévision de la production éolienne. L'objectif du prochain chapitre consistera à utiliser cette courbe de puissance pour convertir les prévisions météorologiques calculées par GEM-LAM en prévisions de la puissance de la centrale éolienne expérimentale.

CHAPITRE 4

DÉVELOPPEMENT DU MODÈLE DE PRÉVISION

4.1 Description et architecture du modèle

Les modèles physiques de prévision de la production éolienne utilisent les prévisions météorologiques évaluées par les modèles de prévision numérique pour prédire la production des centrales éoliennes. La courbe de puissance de la centrale permet de convertir les prévisions météorologiques en prévisions de puissance. Le modèle de prévision développé dans cette étude utilise les prévisions météorologiques de GEM-LAM 2.5 km fournies par Environnement Canada, et la courbe de puissance, et est capable de prédire la puissance produite d'une centrale éolienne jusqu'à quarante-huit heures à l'avance; c'est donc un modèle physique de prévision de la production éolienne. Les étapes nécessaires pour effectuer la prévision de la production de la centrale éolienne sont présentées à la figure 4.1.

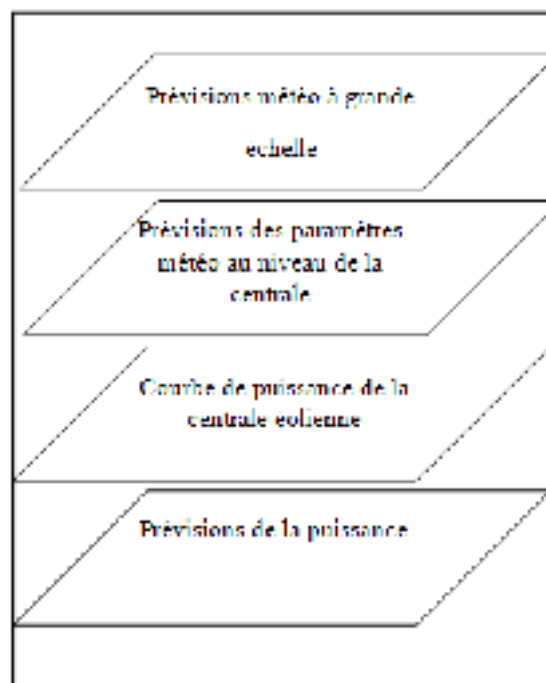


Figure 4.1 Les différentes étapes du modèle de prévision

La première étape consiste à acquérir les prévisions météorologiques calculées par GEM-LAM. Les modèles méso-échelles de prévisions météorologiques donnent les prévisions sur un maillage pouvant aller de quelques kilomètres jusqu'à cinquante kilomètres de résolution (Pinson, 2006). La deuxième étape consiste à se servir des sorties du modèle numérique pour évaluer les prévisions météorologiques au niveau de la centrale éolienne. Un module statistique de traitement des sorties, souvent appelé MOS (Model Output Statistics), est capable d'effectuer cette tâche en corrigeant davantage les erreurs systématiques du modèle. La courbe de puissance de la centrale, qui est la relation entre les variables météorologiques et la production de puissance de la centrale, permet alors de convertir les prévisions météo en prévisions de puissance de la centrale éolienne.

L'objectif de ce chapitre est d'utiliser la courbe de puissance développée au chapitre 3 pour convertir les prévisions météorologiques calculées par GEM-LAM en prévisions de la puissance produite de la centrale éolienne étudiée. Environnement Canada effectue des prévisions météorologiques expérimentales sur un domaine couvrant la côte Est du Canada grâce au modèle de prévision numérique GEM-LAM à 2.5 km de résolution. La centrale éolienne expérimentale est couverte par le domaine de prévision de GEM-LAM. Une année complète (l'année 2009) des données de prévisions météorologiques est disponible ainsi que les données mesurées de vent et de puissance de la centrale éolienne. Tout comme le protocole standard d'évaluation des modèles de prévision de la production éolienne proposé dans le cadre du projet ANEMOS (Madsen *et al.*, 2004), les données utilisées pour bâtir le modèle couvrent neuf mois de l'année (du 1^{er} janvier au 30 septembre 2009). Ainsi, la puissance produite de la centrale éolienne étudiée est prédite durant les trois derniers mois de l'année (1^{er} octobre au 31 décembre 2009) sur les horizons de prévision allant de 1 à 48 heures. La comparaison entre les puissances prédites et les puissances mesurées permet d'évaluer la capacité de prévision du modèle proposé et ainsi, d'évaluer ses performances sur chaque horizon de prévision.

Bédard *et al.* (2011) ont analysé les incertitudes du modèle de prévisions numériques GEM-LAM, et ont proposé un module géophysique de traitement statistique des sorties qui corrige

les erreurs systématiques du modèle numérique. Toutefois, les erreurs du modèle ont été corrigées pour les prévisions de la vitesse du vent en un point précis de la centrale (au mât de mesures météorologiques). Le module développé par Bedard *et al.* (2011) se sert donc des prévisions météorologiques évaluées par GEM-LAM à 2.5 km de résolution pour prédire la vitesse du vent au niveau du mât de mesures météorologiques, précisément à la hauteur des éoliennes, avec la meilleure précision possible.

La courbe de puissance de la centrale développée au chapitre 3 a uniquement pris en compte les intervalles de temps pour lesquels toutes les éoliennes de la centrale sont en mode opérationnel. Bien que cette supposition permette de mieux contrôler les données pour une meilleure analyse des méthodes et méthodologies adéquates pour modéliser la courbe de puissance de la centrale éolienne, il arrive très souvent que plusieurs turbines de la centrale soient arrêtées pour plusieurs raisons. Par conséquent, le fait de supposer toutes les turbines de la centrale éolienne en opération conduit à réduire de moitié la quantité de données disponibles pour mener l'étude. Il serait donc important d'étudier dans un premier temps le phénomène des arrêts des éoliennes de la centrale étudiée dans le but de conserver plus de données pour mener l'étude.

4.2 Les arrêts des turbines

Les arrêts des éoliennes affectent la production de la centrale. Une centrale éolienne fonctionnant avec la moitié des éoliennes arrêtée ne peut pas produire la même quantité d'énergie qu'elle produirait si toutes les éoliennes étaient disponibles dans les mêmes conditions météorologiques. La procédure de contrôle de qualité des données présentée au chapitre 2 a supposé que la centrale éolienne opère avec toutes les turbines disponibles. Ainsi, les intervalles de temps pour lesquels au moins une éolienne était arrêtée avaient été écartés. Repérer les temps d'arrêts des éoliennes de la centrale ainsi que la fréquence des arrêts, et corriger la puissance totale de la centrale dans les intervalles de temps correspondants s'avère donc nécessaire.

La centrale éolienne étudiée est constituée de plusieurs éoliennes. Les données mesurées de 66 éoliennes ont été disponibles pour réaliser cette étude. La figure 4.2 présente la fréquence des arrêts des turbines sur la centrale éolienne durant l'année 2009. Il apparaît que la centrale fonctionne avec toutes les éoliennes disponibles environ 50% du temps sur toute l'année. Le plus souvent, une à cinq éoliennes sont arrêtées simultanément. En conservant les données des intervalles de temps dans lesquels une à cinq turbines sont arrêtées, l'on est capable de conserver un nombre considérable de données; soit, environ 80% des données supprimées à cause des arrêts des turbines.

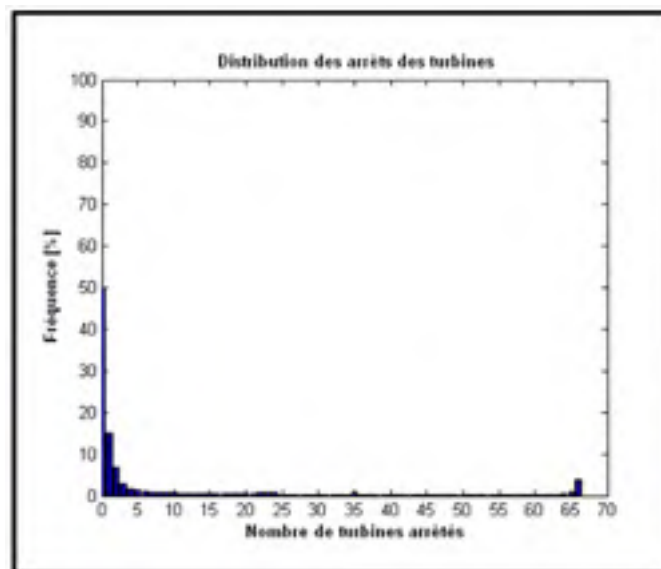


Figure 4.2 Fréquence des arrêts des turbines sur une année complète

Plusieurs raisons peuvent conduire à l'arrêt des turbines dans une centrale éolienne. Pour le cas de la centrale éolienne étudiée, les éoliennes installées sont conçues pour fonctionner pour des vitesses du vent comprises entre 3.5 et 25m/s. Les vitesses du vent en dehors de cet intervalle de vitesses provoqueront donc l'arrêt des turbines. On peut donc dire que la vitesse du vent constitue une raison des arrêts des turbines sur la centrale éolienne (les vitesses de vent trop fortes ou les vitesses de vent trop faibles). La figure 4.3 montre la vitesse du vent

lorsque toutes turbines de la centrale éolienne sont arrêtées simultanément. Elle montre que les arrêts de toutes les turbines surviennent le plus souvent dans les périodes de vitesses de vent faibles. Cependant il est remarquable une période de temps durant laquelle la vitesse du vent est comprise dans l'intervalle des vitesses de fonctionnement des éoliennes installées sur la centrale. Dans ce cas, les arrêts peuvent traduire plusieurs autres raisons telles que les travaux de maintenance ou des conditions météorologiques pouvant causer le phénomène de surpuissance tel que décrit par Dabin (2008).

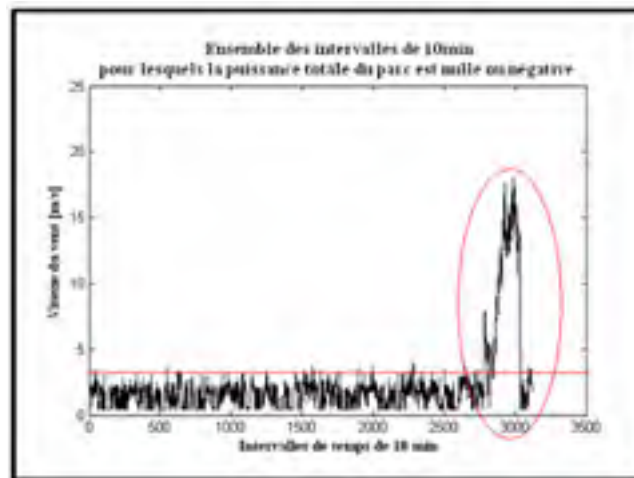


Figure 4.3 Comparaison entre les arrêts des turbines et la vitesse du vent

La puissance totale de la centrale correspondant aux intervalles de temps pour lesquels une à cinq turbines sont arrêtées est corrigée suivant l'équation 4.1. Cette correction suppose que toutes les éoliennes de la centrale produisent la même quantité d'énergie à chaque instant.

$$P_c = \frac{N}{N - n} P_m \quad (4.1)$$

où : P_c est la puissance totale corrigée de la centrale;
 P_m est la puissance totale mesurée de la centrale;

- N est le nombre total de turbines de la centrale; et,
 n est le nombre de turbines arrêtées, ce qui est compris entre 1 et 5.

Si $n=N$, alors la puissance totale corrigée de la centrale est $P_c = 0$, tandis que si $n=0$, alors la puissance totale corrigée est $P_c = P_m$.

La prise en compte des arrêts des turbines, en corrigeant la puissance totale mesurée de la centrale selon l'équation 4.1, a permis de conserver une quantité importante de données sans toutefois affecter la courbe de puissance de la centrale éolienne de manière significative. Les figures 4.4 et 4.5 montrent les deux courbes de puissances de la centrale éolienne avant et après la prise en compte des arrêts des turbines.

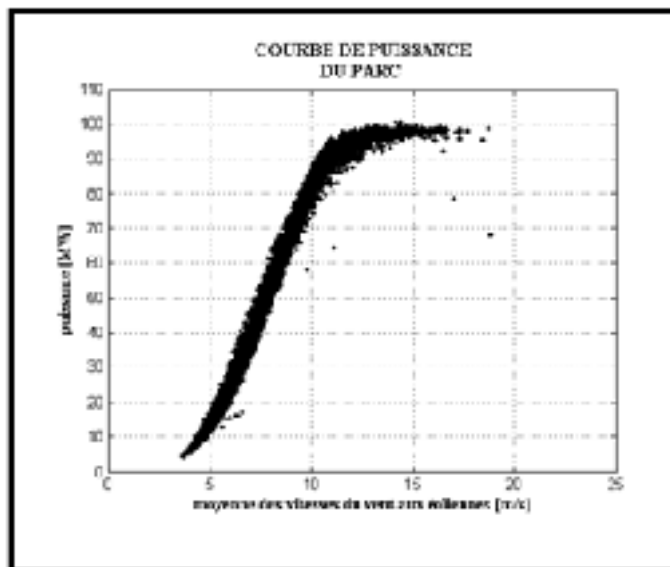


Figure 4.4 Courbe de puissance de la centrale sans arrêts des turbines

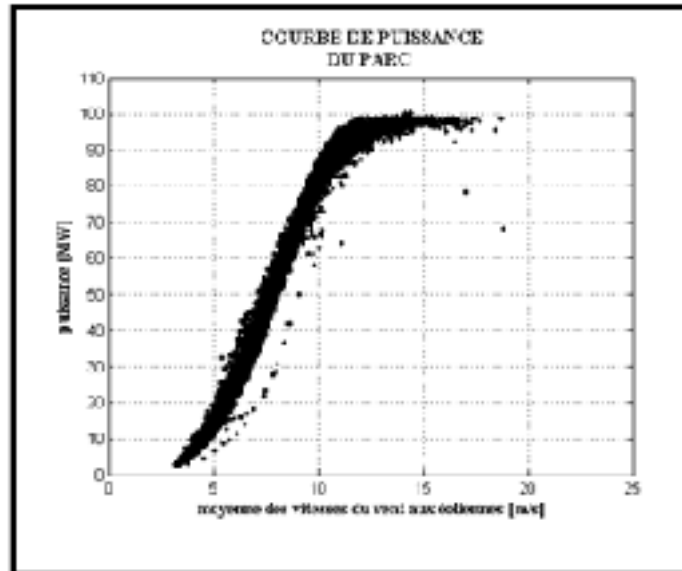


Figure 4.5 Courbe de puissance de la centrale incluant les arrêts des turbines

4.3 Prédiction des variables météorologiques

Environnement Canada effectue les prévisions météorologiques sur un maillage à 2.5 km de résolution. Les variables sont prédites par le modèle pour des horizons de prévision allant d'une heure à quarante-huit heures. Le tableau 4.1 montre les variables prédites par GEM-LAM qui sont utilisées pour évaluer la prévision de la puissance de la centrale éolienne.

Tableau 4.1 Variables météorologiques prédites par GEM-LAM et utilisées pour prédire la puissance de la centrale (niveau de hauteur : I=50m, II=120m)

Paramètres	Niveau de hauteur	Unités
Vitesse du vent	I	m/s
	II	m/s
Température	I	K
	II	K
Pression	I	hPa
	II	hPa
Humidité relative	I	-
	II	-
Direction du vent	-	Degré
Latitude et longitude	-	Degré
Rugosité	-	m

La courbe de puissance optimale de la centrale proposée au chapitre 3 utilise les variables météorologiques suivantes : la moyenne des vitesses aux éoliennes, la masse volumique de l'air au mât et la direction du vent au mât. Il est donc nécessaire de calculer ces variables requises à partir des sorties du modèle de prévisions météorologiques fourni par Environnement Canada.

4.3.1 Prédiction de la moyenne des vitesses aux éoliennes

Les prévisions de la moyenne des vitesses du vent aux éoliennes se calculent en utilisant les prévisions de la vitesse du vent à deux différentes hauteurs (50m et 120m) calculées par GEM-LAM. La première étape consiste à importer sur une grille 7x7, les prévisions de la vitesse du vent aux deux hauteurs disponibles (50m et 120m) à 2.5 km de résolution. Ensuite,

la vitesse du vent est évaluée à 80m en chaque point de la grille. Pour ce faire, une interpolation verticale de la vitesse de vent est nécessaire.

Bédard (2010) a analysé plusieurs méthodes d'interpolation verticale de la vitesse du vent prédite par GEM-LAM. Il a comparé la loi de puissance, le profil logarithmique et le profil linéaire pour différentes situations de l'atmosphère (stable, instable et neutre). Bédard (2010) a abouti à la conclusion selon laquelle les profils linéaires ne sont pas adéquats pour l'interpolation verticale de la vitesse du vent prédite par GEM-LAM. Selon les conclusions de son analyse, le profil logarithmique non-neutre est meilleur et recommandé pour interpoler la vitesse du vent à partir des sorties de GEM-LAM, car il donne de meilleurs résultats pour l'interpolation verticale de la vitesse du vent quel que soit l'état de l'atmosphère (stable, instable ou neutre). Le profil logarithmique non- neutre est donc utilisé pour interpoler la vitesse du vent en chaque point de la grille de prévision de GEM-LAM. L'équation 4.2 est utilisée pour interpoler la vitesse du vent à 80m, à partir des prévisions de la vitesse du vent aux deux hauteurs disponibles (50m et 120m) :

$$U(Z) = \frac{u_*}{k} \left(\ln \left(\frac{Z}{z_0} \right) - \varphi \right) \quad (4.2)$$

$$u_* = \frac{k(U_2 - U_1)}{\ln(z_2/z_1)} \quad (4.3)$$

$$\varphi = \frac{\left(U_2 \ln \left(\frac{z_1}{z_0} \right) - U_1 \ln \left(\frac{z_2}{z_0} \right) \right)}{(U_2 - U_1)} \quad (4.4)$$

- où φ est la fonction de stabilité;
 z_0 est la rugosité prédite par le modèle (en mètre);
 u_* est la vitesse de friction (en m/s);

- U_i est la vitesse du vent à la hauteur z_i ;
- K est la constante de Von Karman (0.4 sans unité); et,
- U est la vitesse du vent évaluée à la hauteur Z .

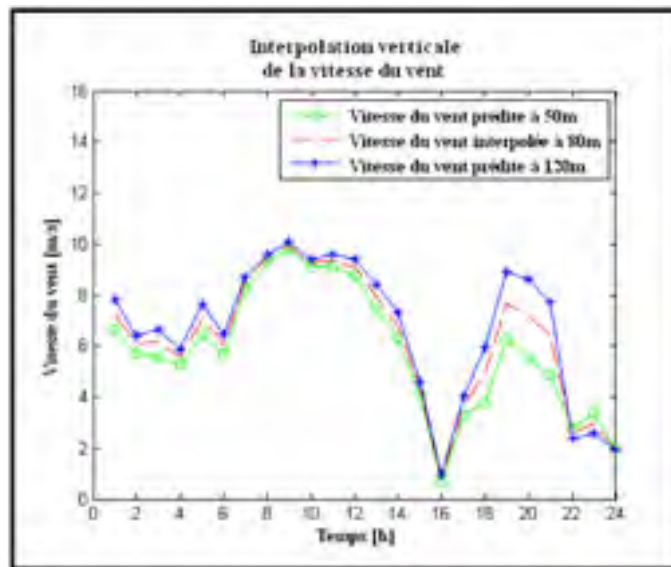


Figure 4.6 Résultats de l'interpolation verticale de la vitesse du vent

La figure 4.6 montre que la vitesse interpolée à 80m à partir des vitesses du vent à 50m et 120m est à chaque instant comprise entre les vitesses du vent à 50m et à 120m. Cela correspond bien à un résultat auquel l'on devrait s'attendre.

Une fois la vitesse du vent interpolée à chaque point de la grille de prévision de GEM-LAM, la seconde étape consiste à évaluer la prévision de la moyenne des vitesses du vent aux éoliennes. La figure 4.7 montre la position des points du maillage de GEM-LAM par rapport à la position des éoliennes sur la centrale. La comparaison entre la moyenne de la vitesse du vent en chaque point de la grille et la moyenne de la vitesse du vent aux éoliennes est représentée par une droite de régression à la figure 4.8. Cette figure montre une bonne corrélation entre la moyenne des vitesses du vent des points de la grille et la moyenne des vitesses du vent aux éoliennes sans aucun post-traitement.

Bédard (2010) a analysé les erreurs de GEM-LAM en comparant les prévisions de la vitesse du vent par GEM-LAM et les vitesses du vent mesurées au mât sur trois centrales éoliennes différentes. Les résultats qu'il a obtenu ont montré qu'en faisant une régression linéaire par secteur de direction du vent, il est possible de réduire les incertitudes de prévision de la vitesse du vent (par GEM-LAM) en un point d'environ 25%. Pour cette étude, la régression linéaire est utilisée afin d'essayer d'augmenter la précision de la prévision de la moyenne des vitesses du vent aux éoliennes. L'équation présente comment la vitesse du vent en un point sur la centrale peut être évaluée en exploitant les prévisions de la vitesse du vent sur la grille de prévision de GEM-LAM.

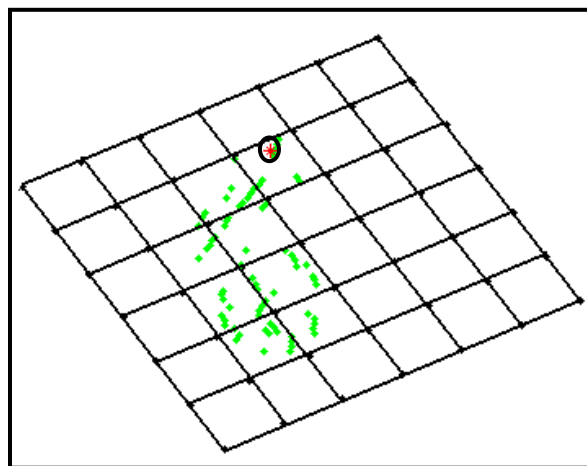


Figure 4.7 Comparaison de la position des points de la grille de prévision avec la position des éoliennes. Les points verts (dispersés) représentent la position des éoliennes et le point encerclé représente la position du mât

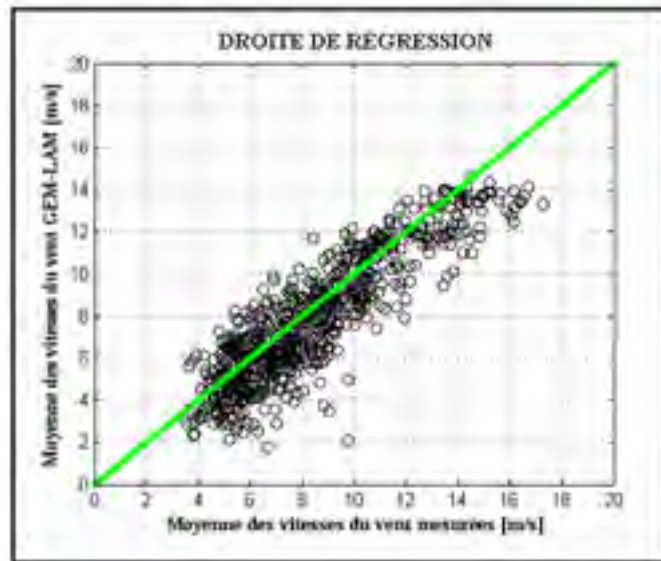


Figure 4.8 Comparaison entre la moyenne des vitesses du vent prédites par GEM-LAM et la moyenne des vitesses du vent aux éoliennes

$$V_{MP \text{ linéaire}} = \sum_i (A_i \times V_{prédite_i}) + B \quad (4.5)$$

Dans l'équation 4.5, la prévision de la moyenne des vitesses du vent aux éoliennes est représentée par $V_{MP \text{ linéaire}}$ (où MP signifie « Multipoint »), alors que $V_{prédite_i}$ représente la vitesse du vent prédite en chaque point i de la grille de prévision de GEM-LAM; A_i et B représentent les paramètres statistiques à optimiser. Pour prédire la vitesse du vent en un point, Bédard (2010) a optimisé les constantes A_i et B en utilisant la technique d'apprentissage par rétro-propagation de l'erreur. Les étapes nécessaires pour évaluer les coefficients sont présentées :

1. Initialiser les paramètres statistiques A_i et B par des valeurs aléatoires.
2. Évaluer le gradient $\left(\frac{\partial E}{\partial A_i}\right)$ de la fonction d'erreur :

$$E = \frac{1}{2} (V_{Mesurée} - V_{MP \text{ linéaire}})^2 \quad (4.6)$$

$$\frac{\partial E}{\partial A_i} = (V_{Mesurée} - V_{MP \text{ linéaire}}) \times \frac{\partial (\sum_i (A_i \times V_{prédite_i}) + B)}{\partial A_i} \quad (4.7)$$

3. Mettre à jour les poids A_i et B en soustrayant au gradient de l'erreur multipliée par une constante d'apprentissage (lr), l'ajustement précédent multiplié par un momentum (mc) :

$$\Delta A_i(t) = -lr \times \frac{\partial E(t)}{\partial A_{i(t)}} - mc \times \Delta A_i(t - 1) \quad (4.8)$$

4. Reprendre les étapes 2 et 3 jusqu'à ce que le gradient de la fonction d'erreur de l'ensemble des données d'entraînement devienne suffisamment petit (10^{-10}) ou si la fonction d'erreur de l'ensemble des données pour la validation croisée ne diminue pas au bout de cinq itérations consécutives.

Le résultat de l'apprentissage est apprécié en comparant la vitesse prédite à la vitesse mesurée. La vitesse prédite étant évaluée en faisant une somme pondérée des vitesses du vent en chaque point de la grille de prévision. Chaque point de la grille de prévision a un poids de A_i qui est déterminé par la technique d'apprentissage ci-dessus mentionnée. La figure 4.9 permet de comparer la moyenne des vitesses du vent prédites et mesurées aux nacelles.



Figure 4.9 Comparaison entre la moyenne des vitesses du vent prédites et la moyenne des vitesses aux éoliennes

En observant la figure 4.10, on peut remarquer que le module de sortie statistique améliore la prévision du vent sur la centrale éolienne, mais toutefois, l'amélioration est très peu significative. GEM-LAM a des difficultés à prédire les vitesses du vent autour de 4m/s, alors qu'il a toujours tendance à les sous-estimer. Ce phénomène est moins observé lorsque que le module de traitement statistique a corrigé le biais systématique de GEM-LAM pour ces vitesses de vent faibles. On peut aussi remarquer que GEM-LAM prédit assez bien la vitesse du vent sur la centrale éolienne étudiée. Pour l'illustrer, la figure 4.12 montre une droite de régression comparant les vitesses mesurées au mât et les vitesses prédites par GEM-LAM sur le point de la grille de prévision le plus proche du mât. En comparant la figure 4.12 à la figure 4.11, tirée du mémoire de maîtrise de Bédard (2010), il est remarquable que les prévisions de la vitesse du vent de GEM-LAM sur la centrale éolienne étudiée paraissent plus précises que le modèle corrigé avec un MOS sur la centrale éolienne ayant servi aux études expérimentales de Bédard (2010). Cela pourrait expliquer pourquoi le MOS n'a pas apporté une amélioration très significative, les prévisions elles-mêmes étant déjà corrélées aux mesures sans aucun post-traitement.

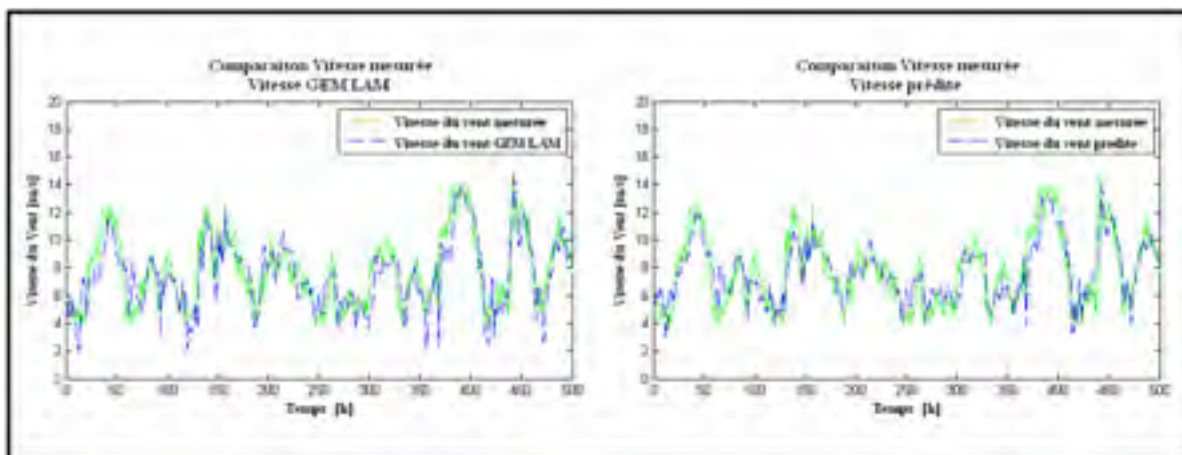


Figure 4.10 Comparaison entre les prévisions de la vitesse du vent et les mesures

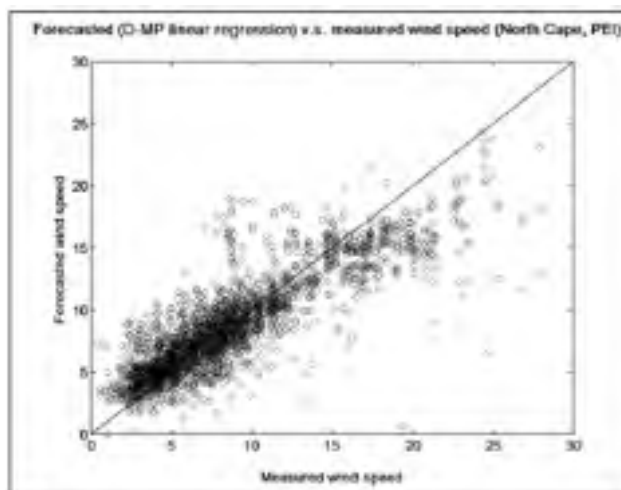


Figure 4.11 Comparaison entre les prévisions de la vitesse du vent au mât et les vitesses du vent mesurées. Les prévisions sont corrigées par une régression linéaire par secteur de direction de vent
Tirée de Bédard (2010)

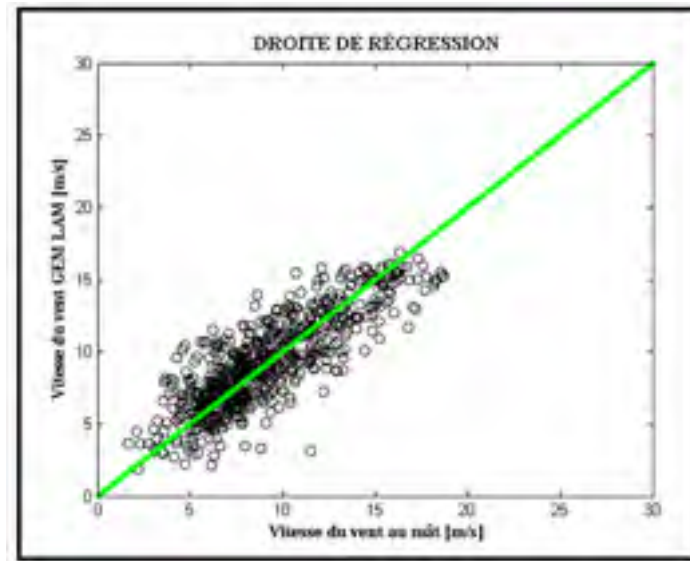


Figure 4.12 Comparaison entre les prévisions de la vitesse du vent au mât et les vitesses du vent mesurées. Les prévisions ne sont pas corrigées

Dans son analyse, Bédard (2010) a montré qu'en faisant l'apprentissage des coefficients A_i et B par secteur de direction, l'on est capable de prédire la vitesse du vent au mât avec une meilleure précision. Cette approche a donc été testée pour le cas de la prévision de la moyenne des vitesses du vent aux éoliennes. Mais contrairement aux résultats obtenus par Bédard (2010), l'approche de l'apprentissage par secteur de direction de vent a conduit à une augmentation de l'erreur globale. Plusieurs raisons peuvent expliquer ce résultat. Premièrement, la vitesse du vent mesurée en un point est beaucoup plus sensible aux phénomènes de sillage qui sont mis à l'évidence au chapitre 3. Par conséquent, effectuer un apprentissage par secteur de direction de vent permettrait de capturer davantage les erreurs systématiques dues aux phénomènes de sillage et à l'influence de la topographie, ce qui est plus difficilement observé lorsque la vitesse du vent est moyennée sur une grande région géographique (telle qu'elle est évaluée dans la présente étude). D'autre part, Bédard (2010) a utilisé pour ses analyses, une année complète de données pour implanter son modèle et une année complète pour le tester. Par contre, une année a servi à mener cette étude dont neuf mois pour l'apprentissage et trois mois pour les tests. De plus, il a été observé plusieurs intervalles de temps dans lesquelles certaines variables nécessaires pour bâtir le modèle n'ont

pas été enregistrées, soit environ 19% du nombre total de données, tel que présenté au chapitre 2. Ce problème a conduit à réduire l'ensemble de données disponibles pour mener la présente étude. En conséquence, dans certains secteurs de direction de vent, le nombre de données n'a pas été suffisant pour faire un apprentissage par secteur de direction du vent, par la technique de rétro-propagation de l'erreur qui nécessite une grande quantité de données pour pouvoir capturer la relation entre la moyenne des vitesses aux éoliennes et celles en chacun des points de la grille de prévision pour chaque secteur de direction du vent. Il serait donc judicieux de tester l'approche soit sur une autre centrale éolienne, ou alors sur la même centrale éolienne en utilisant un nombre de données beaucoup plus élevé que celui qui a été disponible pour la présente étude. Des conclusions plus claires pourraient alors être tirées.

4.3.2 Prévision de la masse volumique de l'air

La masse volumique de l'air dépend de trois variables météorologiques qui sont la température, la pression de l'air et l'humidité relative. Les prévisions de la masse volumique de l'air (ρ) s'effectuent en utilisant les prévisions de la température (T), les prévisions de la pression (p) et les prévisions de l'humidité relative à 50m de hauteur fournies par GEM-LAM. La pression atmosphérique et la température sont d'abord élevées à la hauteur des nacelles des éoliennes (80 m) à l'aide des équations 4.11 et 4.10. La masse volumique de l'air se calcule ensuite en fonction de la température, la pression et la constante spécifique de l'air (R_h) à l'aide de l'équation 4.9. Les prévisions de l'humidité relative à 50m sont directement utilisées car dans sa thèse de doctorat, Dabin (2008) mentionne que la variation de la masse volumique de l'air due à l'humidité est négligeable dans le cas des éoliennes. Les mêmes équations ont été utilisées pour évaluer les prévisions de la masse volumique de l'air dans la présente étude.

$$\rho(p, T, z) = \frac{p(z)}{R_h T(z)} \quad (4.9)$$

$$T(z) = T(z_{Ref}) - \frac{n-1}{n} \frac{g}{R} (z - z_{Ref}) \quad (4.10)$$

$$p(z) = p(z_{Ref}) \left(1 - \frac{g(z - z_{Ref})}{RT(z_{Ref})} \frac{n-1}{n} \right)^{\frac{n}{n-1}} \quad (4.11)$$

- où :
- n est l'exposant polytropique et vaut 1.2349 en atmosphère standard;
 - R est la constante des gaz parfait pour l'air et vaut $275.05 \text{ m}^2 \cdot \text{K}^{-1} \cdot \text{s}^{-1}$;
 - z_{Ref} est la hauteur de référence;
 - g est la constante gravitationnelle et vaut $9.81 \text{ m} \cdot \text{s}^{-2}$; et,
 - R_h est la constante spécifique de l'air et vaut :

$$R_h = \frac{R_s}{1 - \left(\varphi \cdot \frac{p_{sat}}{p} \right) \cdot \left(1 - \frac{R_s}{R_v} \right)} \quad (4.12)$$

- où :
- φ représente l'humidité relative;
 - R_v est la constante spécifique de la vapeur d'eau ($461 \text{ J} \cdot \text{Kg}^{-1} \cdot \text{K}^{-1}$);
 - R_s est la constante spécifique de l'air sec ($287.06 \text{ J} \cdot \text{Kg}^{-1} \cdot \text{K}^{-1}$); et,
 - p_{sat} est la pression de la vapeur saturante de l'eau dans l'air et s'exprime de la manière suivante :

$$p_{sat} = 611.213 \cdot \exp\left(\frac{17.5043 \cdot \vartheta}{241.2 + \vartheta}\right) \quad (4.13)$$

avec ϑ , la température en degrés Celsius.

La figure 4.13a présente la droite de régression linéaire entre les données mesurées et celles prédites de la température évaluée en degré Kelvin. Les prévisions sont calculées à partir des

sorties de GEM-LAM au niveau I (à 50m de hauteur). La comparaison entre les prévisions au point du maillage le plus proche du mât et les valeurs mesurées au mât montre que le modèle numérique est capable de prédire la température avec une bonne précision sur la centrale éolienne étudiée, avec un coefficient de régression qui vaut 0.97. Par conséquent, aucun post-traitement n'est nécessaire pour corriger les valeurs de températures prédites par GEM-LAM. La figure 4.13b, qui montre la comparaison entre la pression prédite et celle mesurée, montre que le modèle prédit aussi bien la pression atmosphérique sur la centrale éolienne étudié, avec un coefficient de régression qui vaut 0.98. Les prévisions de pression et de température étant de bonne qualité sur la centrale éolienne, prédire la masse volumique de l'air revient à extrapoler les deux variables à la hauteur des éoliennes et à évaluer la masse volumique de l'air sur la centrale éolienne grâce à l'équation 4.9 en tenant compte de l'humidité relative.

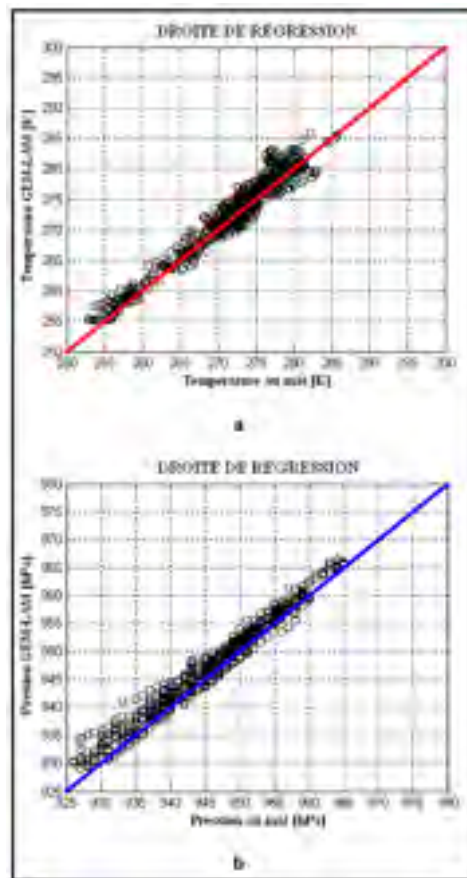


Figure 4.13 Comparaison des prévisions des variables météorologiques avec les valeurs mesurées: a- Température
b- Pression atmosphérique

4.3.3 Prévision de la direction du vent

Pour évaluer les prévisions de la direction du vent sur le site, les sorties de GEM-LAM sont utilisées. Elles donnent les prévisions de la direction du vent, en supposant que la direction du vent ne varie pas avec la hauteur. Cette hypothèse peut affecter la précision des prévisions selon la nature du terrain et la situation météorologique. En terrain complexe, la direction du vent en un point peut varier en fonction de la hauteur. Ce phénomène peut aussi être observé pour certaines conditions météorologiques. Mais pour la centrale éolienne étudiée, la comparaison entre les prévisions de la direction du vent (en utilisant les prévisions au point du maillage le plus proche du mât) et les données mesurées de la direction du vent à la hauteur des nacelles des éoliennes montre que dans la plupart du temps, le modèle de prévision météorologique prédit bien la direction du vent au mât de mesures météorologiques, avec un coefficient de régression qui vaut 0.91. Il n'a donc pas été nécessaire de développer un modèle de correction qui permet de corrélérer les sorties du modèle (GEM-LAM) avec les valeurs de la direction du vent relevées au niveau du mât de mesures météorologiques.

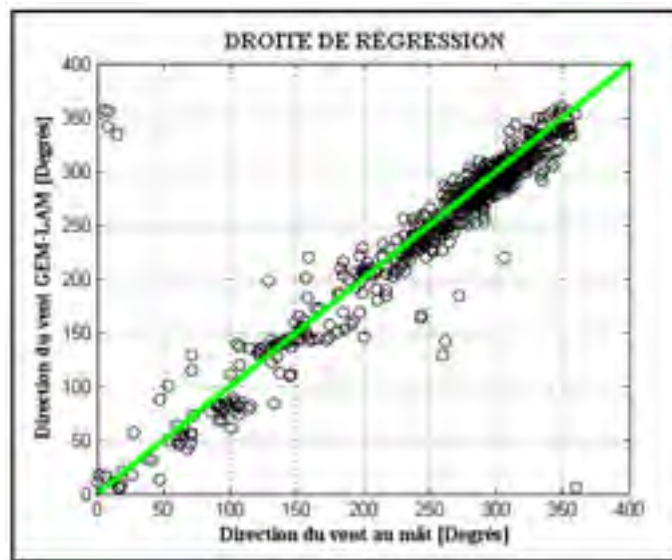


Figure 4.14 Comparaison des prévisions de la direction du vent avec les valeurs mesurées

4.4 La conversion des prévisions météorologiques en prévisions de puissance

La conversion des prévisions météorologiques en prévisions de puissance se fait à l'aide de la courbe de puissance de la centrale développée au chapitre 3. L'analyse effectuée sur la relation entre les variables météorologiques et la puissance produite de la centrale éolienne a permis de mettre à l'évidence l'importance de considérer trois principales variables pour atteindre une bonne précision : la moyenne des vitesses du vent aux éoliennes, la masse volumique de l'air et la direction du vent. GEM-LAM fournit des prévisions météorologiques pour des horizons de prévision allant d'une heure à 48 heures et sur un maillage à 2.5 km de résolution. La comparaison entre les prévisions de la température de GEM-LAM et les valeurs mesurées a montré que le modèle prédit bien la température. Le même constat a pu être fait en ce qui concerne la pression de l'air. La prévision de ces deux variables permet de calculer la prévision de la masse volumique de l'air en prenant en compte la prévision de l'humidité relative tel que décrit à la section 4.3.2. Une comparaison entre les prévisions de la vitesse du vent en un point et les prévisions de la moyenne des vitesses aux éoliennes a montré que le modèle de prévision météo prédit mieux la vitesse du vent sur une région en comparaison à la prévision de la vitesse du vent en un point. L'ensemble des variables météorologiques prédites (moyenne des vitesses aux éoliennes, masse volumique de l'air au mât de mesures et direction du vent au mât de mesures) est introduite dans la courbe de puissance. La précision de la prévision de puissance est appréciée en comparant les séries temporelles des puissances mesurées et des puissances prédites à la figure 4.15. Pour sa part, la figure 4.16 montre les performances du modèle de prévision pour les 48 horizons de prévision disponibles.

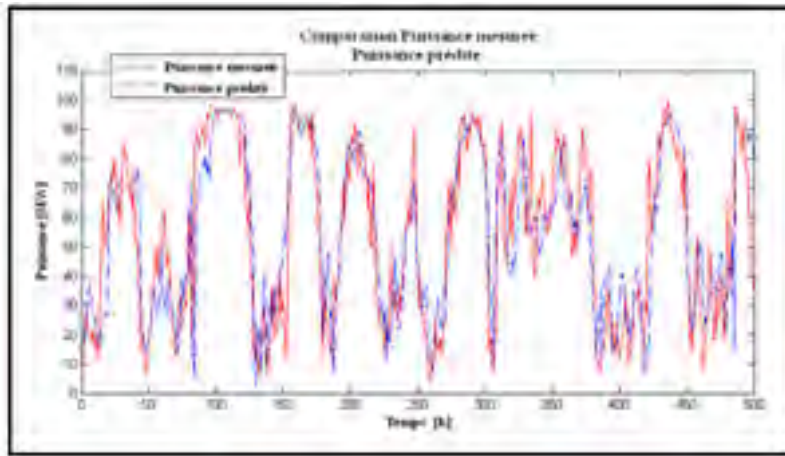


Figure 4.15 Comparaison des séries temporelles des prévisions de la puissance avec les valeurs mesurées

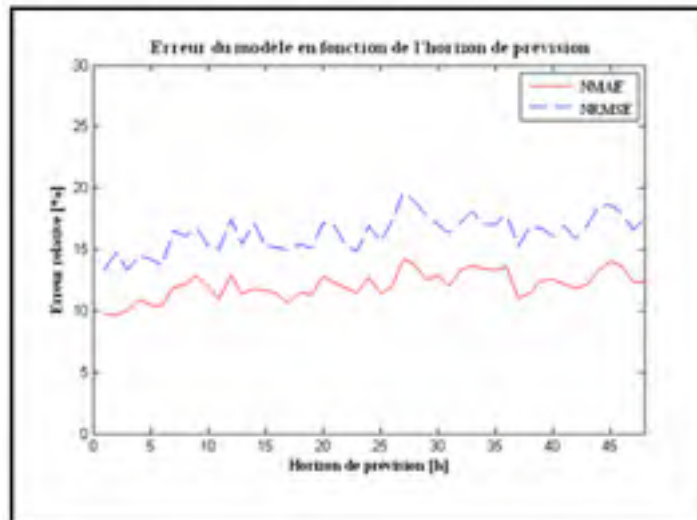


Figure 4.16 Performances du modèle de prévision en fonction de l'horizon de prévision

Les résultats obtenus montrent que les performances du modèle ne varient pas beaucoup en fonction de l'horizon de prévision. L'erreur absolue moyenne (Mean Absolute Error, MAE) varie entre 9.6 et 14.1 % de la puissance totale installée de la centrale, alors que l'erreur quadratique moyenne (Root Mean Square Error, RMSE) varie de 13.2 à 19.4%. Une analyse du modèle est effectuée au chapitre 5, en suivant une procédure bien précise et détaillée. Le

modèle sera d'abord comparé à deux modèles de référence de prévision de la production éolienne. Ensuite, les performances du modèle seront discutées avec des résultats retrouvés dans la littérature.

CHAPITRE 5

ÉVALUATION DU MODÈLE

5.1 Introduction

Le modèle de prévision proposé au chapitre 4 est capable de prédire la puissance produite d'une centrale éolienne jusqu'à quarante-huit heures à l'avance. Il utilise les prévisions météorologiques calculées par GEM-LAM, un modèle numérique de prévisions météorologiques à 2.5km de résolution. Le calcul des paramètres statistiques indicateurs de la précision du modèle a montré que le modèle a une erreur qui ne varie pas beaucoup en fonction de l'horizon de prévision. Cependant, il est nécessaire de l'analyser en détail afin de mieux apprécier les performances observées.

Les performances d'un modèle physique de prévision de la production éolienne dépendent de la précision du modèle numérique de prévisions météorologiques utilisé (Giebel *et al.*, 2011), principalement de la précision de la prévision de la vitesse du vent. D'après la nature de la courbe de puissance de la centrale éolienne, la relation qui lie la vitesse du vent à la puissance produite est de nature cubique. Par conséquent, une petite incertitude de la vitesse du vent peut produire une grande incertitude de la prévision de la puissance, surtout pour des vitesses du vent comprises entre 6 et 13m/s. En utilisant une prévision du vent de mauvaise qualité, il serait donc difficile d'obtenir une prévision de la puissance de bonne qualité. Le modèle de puissance, qui est chargé de convertir les prévisions du vent en prévisions de la puissance, peut aussi être une source d'incertitude du modèle de prévision (Monteiro *et al.*, 2009), car si la relation entre les variables météorologiques et la puissance produite de la centrale éolienne est mal définie, une bonne prévision du vent pourra conduire à une mauvaise prévision de la puissance.

Dans cette étude, la courbe de puissance a été modélisée avec les données mesurées de la centrale éolienne étudiée. La précision du modèle est donc aussi dépendante de la qualité et

du nombre de données disponibles. Si les données utilisées ne sont pas représentatives de la centrale éolienne étudiée, le modèle de puissance pourrait donc prédire la puissance de la centrale avec des incertitudes significatives. D'autres facteurs peuvent aussi influencer les performances d'un modèle de prévision de la production éolienne. On peut citer la nature du terrain ou la capacité totale de production de la centrale éolienne (ou de l'ensemble des centrales éoliennes d'une région si la prévision est régionale) (Martí, 2006), car plus la capacité de la centrale éolienne est grande, moins on observe les incertitudes en utilisant le même modèle de prévision. Il est donc important d'analyser le modèle de manière à rendre les résultats obtenus facilement interprétable par tous les acteurs de la prévision éolienne.

Dans sa thèse de doctorat, Siebert (2008) présente l'évaluation des performances d'un modèle de prévision de la production éolienne comme une analyse des paramètres statistiques indicateurs de la précision. Afin d'analyser la précision de son modèle, il évalue plusieurs paramètres statistiques tels que le biais (*Bias*), l'erreur absolue moyenne (*MAE*), l'erreur quadratique moyenne (*RMSE*) et le coefficient de régression (R^2) en montrant comment ils peuvent être évalués. Cependant, l'évaluation des paramètres statistiques ne peut pas être suffisante pour caractériser un modèle, car un même modèle peut produire des résultats différents dépendamment de la complexité du terrain, de la précision du modèle de prévisions météorologiques utilisé et dépendamment de plusieurs autres facteurs évoqués au paragraphe précédent. La validation du modèle de prévision de la production proposé par Siebert (2008) est alors effectuée en comparant ses résultats avec ceux d'un modèle servant de référence. Landberg *et al.* (2003) présentent la manière d'évaluer le coefficient d'amélioration d'un nouveau modèle par rapport à Persistance ou Climatologie, les deux modèles de référence de prévision de la production éolienne les plus utilisés. Dans sa thèse de doctorat, Pinson (2006) valide son modèle de prévision de la production éolienne de la même façon que Siebert (2008), soit en comparant les paramètres statistiques obtenus avec ceux observés en prédisant la puissance de la même centrale éolienne avec les modèles de références.

Monteiro *et al.* (2009) ont rédigé une revue de la littérature récente et exhaustive des modèles de prévision de la production éolienne pour le centre de recherche américain '*Argonne National Laboratory*'. Dans ce recueil d'informations, Monteiro *et al.* (2009) présentent trois étapes pour évaluer un modèle de la prévision éolienne :

1. **La séparation des données en deux groupes** : Il est très important que l'ensemble des données ayant servi à construire les modèles ne soit pas utilisé pour son évaluation. Les auteurs suggèrent donc de diviser l'ensemble de données disponibles en deux groupes distincts avant de construire les modèles. Ainsi, il pourra être utilisé un ensemble de données pour tester le modèle n'ayant pas servi à l'apprentissage.
2. **L'évaluation des paramètres statistiques** : Tout comme Siebert (2008), les auteurs mentionnent qu'il est important de calculer les paramètres statistiques de performance pour évaluer le modèle de prévision analysé. De plus, les auteurs suggèrent de normaliser les erreurs par rapport à la puissance totale installée de la centrale. La normalisation pourra permettre de comparer les résultats obtenus, quelque soit la capacité totale installée des centrales éoliennes.
3. **La comparaison de la précision avec différents modèles** : Les auteurs suggèrent de comparer le modèle de prévision développé avec des modèles existants, au lieu de les comparer avec les modèles de référence tels que Persistance ou Climatologie. Mais pour l'appliquer, il est nécessaire d'avoir un modèle développé au préalable car la plupart des modèles de prévision de la production éolienne existants sont commercialisés et paraissent comme une boîte noire pour le public.

Dans le souci d'uniformiser la procédure d'évaluation des modèles de prévision de la production éolienne, Madsen *et al.* (2004) proposent une méthodologie mentionnant un ensemble de procédures qui permettent de mieux apprécier les performances d'un modèle de prévision de la production éolienne. L'application de cette méthode couvre l'ensemble des méthodes d'analyse des modèles de prévision de la production éolienne mentionnées au

paragraphe 5.1. Elle est donc utilisée pour évaluer le modèle de prévision proposé dans cette étude. Avant d'appliquer la méthodologie, il est important de la décrire de manière précise.

5.2 Procédure d'évaluation du modèle

Le protocole d'évaluation des modèles de prévision de la production éolienne proposée par Madsen *et al.* (2004) permet de garantir l'appréciation des nouveaux modèles de prévision de la production éolienne. La documentation comprend un ensemble de directives qui permettent de mieux utiliser les critères statistiques pour une meilleure évaluation des modèles. En suivant le protocole, les étapes suivantes doivent être respectées :

1. Décrire la centrale éolienne pour lequel le modèle est bâti et évalué.

Tel que mentionné précédemment, les performances d'un modèle de prévision de la production éolienne dépendent de plusieurs facteurs. Il est donc important de fournir un ensemble d'informations sur la procédure utilisée, ainsi que sur la centrale éolienne expérimentale. Les informations suivantes doivent être disponibles :

- Le nombre de turbines éoliennes et la capacité totale d'énergie installée.
- Les horizons de prévision disponibles.
- Les intervalles de temps dans lesquels les données ont été observées (moyennes de 10min, moyennes horaires, etc.).
- La fréquence de la mise à jour des prévisions des modèles numériques météorologiques. Les modèles numériques météorologiques mettent généralement plusieurs heures de calculs pour évaluer les prévisions météorologiques. Les prévisions météorologiques ne sont donc pas mise à jour à chaque heure de la journée.
- Les caractéristiques du modèle de prévision météorologique. Il doit être mentionné les informations telles que les horizons de prévisions disponibles, la résolution du maillage, et aussi mentionner si les prévisions sont interpolées au niveau de la centrale éolienne ou pas.

- Les données mesurées comme entrées du modèle. Certains modèles de prévision de la production éolienne utilisent les données mesurées pour ajuster la prévision. Dans le cas échant, cela doit être mentionné selon le protocole d'évaluation des modèles utilisé.
2. S'assurer que le modèle est évalué sur un ensemble de données tests n'ayant pas servi à l'apprentissage.

La procédure demande de spécifier de manière claire la période de temps sur laquelle les données sont disponibles, et surtout la manière dont les données sont séparées en deux groupes distincts. Un groupe devra servir pour bâtir le modèle et le deuxième servira à évaluer ses performances.

3. Utiliser un minimum de critères statistiques pour évaluer le modèle.

Le modèle doit être évalué avec des critères statistiques tels que le biais (*Bias*), l'erreur absolue moyenne (*MAE*) et l'erreur quadratique moyenne (*RMSE*). Ces valeurs doivent être normalisées par rapport à la capacité totale installée de la centrale éolienne.

4. Comparer le modèle par rapport à d'autres modèles.

Le modèle doit être comparé avec les modèles de référence ou avec d'autres modèles précédemment développés. Le coefficient d'amélioration doit être calculé et présenté pour chaque horizon de prévision. Il est suggéré non seulement d'évaluer le modèle sur tout l'ensemble de données de tests, mais aussi de présenter l'amélioration du modèle pour des périodes de temps plus courtes telles que des mois par exemple.

Le modèle de prévision de la production éolienne qui est proposé dans cette étude est évalué en suivant cette procédure. Pour des raisons de confidentialité, le nom de la centrale éolienne

et sa situation géographique ne sont pas donnés avec exactitude et les critères statistiques sont normalisés par rapport à la puissance totale de la centrale éolienne.

5.3 Application de la procédure d'évaluation du modèle

L'application du protocole standard d'évaluation des modèles de prévision de la production éolienne proposé par Madsen *et al.* (2004) commence par la description d'ordre générale de la centrale éolienne, et de la procédure utilisée pour prédire la puissance éolienne. La centrale éolienne expérimentale est composée de plusieurs éoliennes et a une capacité totale installée supérieure à 100 MW. Pour la présente étude, les données mesurées de 66 éoliennes ont été disponibles. Les données mesurées de la centrale sont moyennées sur des périodes de temps de 10min. Ces données ont suivi une procédure de contrôle de la qualité qui est décrite au chapitre 2, afin de repérer et d'enlever les données de mauvaise qualité. Le modèle de prévision météorologique utilisé (GEM-LAM 2.5km) donne les prévisions météorologiques pour chaque heure de la journée, et sur des horizons de prévision allant d'une heure à quarante-huit heures. Les prévisions sont évaluées sur un maillage à 2.5 km de résolution, tel que présenté au chapitre 4. La procédure de contrôle de qualité des données a parmi d'observer plusieurs intervalles de temps où toutes les variables n'ont pas été relevées. Ce problème a conduit à réduire l'ensemble des données disponibles pour l'étude. Le modèle n'exploite donc pas les données mesurées en temps réelle pour contribuer à davantage réduire l'erreur de prévision. Il serait aussi judicieux de mentionner que la centrale éolienne étudiée est sur un terrain complexe, car la nature du terrain influence les performances des modèles.

Les données disponibles pour l'étude couvrent l'année 2009 au complet. Les données couvrant les neuf premiers mois de l'année (du 1^{er} janvier au 30 septembre 2009) ont été utilisées pour construire la courbe de puissance de la centrale, et pour bâtir le modèle physique de prévision de la production. Les données des trois derniers mois (1^{er} octobre au 31 décembre 2009) ont servi à évaluer les performances du modèle, en calculant les paramètres statistiques de précision. L'ensemble des données utilisées pour tester le modèle

n'a donc pas servi à apprendre le modèle, ce qui est conforme au protocole d'évaluation des modèles de prévision de la production suivi.

L'utilisation des critères statistiques permet d'analyser et de comprendre les incertitudes du modèle. Le protocole requiert l'évaluation d'un minimum de paramètres statistiques incluant le biais (*Bias*), l'erreur absolue moyenne (*MAE*) et l'erreur quadratique moyenne (*RMSE*). Les détails des équations utilisées pour évaluer ces paramètres statistiques sont présentés au premier chapitre. Les résultats de l'évaluation de ces trois paramètres sont normalisés par rapport à la capacité totale de la centrale, et présentés sous forme graphique par la figure 5.1, 5.2 et 5.3.

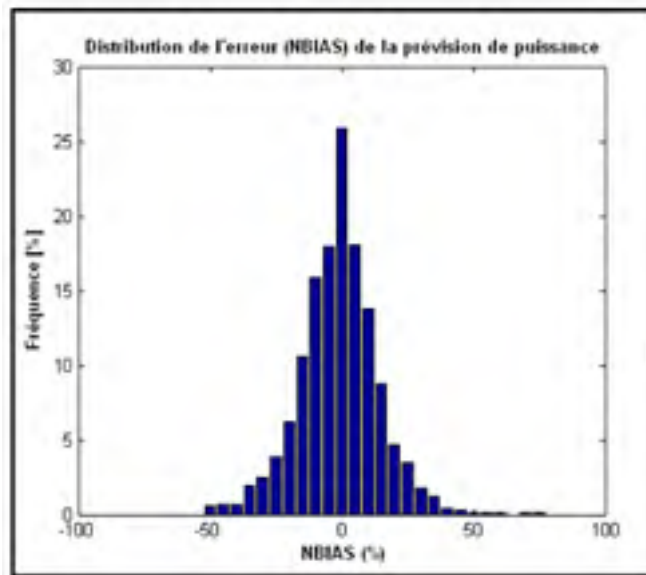


Figure 5.1 Distribution de l'erreur de la prévision de puissance normalisée pour l'ensemble des horizons de prévision allant de une à douze heures

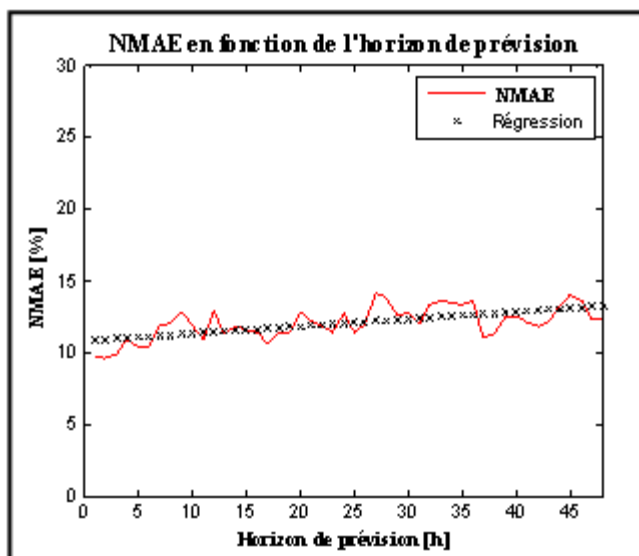


Figure 5.2 Erreur absolue moyenne normalisée (NMAE) du modèle en fonction de l'horizon de prévision

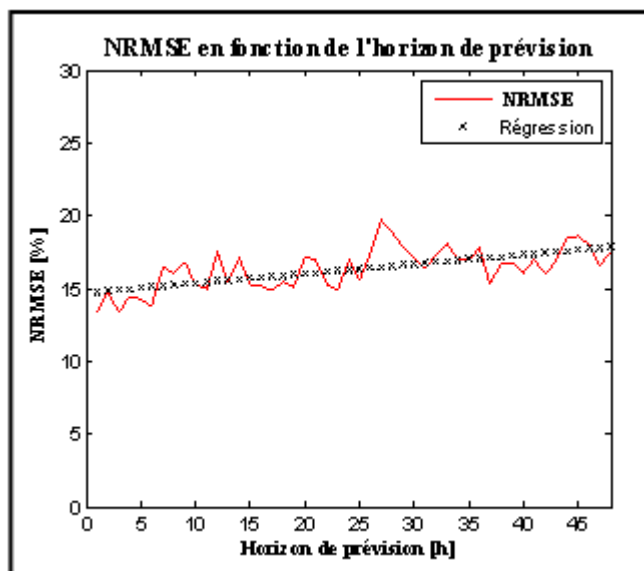


Figure 5.3 Erreur quadratique moyenne normalisée (NRMSE) du modèle en fonction de l'horizon de prévision

La figure 5.1 montre la distribution de l'erreur absolue moyenne du modèle sur les 12 premiers horizons de prévision. La distribution de l'erreur obtenue suit la loi normale, ce qui

montre que le modèle a une probabilité élevée de prédire la puissance avec une bonne précision sur les 12 premiers horizons de prévision. Quelques fois, il peut arriver que le modèle prédise la puissance avec une erreur très grande, soit près de 50% de la puissance totale installée de la centrale. Mais cet événement a une probabilité de se réaliser qui est inférieure à 2%. La figure 5.2, montre la variation de l'erreur absolue moyenne en fonction de l'horizon de prévision. Le modèle présente un NMAE de 9.73% pour les prévisions effectuées une heure à l'avance, 12.93% pour les prévisions effectuées 12h à l'avance, 12.73% pour une prévision de 24h et 14.05% pour une prévision effectuée 45h à l'avance. L'erreur absolue moyenne du modèle (NMAE) évolue donc avec les horizons de prévision. Un comportement similaire est observé pour l'erreur quadratique moyenne du modèle (NRMSE) qui varie entre 13.29% et 19.64% pour les horizons de prévision allant d'une heure à 48 heures. Ces performances ont été obtenues en prédisant la puissance d'une centrale éolienne située en terrain complexe. Afin de mieux les apprécier, il est important de comparer les résultats obtenus d'abord avec les modèles de référence sur la même centrale éolienne, ensuite de discuter des résultats obtenus par d'autres modèles retrouvés dans la littérature.

Le modèle de référence le plus utilisé pour comparer les modèles de prévision de la production éolienne est le modèle Persistance. Ce modèle de référence est décrit en détail au premier chapitre. Le deuxième modèle de référence qui est utilisé pour comparer les résultats du modèle proposé dans la présente étude est le nouveau modèle de référence (NMR) proposé par Nielsen *et al.* (1998) qui est aussi décrit au premier chapitre. La puissance de la centrale éolienne étudiée a été prédite avec les deux modèles de références mentionnés pour chacun des horizons de prévision allant d'une heure à 48 heures. Les performances observées (NMAE et NRMSE) sont présentés aux figures 5.4 et 5.6. Persistance a une très bonne précision (en dessous de 10%) pour les premiers horizons de prévision. Ce modèle suppose que la puissance observée de la centrale à un temps t est la même à un temps $t+k$, où k est l'horizon de prévision. L'état de l'atmosphère ne change pas beaucoup sur une période de temps courte (de l'ordre de quelques heures), ce qui explique la forte probabilité pour laquelle l'état de l'atmosphère à un temps t ne soit pas très différent une ou deux heures plus

tard. C'est la raison pour laquelle Persistance a une bonne précision pour les tout premiers horizons de prévision. Toutefois, l'erreur de Persistance évolue de manière exponentielle avec les horizons de prévision. La figure 5.5 montre que Persistance est capable de battre le modèle développé uniquement pour les deux premières heures de prévision. Une comparaison des performances des deux modèles pour des prévisions effectuées 48 heures à l'avance montre que le modèle est capable d'apporter une amélioration de 60% sur le modèle Persistance (tableau 5.1). L'évaluation du coefficient d'amélioration est effectuée sur tous les horizons de prévision disponibles et présentée à la figure 5.5, ce qui permet de comparer les deux modèles pour chaque horizon de prévision.

Une comparaison similaire est effectuée entre le modèle proposé et le « nouveau modèle de référence » (NMR). Contrairement à persistance, le nouveau modèle de référence peut battre le modèle jusqu'aux quatre premiers horizons de prévision (figures 5.6 et 5.7). Aussi, le nouveau modèle de référence a une performance inférieure à 5% pour les prévisions effectuées une heure à l'avance. Toutefois, le modèle proposé dans cette étude présente aussi une amélioration nette, soit une moyenne de 48% du NMAE et 40% du NRMSE sur l'ensemble des horizons de prévision disponibles allant d'une heure à 48 heures (tableau 5.2).

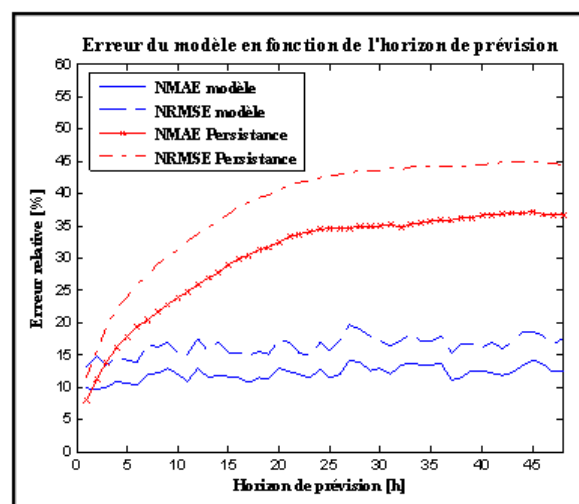


Figure 5.4 Comparaison du modèle avec Persistance en fonction de l'horizon de prévision

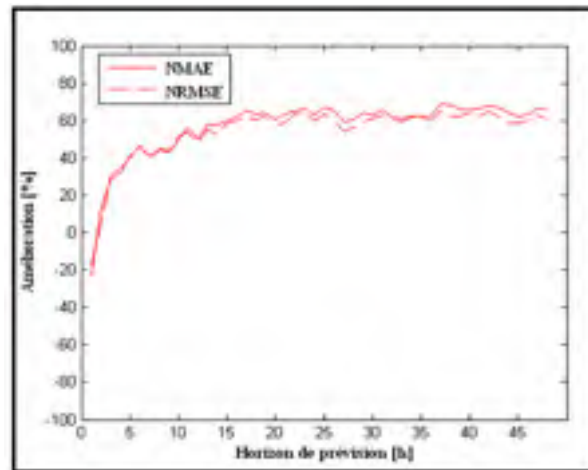


Figure 5.5 Pourcentage d'amélioration du modèle proposé en comparaison à la Persistance en fonction de l'horizon de prévision

Tableau 5.1 Moyenne des erreurs du modèle proposé et de Persistance pour les horizons de prévision de une à quarante-huit heures

Méthodologie	NMAE (%)	NRMSE (%)
Persistence	30.39	38.10
Modèle développé	12.07	16.33
Moyenne de l'amélioration sur 48h	60.26	57.15

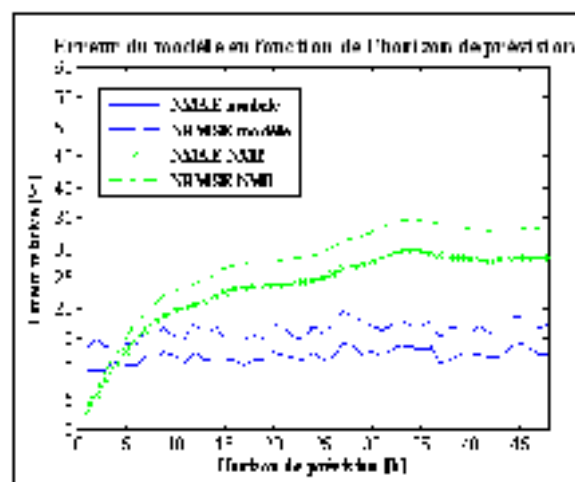


Figure 5.6 Comparaison du modèle proposé avec le NMR en fonction de l'horizon de prévision

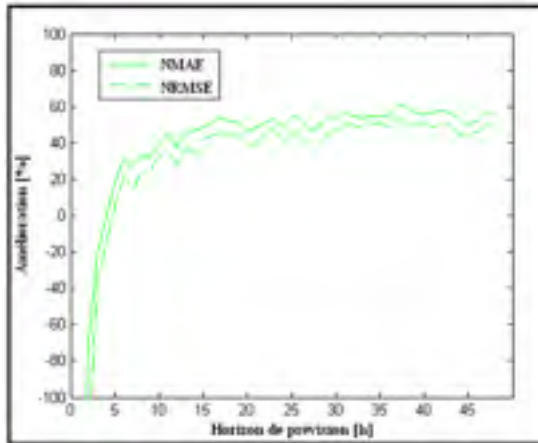


Figure 5.7 Pourcentage d'amélioration du modèle proposé en comparaison au NMR en fonction de l'horizon de prévision

Tableau 5.2 Moyenne des erreurs du modèle proposé et du NMR pour les horizons de prévision de une à quarante-huit heures

Méthodologie	NMAE (%)	NRMSE (%)
Nouveau modèle de référence	23.38	27.40
Modèle développé	12.07	16.33
Moyenne de l'amélioration sur 48h	48.35	40.40

5.4 Quelques résultats de la littérature

Pour mieux apprécier les performances du modèle proposé, il serait important de mentionner quelques résultats retrouvés dans la littérature. Pinson *et al.* (2004) ont présenté une analyse des incertitudes d'un modèle physique de prévision de la production éolienne en prédisant la production éolienne d'une centrale éolienne située en Irlande. La distribution de l'erreur absolue moyenne (*MAE*) du modèle analysé est représentée par la figure 5.8 pour les horizons de prévision d'une heure et 24 heures.

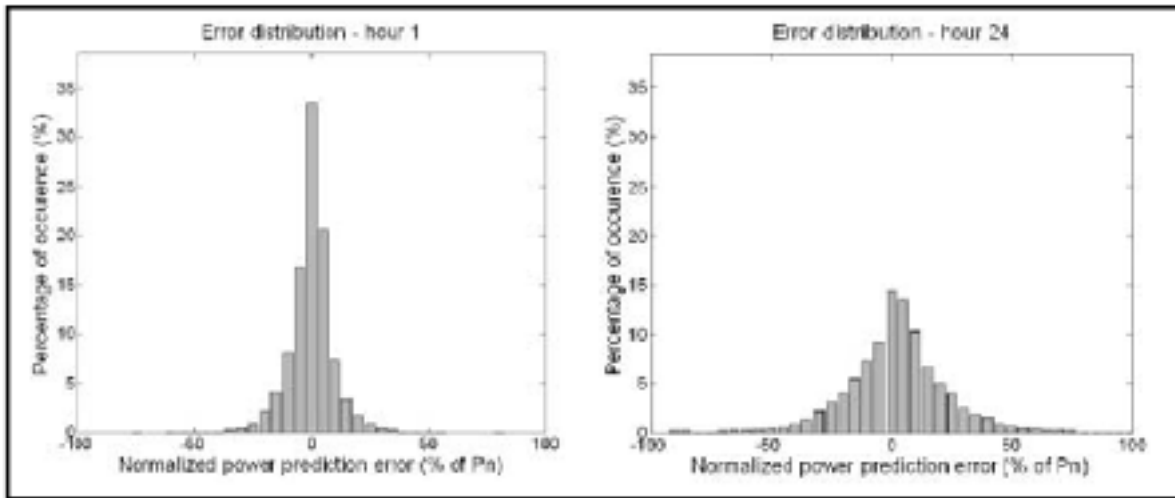


Figure 5.8 Distribution de l'erreur de la prévision de puissance d'un modèle pour les horizons de prévision d'une et 24 heures
Tirée de Pinson *et al.* (2004)

Ces résultats montrent que, comme le modèle proposé dans cette étude, le modèle qu'ils ont développé a une précision qui varie en fonction de l'horizon de prévision. L'allure de la distribution de l'erreur est de nature normale, ce qui révèle que leur modèle a plus de chance de prédire la puissance avec une petite incertitude pour la première heure de prévision. La distribution de l'erreur devient plus grande pour une prévision effectuée à un horizon de 24 heures. Ces graphiques sont assez comparables à l'analyse de l'erreur effectuée du modèle proposé au paragraphe 5.2, (figure 5.1) qui montre la distribution de l'erreur pour la moyenne des horizons de prévision allant d'une heure à douze heures. le modèle proposé dans cette étude a 25% de chance de prédire la puissance totale de la centrale avec une erreur négligeable et moins de 2% de chance de prédire la puissance avec 50% d'erreur pour les prévisions effectuées d'une heure à 12 heures à l'avance. Le modèle analysé par Pinson *et al.* (2004) a pour sa part 33% de chance de prédire la puissance totale de la centrale une heure à l'avance avec une erreur négligeable, contre 15% de chance 24 heures à l'avance.

Afin de comparer l'allure de l'évolution de l'erreur du modèle proposé avec les erreurs des modèles retrouvés dans la littérature en fonction des horizons de prévision, la figure 5.9 présente un exemple de l'évolution de l'erreur quadratique moyenne (NRMSE) du modèle Prediktor en fonction de l'horizon de prévision. Le modèle Prediktor, qui est décrit au

premier chapitre, est utilisé pour prédire la puissance d'une ancienne centrale éolienne située au Danemark, d'une capacité totale de 5 MW. Prediktor, tout comme la méthode implantée dans cette étude, utilise les prévisions météorologiques pour calculer les prévisions de la puissance de la centrale. L'évolution de l'erreur de la prévision de la puissance, en fonction de l'horizon de prévision a un comportement similaire aux erreurs observées du modèle qui sont présentées aux figures 5.2 et 5.3; elles ne varient pas beaucoup en fonction de l'horizon de prévision, et l'erreur quadratique moyenne de leur modèle varie entre 17% et 20% sans correction avec un MOS. Après correction, l'erreur du modèle varie alors entre 15 et 19% pour des horizons de prévision allant jusqu'à 48 heures. Ce modèle a aussi fait l'objet d'une comparaison avec les modèles de référence (Persistance et le « nouveau modèle de référence »). La figure 5.9 montre que les allures des courbes d'erreurs tant des modèles de référence que du modèle présenté, sont comparables à celles du modèle qui a été proposé dans cette étude.

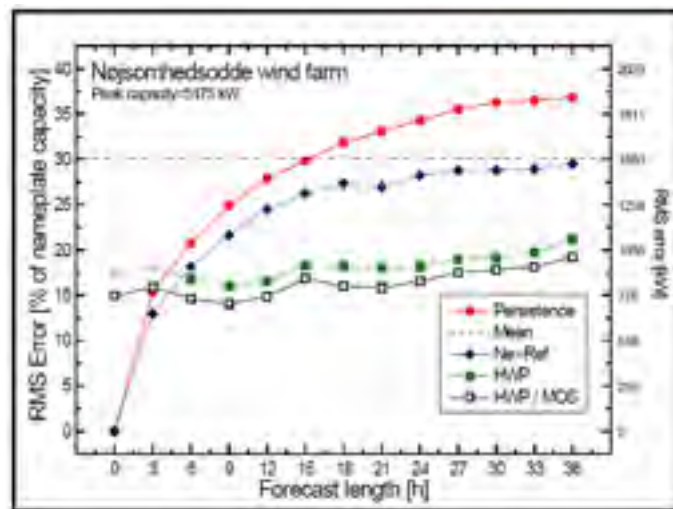


Figure 5.9 Analyse des performances du modèle Prediktor et comparaison avec les modèles de référence
Tirée de Giebel *et al.* (2011)

Pinson (2006), dans sa thèse de doctorat portant sur l'analyse de l'incertitude des modèles de prévision de la production éolienne, présente les performances de cinq modèles différents de prévision de la production éolienne, en comparant les performances des différents modèles avec celles du modèle de référence le plus utilisé : Persistance. D'après les résultats qu'il

présente, les modèles comparés prédisent la puissance avec un *MAE* qui varie de 7% à 15 % par rapport à la puissance totale installée pour les horizons de prévision allant d'une heure à quarante-huit heures. La comparaison entre les différents modèles utilisés et les modèles de référence sont présentés à la figure 5.10. L'amélioration par rapport à Persistance varie entre 40% et 55% après le cinquième horizon de prévision, ce qui est du même ordre que l'amélioration qui a été observé dans cette étude.

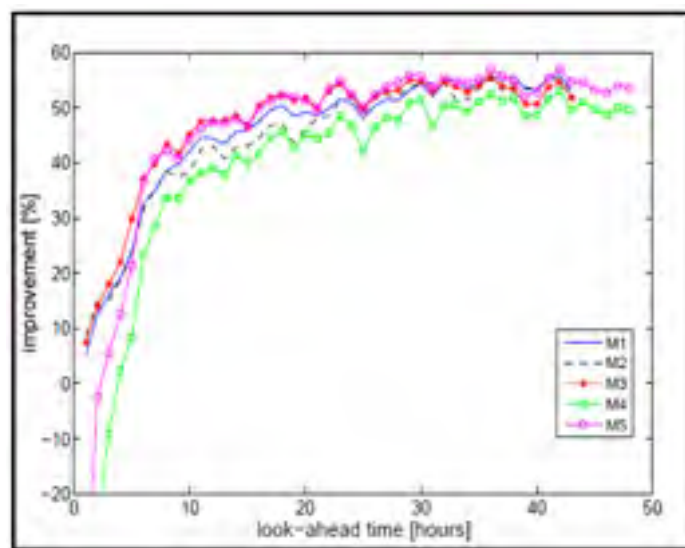


Figure 5.10 Comparaison de cinq modèles différents avec Persistance
Tirée de Pinson (2006)

Monteiro *et al.* (2009) présentent un résumé des conclusions tirées dans le cadre du projet ANEMOS. D'après leurs conclusions, les modèles de prévision de la production éolienne ont typiquement une incertitude (*MAE*) qui varie entre 10% et 21% pour un site, partant d'un terrain plat à un terrain complexe. Cependant, aucune précision n'a été apportée sur les horizons de prévision pour lesquelles ces valeurs sont observées. Kariniotakis *et al.* (2004), Kariniotakis *et al.* (2006), et Martí (2006) insistent sur l'influence de la nature du terrain sur la précision des modèles de prévision de la production. D'après les conclusions de leurs études, l'erreur quadratique des modèle peut dépasser les 25%, jusqu'à atteindre 35% si la centrale éolienne est située en terrain très complexe. Les prévisions de la production de la centrale éolienne de cette étude, située en terrain complexe et en utilisant le modèle proposé, ont permis d'enregistrer une erreur *MAE* de 11% sur les 12 premières heures de prévision,

avec une valeur maximale de moins de 15% pour les horizons de prévision allant jusqu'à 48 heures.

La firme américaine AWS Truewind a effectué des prévisions de la production éolienne de l'ensemble des centrales éoliennes de la province d'Alberta au Canada. Ils ont enregistré des erreurs (NMAE et NRMSE) qui varient entre 20% et 30% de la capacité totale installée (Monteiro *et al.*, 2009). Les raisons de cette mauvaise performance ont été mentionnées dans leur étude. Notamment, les prévisions météorologiques sont très difficiles en Alberta, à cause d'un régime climatique atypique, ce qui conduit à une mauvaise prévision de la vitesse du vent par le modèle numérique de prévisions météorologiques et par conséquent, une prévision de la puissance de mauvaise qualité avec un modèle qui a obtenu des bons résultats dans d'autres régions. Dans leur rapport faisant état de l'état de l'art des prévisions de la production éolienne en Amérique du nord, Porter *et al.* (2009) mentionnent que, pour une capacité de 657 MW d'énergie éolienne générée par des centrales éoliennes, Hydro-Québec enregistre une erreur (*MAE*) de 8.5% pour la première heure, de 13% pour la deuxième heure, de 14.4% pour une prévision de la puissance 12 heures à l'avance et de 15% pour 24 heures à l'avance.

L'ensemble de ces données peuvent permettre d'avoir une idée générale sur les performances des modèles de prévision de la production éolienne. Les erreurs enregistrées montrent que le modèle a des performances qui se comparent bien avec la littérature. Les bonnes performances s'expliquent par des prévisions météorologiques de bonne qualité par GEM-LAM 2.5km sur la centrale éolienne étudiée, et la courbe de puissance de la centrale qui a permis de convertir ces prévisions météorologiques en prévisions de puissance sans apporter une dégradation de la précision du modèle. Pour développer et analyser le modèle, il a été supposé que la centrale éolienne fonctionne avec toutes les turbines disponibles. Les données correspondant à 80% des arrêts des turbines ont été prises en compte en corrigeant la puissance totale de la centrale par rapport au nombre de turbines arrêtées. La conséquence majeure de cette action est qu'en opération, le modèle pourra prédire une puissance supérieure à celle enregistrée à l'heure correspondante, si plusieurs turbines sont arrêtées simultanément car le modèle supposera toujours que toutes les turbines fonctionnent

normalement. Pour prendre en compte l'arrêt des turbines dans le modèle proposé, l'on devra d'abord être capable de connaître le nombre de turbines qui seront arrêtées. Ainsi, la puissance prédite par le modèle proposé sera corrigée en fonction du nombre de turbines opérationnelles selon les sorties du module de prévision des arrêts des turbines. Un autre phénomène qui pourra induire les erreurs au modèle en opération est la situation de givre. Les données mesurées dans les conditions de givrage ont été écartées dans cette étude, pour la simple raison que les instruments de mesures fournissent des données de très mauvaise qualité dans ces conditions, ce qui ne reflète pas souvent les conditions météorologiques de la centrale. Pour prendre en compte ce phénomène dans le modèle proposé, nous pourrions utiliser un outil qui sera capable de prédire l'intensité du givrage; et dans ces conditions, utiliser une courbe de puissance qui compare directement la vitesse prédite par GEM-LAM, à la puissance totale de la centrale. La précision du modèle proposé pourra aussi être améliorée en introduisant les puissances mesurées antérieures de la centrale pour développer un module statistique de sortie après la prévision de puissance. Cette approche n'a pas été implantée dans la présente étude à cause du nombre de données disponibles, mais permettra certainement de faire un gain en précision car d'après les résultats présentés Giebel *et al.* (2011) (figure 5.9), elle peut apporter un gain en précision de 1 à 2% du *RMSE*.

CONCLUSION

L'ensemble des expériences effectuées au cours de cette étude ont visé à établir une méthodologie capable de prédire la puissance produite par une centrale éolienne jusqu'à 48 heures à l'avance. Une année complète de données mesurées d'une centrale éolienne canadienne et des données de prévisions météorologiques de GEM-LAM 2.5km ont été utilisées pour aboutir à ces fins. En dépit de 19% de données manquantes pour la centrale éolienne, près de 80% du nombre total de données ont été retenues après la procédure de contrôle de la qualité des données. La procédure de contrôle de la qualité des données a donc permis de s'assurer que les données utilisées sont de bonne qualité.

Pour développer la courbe de puissance de la centrale éolienne, deux méthodes ont été comparées. La méthode des Bins est présentée dans la littérature comme étant une approche simple servant de référence. Plusieurs études antérieures ont donc modélisé la courbe de puissance des éoliennes et des centrales avec d'autres méthodes tout en l'utilisant comme méthode de base. Parmi les différentes méthodes qui ont déjà été utilisées pour modéliser la courbe de puissance des éoliennes ou des centrales, les réseaux de neurones artificiels (RNA) ont toujours été identifiés comme une méthode supérieure. Plusieurs études précédentes ont par conséquent été conclues en soulignant les avantages des réseaux de neurones et en insistant sur leur capacité à modéliser la courbe de puissance de manière plus appropriée. Un constat a pu être effectué : la plupart du temps, l'utilisation des réseaux de neurones pour modéliser la courbe de puissance des éoliennes ou des centrales se fait en modifiant la méthodologie ou en introduisant plus de variables, ce qui peut être plus difficile à faire avec la méthode des Bins, surtout si le nombre de données est limité.

Ainsi, il est possible que la modification de la méthodologie d'évaluation, ou l'ajout de nouvelles variables, soient en partie responsables du gain observé en précision de la modélisation de la courbe de puissance. C'est pour cette raison que dans un premier temps, la courbe de la puissance de la centrale éolienne a été modélisée avec les mêmes paramètres que ceux utilisés avec la méthode des Bins, et en utilisant la même méthodologie d'évaluation.

En utilisant la vitesse du vent lue au mât de mesures météorologiques et la masse volumique de l'air, il a été possible de modéliser la puissance totale produite par la centrale éolienne en utilisant ces deux méthodes. Un *MAE* de 9.67% est observé pour les réseaux de neurones tandis qu'un *MAE* de 9.89% est observé avec la méthode des Bins. En comparant une méthode avancée avec une méthode de base, il a donc été observé une amélioration de 2.18% en utilisant la même méthodologie et les mêmes paramètres.

La seconde expérience a visé à comparer deux différentes méthodologies tout en utilisant la méthode simple de base. La courbe de puissance de la centrale éolienne a donc été modélisée en comparant d'abord les vitesses lues au mât de mesures avec la puissance totale de la centrale, ensuite la moyenne des vitesses aux éoliennes a été comparée avec la puissance totale de la centrale. La deuxième approche, bien qu'implantée avec la méthode des Bins, a permis d'observer une amélioration de 85.74%. Cette amélioration est nettement supérieure à celle observée avec les RNA en utilisant les vitesses du vent lues au mât de mesures.

De ces deux expériences, la conclusion suivante peut être tirée : Quelque soit l'outil mathématique utilisé pour modéliser la courbe de puissance de la centrale, le véritable gain de la précision s'obtient en choisissant les entrées du modèle qui représentent le mieux l'état du vent sur la centrale éolienne, principalement la vitesse du vent. L'approche avec la moyenne des vitesses aux éoliennes permet donc d'avoir une meilleure description du vent sur l'ensemble de la centrale éolienne en comparaison à celle avec les vitesses lues en un seul point (le mât) qui peut être très influencé par plusieurs facteurs tels que les effets du terrain par exemple.

La troisième expérience qui a utilisé les réseaux de neurones pour tester plusieurs variables (vitesse du vent, masse volumique de l'air, direction du vent, intensité de turbulence et cisaillement), a permis de sélectionner les trois premières variables. En effet, la vitesse du vent est la variable la plus influente sur la relation entre le vent et la puissance de la centrale éolienne car elle exprime l'énergie cinétique du vent qui est convertie en puissance. La masse volumique de l'air apporte une information sur la masse de l'air qui est convertie en

puissance. Elle aussi influence la quantité d'énergie cinétique convertie en puissance. La direction du vent quant à elle permet de capturer les effets du sillage et de la topographie, car la centrale éolienne ne peut pas produire de la même façon dans toutes les directions à cause de ces effets. Les expériences effectuées avec le cisaillement n'ont pas permis de déceler une influence claire sur la puissance totale de la centrale. L'intensité de turbulence quant à elle a conduit à une mauvaise généralisation des données par le réseau de neurones. En effet, étant donné que le cisaillement a été évalué en un seul point de la centrale, il est possible que le cisaillement en ce point ne représente pas l'état du cisaillement sur la centrale éolienne au complet. Dans ce cas, il pourrait ne pas exister de relation entre le cisaillement en ce point et la puissance totale produite de la centrale.

Par ailleurs, en approximant une fonction avec les réseaux de neurones, on suppose a priori qu'il existe une relation entre la sortie et les entrées présentées au réseau. Si tel n'est pas le cas, le réseau éprouve des difficultés lors de l'apprentissage et aboutit à une mauvaise généralisation des données, telle qu'observée. Une approche a été envisagée pour vérifier si l'intensité de turbulence calculée en plusieurs points différents de la centrale éolienne présente une certaine corrélation. Cette approche n'a pas pu être testée parce que les données des autres mâts de mesures de la centrale éolienne n'ont pas été disponibles; et les données aux éoliennes ne comportaient pas l'écart-type de la vitesse du vent qui est nécessaire pour l'évaluer.

La courbe de puissance proposée de la centrale a donc été modélisée en utilisant une méthode avancée (les RNA) et trois variables (vitesse du vent, masse volumique de l'air, direction du vent); la vitesse du vent faisant référence à la moyenne des vitesses aux éoliennes. Cette courbe de puissance a permis de modéliser la puissance de la centrale éolienne sur un ensemble de données n'ayant pas servi à l'apprentissage et d'obtenir un *MAE* et un *RMSE* équivalents à 1% de la capacité totale installée de la centrale éolienne, et un coefficient de régression de 0.99. La méthode proposée a été validée sur la centrale éolienne d'East Point de capacité totale 30MW. Les résultats sont présentés à annexe 2.

Une fois la courbe de puissance développée, le principal défi a été de prédire les variables requises par la courbe de puissance à partir des sorties de GEM-LAM à 2.5km de résolution. Il a été question de prédire la moyenne des vitesses du vent aux éoliennes, la masse volumique de l'air au mât et la direction du vent au mât. D'après les résultats obtenus, il a été observé une corrélation entre la moyenne des vitesses du vent des points de GEM-LAM qui sont sur le site et la moyenne des vitesses du vent aux éoliennes. Cette observation a permis de confirmer l'hypothèse selon laquelle il est possible de prédire cette vitesse du vent requise par la courbe de puissance. Ayant appliqué la méthode proposée par Bédard *et al.* (2011), une régression linéaire a permis de partir des prévisions de GEM-LAM aux prévisions de la moyenne des vitesses aux éoliennes. Un fait a cependant été relevé : Bédard *et al.* (2011) a pu montrer qu'en faisant cette régression linéaire par secteur de direction du vent, il est possible d'atteindre une précision plus importante. D'après ces résultats, l'approche de l'apprentissage par secteur de direction du vent permet de capturer les effets de la topographie sur la vitesse du vent au mât de mesures. Pour ses travaux, Bédard *et al.* (2011) a utilisé deux années complètes de données : l'une pour modéliser et l'autre pour tester. Le nombre de données dans ce cas a été suffisant pour obtenir, dans chaque secteur de direction, un nombre suffisant de données pour faire un apprentissage. Dans cette étude, une année a été utilisée pour toute l'étude (modélisation et test), avec 19% de données manquantes. Dans ce contexte, l'approche par secteur de direction n'a pas apporté une amélioration à la prévision de la moyenne des vitesses du vent aux éoliennes. Cette approche pourrait encore être testée, mais en utilisant le maximum de données possibles. Des conclusions plus claires pourraient alors être tirées à ce propos. Toutefois, les effets de la topographie évoqués par Bédard *et al.* (2011) sont pris en compte par la courbe de puissance qui a été modélisée, en comptant parmi ses entrées la direction du vent. Les prévisions de la direction du vent ont été prises sur le point de la grille de GEM-LAM la plus proche du mât; et les prévisions de la masse volumique de l'air ont été calculées en utilisant les prévisions de la température, de la pression et de l'humidité sur le point de la grille le plus proche du mât. L'ensemble de ces expériences a donc abouti à une méthodologie qui a prédit la puissance de la centrale éolienne avec une moyenne de *MAE* de 12% sur l'ensemble des 48 horizons de prévision disponibles. Ces résultats sont comparables à ceux retrouvés dans la littérature.

Des améliorations peuvent être apportées à la méthodologie pour atteindre une meilleure précision. La courbe de puissance a été développée en supposant que toutes les turbines fonctionnent normalement. Des arrêts des turbines ont été pris en compte, mais tout en normalisant la puissance de la centrale dans ces conditions. En opération, cette courbe de puissance prédira donc la puissance optimale de la centrale éolienne. Des travaux sont actuellement effectués à Environnement Canada pour évaluer avec une bonne précision les prévisions des arrêts des éoliennes causées par les conditions météorologiques. Avec une telle information, la puissance prédite par le modèle sera normalisée de nouveau (en fonction du nombre de turbines arrêtées) et dans ce cas, la précision de la méthode ne devrait pas être grandement affectée. Il serait enfin important de mentionner la possibilité de partir de cette méthodologie pour aboutir à une méthode capable de prédire la production de l'ensemble des centrales éoliennes d'une région. Cette prévision se fera avec une meilleure précision et contribuera encore plus à intégrer l'énergie éolienne au sein des grands réseaux de distribution d'énergie.

ANNEXE I

LISTE DE QUELQUES MODÈLES GLOBAUX DE PRÉVISIONS MÉTÉOROLOGIQUES OPÉRATIONNELS DANS LE MONDE

Tableau-A I-1 Caractéristiques de quelques modèles globaux
de prévisions météorologiques opérationnels
Tiré de Giebel *et al.* (2011)

Symbol Figure 29	Institution	model name	resolution/ model levels	approx. horiz. res.	Grid	Planned
ECMWF	European Center for Medium Range Weather Forecast	IFS	TL799/L91	~25km	Spectral	TL1279/L150 in 2009/10
METOF	Meteorological Office, UK	UM	0.375°x0.5625°/L 50	~40km	Gaussian grid	25km/L70 2009/10
MSC	Meteorological Service of Canada	GEM	0.3°x0.45°/L58	~30km	Gaussian grid	Global 20-25 km uniform resolution, 90 levels, <2015
NCEP	National Center for Environmental Prediction, USA	GFS	TL382/L64	~50km	Spectral	
METFR	Meteo France	ARPEGE	TL538/L60	~ 15km over France	Spectral + gaussian grid with stretching factor	TL798/L70 2009/10
DWD	Deutscher Wetterdienst, Germany	GME	40km/L40	40km	Icosaeder	new model ICON (>2010)
AUSBM	Bureau of Meteorology, Australia	GASP	TL239/L29	~80km	Spectral	
JMA	Japan Meteorological Agency	JMA-GSM	TL319/L40	~60km	Spectral	
KMA	Korea Meteorological Agency	GDAPS	TL426/L40	~45km	Spectral	

ANNEXE II

ÉVALUATION DE LA COURBE DE PUISSANCE DE LA CENTRALE ÉOLIENNE D'EAST POINT (IPE)

1. Cartes géographiques de la centrale éolienne d'East Point

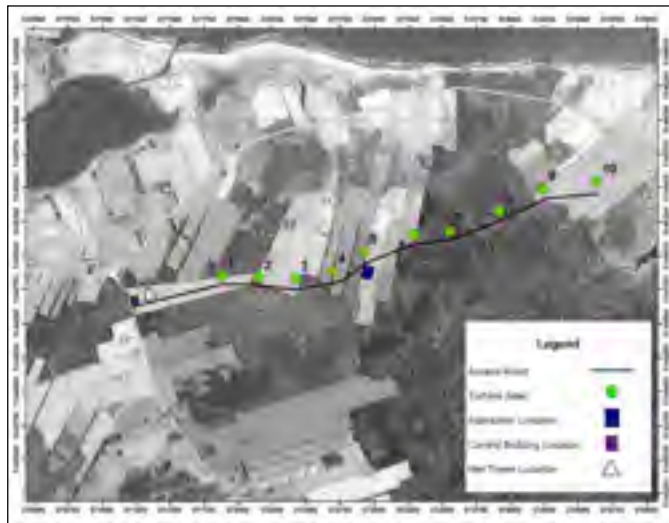


Figure-A II-1 Carte aérienne de la centrale éolienne d'East Point

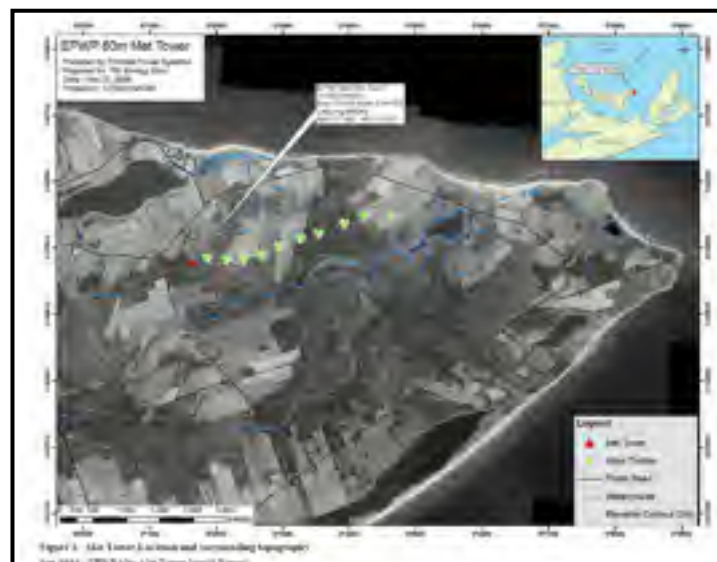


Figure-A II-2 Position du mât de mesures météorologiques et présentation de la topographie (East Point)

2. Comparaison de deux différentes méthodes : Bins et RNA

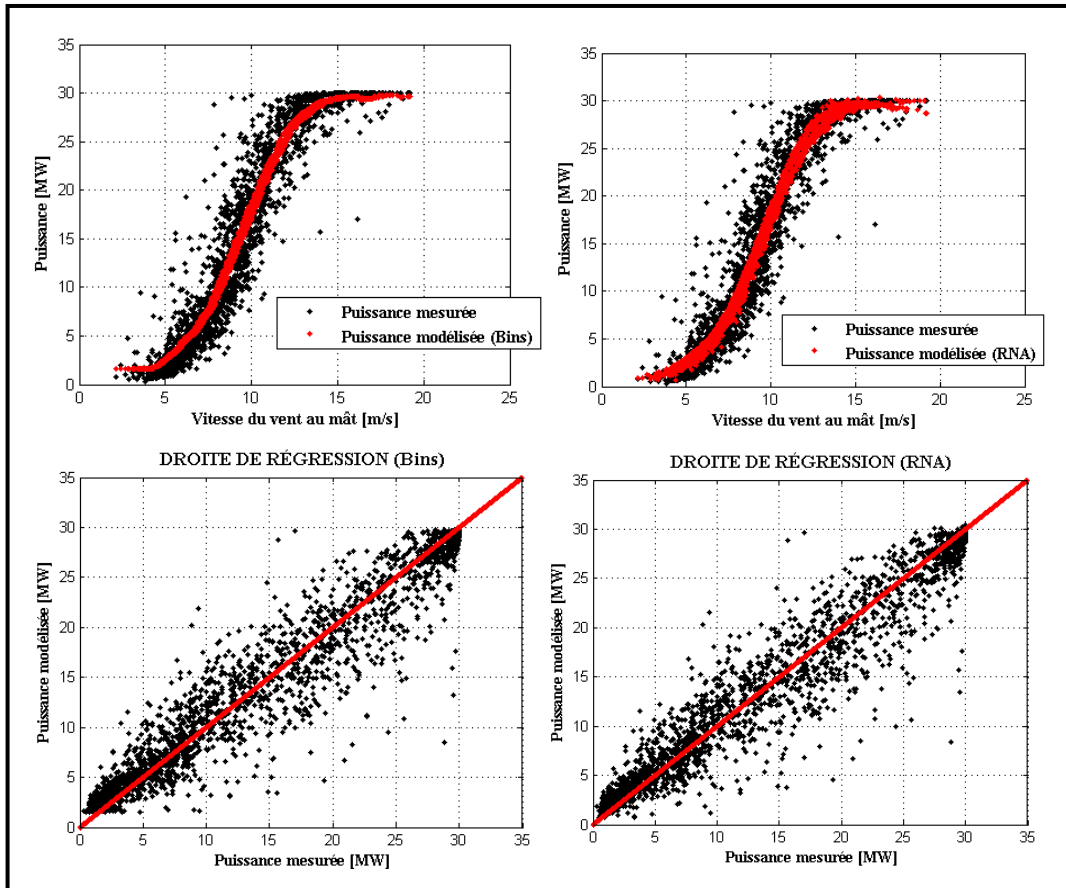


Figure-A II-3 Comparaison des résultats (méthode des Bins vs RNA), Est Point

Tableau-A II-1 Comparaison des performances de deux méthodes :
Bins vs RNA (East Point).

Méthode	NMAE (%)	NRMSE (%)	R ²
Méthode des Bins	6.41	9.46	0.9570
RNA	6.36	9.42	0.9574
Amélioration (%)	0.78	0.42	0.004

3. Comparaison de deux différentes méthodologies en utilisant la méthode des Bins: Vitesse du vent lue au mât vs Moyenne des vitesses du vent aux éoliennes

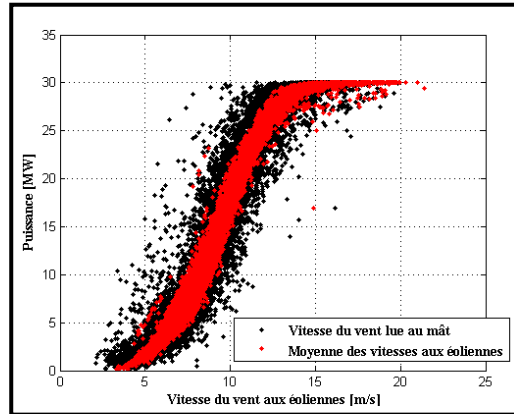


Figure-A II-4 Comparaison des courbes de puissance de la centrale (vitesse lue au mât vs moyenne des vitesses aux éoliennes), East Point

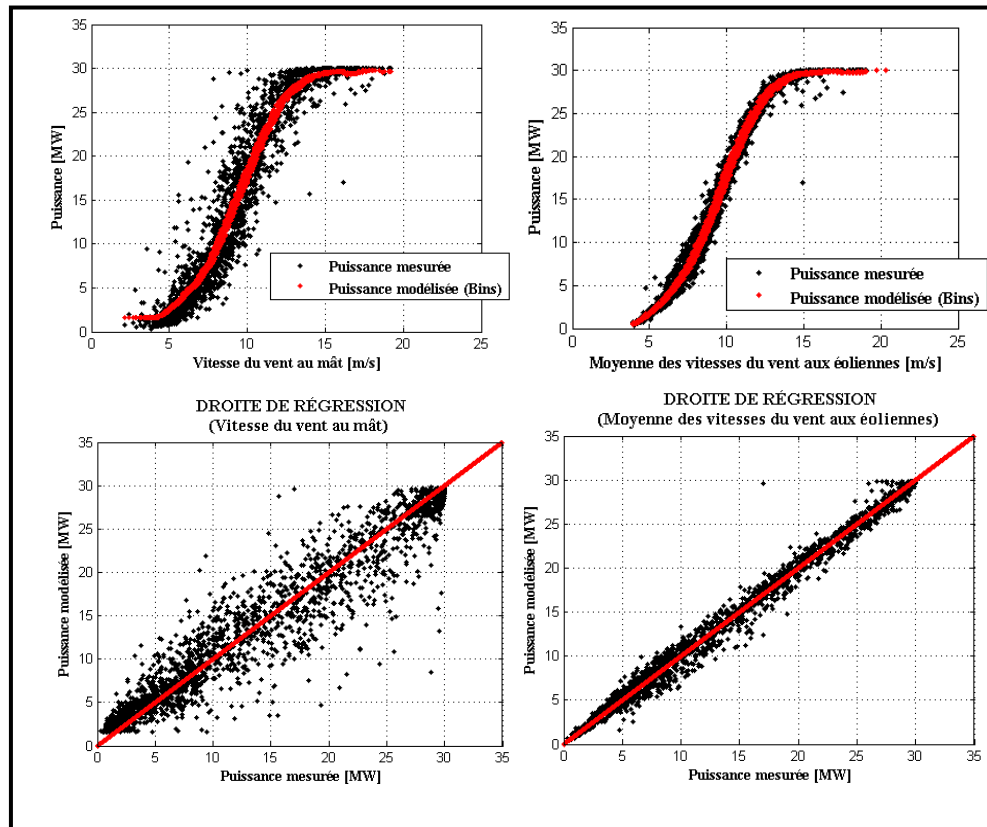


Figure-A II-5 Comparaison des résultats de deux différentes méthodologies (vitesse lue au mât vs moyenne des vitesses aux éoliennes), East Point

Tableau-A II-2 Comparaison des performances de deux méthodologies en utilisant la méthode des Bins : vitesse du vent au mât vs moyenne des vitesses aux éoliennes (East Point).

Méthodologie	NMAE (%)	NRMSE (%)	R ²
Vitesse lue au mât	6.41	9.46	0.957
Moyenne des vitesses aux éoliennes	2.24	3.18	0.995
Amélioration (%)	65.05	66.38	0.038

4. Courbe de puissance proposée

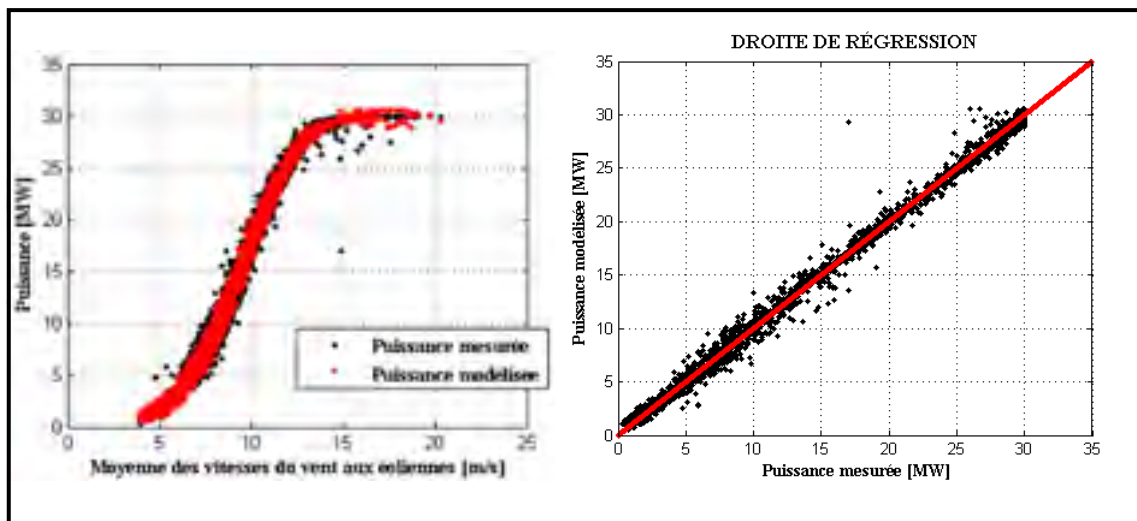


Figure-A II-6 Comparaison entre la puissance mesurée et la puissance modélisée par la courbe de puissance proposée (avec les RNA, East Point)

Tableau-A II-3 Performances de la courbe de puissance proposée (East Point)

Critère d'évaluation	Score (%)
NMAE	1.69
NRMSE	2.51
R ²	0.997

BIBLIOGRAPHIE

- Abreu, L.. 2008. « Wind power generation and dispatch in competitive power markets ». PhD. dissertation, Chicago, IL., Illinois Institute of Technology.
- AWS Scientific Inc.. 1997. « Wind resource assessment handbook: Fundamentals for conducting a successful program ». Golden, Colorado: National Renewable Energy Laboratory, 79p.
- Bédard, J.. 2010. « Improvement of short-term numerical wind prediction ». Master's Thesis, Montréal, Qc., École de technologie supérieure, 134p.
- Bédard, J., W.Yu., Y. Gagnon. et C. Masson. 2011. « Development of a geophisic model output statistics module improving short-term numerical wind predictions over complex sites ». *Submitted in the Wind Energy Journal*.
- Benahmed, N.. 2002. « Optimisation de réseaux de neurones pour la reconnaissance de chiffres manuscrits isolés : Sélection et pondération des primitives par algorithmes génétiques ». Mémoire de maîtrise, Montréal, École de technologie supérieure, 137p.
- Bunn, D.. 2000. « Forecasting loads and prices in competitive power markets ». *In Proceedings of the IEEE*, vol 88, no2, pp.163-169.
- Bunn, D., et E.D. Farmer. 1985. « Comparative models for electrical load forecasting ». Wiley, New York, 332p.
- Cabezon, D., I. Marti., M^a J. S. Isidro, et I. Perez. 2004. « Comparison of methods for power curve modeling ». *Session forecasting challenges and solutions*, (March 31, 2004), p. 9.
- CanWEA. 2010. « 2025 - La force du vent : Une stratégie pour le Québec ». In *Canadian Wind Energy Association (CanWEA) 26e annual congress*. Montréal.
- Collins, J., J. Parkes et A. Tindal. 2009. « Forecasting for utility-scale wind farms; the power model challenge ». In *Integration of Wide-Scale Renewable Resources Into the Power Delivery System, 2009 CIGRE/IEEE PES Joint Symposium (29-31 July 2009)*. pp. 1-10.
- Dabin, N. 2008. « Modélisation et optimisation de la production d'un grand parc éolien en climat nordique ». Thèse de Ph.D., Montréal, École de technologie supérieure, 285 p.
- Demuth, H., N. Beale et M. Hagan. 2008. *Neural Network Toolbox™ 6 User's Guide*. The MathWorks, Inc.

- Focken, U., et M. Lange. 2007. « Summary of forecast methodology of Previento ». Oldenburg: Energy & Meteo Systems Gmbh.
- Focken, U., M. Lange et H.-P. Waldl. 2001. « Reduction of wind power production error by spatial smoothing effects ». In *European Wind Energy Conference EWEC'01* (June 2-6, 2001), pp. 822-825. Copenhagen, Danmark.
- Giebel, G., R. Brounword, G. Kariniotakis, M. Denhard et C. Draxl. 2011. « The state-of-the-art in short-term prediction of wind power - A litterature overview, 2nd Edition. ». Riso, DTU . Roskilde, Danmark: ANEMOS.
- Giebel, G., L. Landberg, G. Kariniotakis et R. Brounword. 2003. « State-of-the-art on methods and software tools for short-term prediction of wind energy production ». In *European Wind Energy Conference, EWEC'2003*. Madrid, Spain.
- Hofman, S.. 2007. « Power curves - The effect of environmental conditions ». In *AWEA Wind Speed and Energy Workshop* (September 2007). Portland.
- IEA. 2011. « Clean energy progress report ». Abu Dhabi: International Energy Agency (IEA).
- IEC-61400-12-1. 2005. « Wind turbine-part 12-1: Power performance measurements of electricity producing wind turbines ». International Electrotechnics Commission (IEC).
- Kariniotakis, G., I. Martí et al. 2004. « What performance can be expected by short term wind power prediction models depending on sites characteristics ». In *European Wind Energy Conference, EWEC'04*. London, UK.
- Kariniotakis, G., et al. 2006. « Next generation short-term forecasting of wind power - Overview of the ANEMOS project ». In *European Wind Energy Conference EWEC'06*. Athens, Grece.
- Landberg, L., G. Giebel, H. A. Nielsen, T. Nielsen et H. Madsen. 2003. « Short-term Prediction—An Overview ». *Wind Energy*, vol. 6, n° 3, pp. 273-280.
- Lange, M. 2003. « Analysis of the uncertainty of wind power predictions ». Ph.D. Dissertation, Oldenburg, University of Oldenburg. 127p.
- Lange, M., et U. Focken. 2005. « Physical approach to short-term wind power prediction ». Berlin: Springer, 208p.
- Li, Mengyan, X. Liu, Z. Mi et J. Qian. 2010. « The contrastive study on the modeling of wind turbine power ». In *Computer Application and System Modeling (ICCASM), 2010 International Conference on* (22-24 Oct. 2010). Vol. 5, pp. V5-294-V5-297.

- Li, S.. 1999. « Artificial neural networks applied for wind power estimation and forecast ». Ph.D. dissertation, Texas, Texas Tech University, 140p.
- Llombart, S.J.Watson, D. Llombart, J.M. Fandos et A. Talayero. 2006. « Power curve characterization: Stochastic methods improving the Bin method and using polynomial regression ». In *European Wind Energy Conference and Exhibition (EWEC)*. Athens,Grece.
- Madsen, H. , G. Kariniotakis, H.Aa. Nielsen, T.S. Nielsen et P. Pinson. 2004. « A protocol for standardizing the performance evaluation of short-Term wind power prediction models ». Coll. « Project ANEMOS ». Lyngby, Denmark, 17p.
- Madsen, H., H.A. Nielsen et T.S. Nielsen. 2005. « A tool for predicting the wind power production of off-shore wind plants ». In *Copenhagen Offshore Wind Conference & Exhibition*. Copenhagen, Danmark: Danish Wind Energy Association.
- Manwell, J.F., J.G. McGohan et A.L. Rogers. 2009. « *Wind energy explained*, second edition ». Coll. « John Wiley & Sons Ltd », 689p.
- Martí, I. 2006. « Evaluation of advanced wind power forecasting models. The results of the ANEMOS project ». In *European Wind Energy Association*, sous la dir. de EWEC'06. Athens, Greece.
- Monteiro, C., R. Bessa, V. Miranda, J. Botterud et G. Conzelmann. 2009. « Wind power forecasting: State-of-the-art 2009 ». Argonne, IL: Argonne national laboratory, 216p.
- Montes, E., A. Arnedo, R. Cordón et R. Zubiaur. 2009. « Influence of wind shear and seasonality on the power curve and annual energy production of wind turbines ». In *European Wind Energy Conference, EWEC 2009*. Marseille, France.
- Negnevitsky, M., et P. Johnson. 2008. « Very short term wind power prediction: A data mining approach ». In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE (20-24 July 2008)*. pp. 1-3.
- Nielsen, T. S., A. Joensen, H. Madsen, L. Landberg et G. Giebel. 1998. « A new reference for wind power forecasting ». *Wind Energy*, vol. 1, n° 1, pp. 29-34.
- Nørgård, P., G. Giebel, H. Holttinen, L. Söder et A. Petterteig. 2004. « Fluctuations and predictability of wind and hydropower ». WILMAR Deliverable D2.1. Risø. Danmark. <<http://130.226.56.153/rispubl/VEA/veapdf/ris-r-1443.pdf>>.
- Pinson, P., et G. Kariniotakis. 2004. « On-line assessment of prediction risk for wind power production forecasts ». *Wind Energy*, vol. 7, n° 2, pp. 119-132.

- Pinson, P. 2006. « Estimation of the uncertainty in wind power forecasting ». Ph.D. dissertation, Paris, École de mines de Paris, 266p.
- Porter, K., et J. Rojers. 2009. « Central wind power forecasting programs in north America by regional transmission organizations and electric utilities ». Columbia, Maryland: Exerter Associates, Inc., 24p.
- Saleck, N., et L. Bremen. 2007. « Wind power forecast error smoothing within a wind farm ». *Journal of Physics*, vol. 75, p. 9.
- Siebert, N.. 2008. « Development of methods for regional wind power forecasting ». Ph.D. dissertation, Paris, École des mines de Paris, 287p.
- Tambke, J., M. Lange, U. Focken, J. Wolff et J. Bye. 2005. « Forecasting offshore wind speeds above the north see ». *Wind Energy* 8(1), pp. 3-16.
- Touani, I., W.Yu., Y. Gagnon. et C. Masson. 2010. « Contribution to the Development of Wind Power Production Forecasting Capacity ». In *26th Annual Conference and Trade Show*. Montréal, Canada: Canadian Wind Energy Association (CanWEA).
- Wagner R., S. M. Courtney, J. T. Larsen et S. U. Paulsen. 2010. « Simulation of shear and turbulence impact on wind turbine performance ». Coll. « Risø-R-1722(EN) ». Risø DTU.