

RÉPUBLIQUE TUNISIENNE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITÉ DE TUNIS EL MANAR



المدرسة الوطنية للمهندسين بتونس

École Nationale d'Ingénieurs de Tunis

THÈSE

Présentée en vue de l'obtention du diplôme de

Docteur en Génie Électrique

Par

Ali LAHOUAR

Ingénieur en Électronique Industrielle, ENISo, 2011

Intitulée

DÉVELOPPEMENT DES TECHNIQUES DE PRÉVISION
DE LA CHARGE ET DE LA PRODUCTION DANS
LE SMART GRID : APPLICATION AU CONTEXTE TUNISIEN

A été dûment acceptée le 06 avril 2016 par le jury constitué de

Président : M. Mohamed ELLEUCH	Professeur, ENIT
Rapporteur : Mme Meriem JAIDANE	Professeur, ENIT
Rapporteur : M. Amine LAHYANI	Maître de conférences, INSAT
Examineur : M. Joseph HAGGÈGE	Maître de conférences, ENIT
Directeur de thèse : M. Jaleddine BEN HADJ SLAMA	Professeur, ENISo
Invitée : Mme Wided MAALEL	Directrice DMT, STEG

Systèmes Avancés en Génie Électrique (SAGE), ENISo

DÉDICACE

*À mon cher père,
À ma chère mère,
À ma chère sœur,
À tous les membres de ma famille, étroite et élargie,
À tous ceux qui m'ont appris un jour quelque chose,
Et à mes chers amis décédés Naïm Korbi et Bechir Jedidi,
Qu'Allah leur accorde sa miséricorde.*

REMERCIEMENTS

Je tiens à exprimer toute ma gratitude à M. Jaleleddine Ben Hadj Slama qui a dirigé mon travail de thèse depuis octobre 2012. Je lui suis éternellement reconnaissant pour son aide, sa disponibilité, sa patience et la confiance qu'il m'a accordée depuis mon arrivée au monde de la recherche scientifique.

Je remercie chaleureusement Mme Meriem Jaidane et M. Amine Lahyani qui ont accepté être les rapporteurs de ce mémoire, pour leurs conseils et leurs remarques pertinentes. Je remercie également M. Mohamed Elleuch le président de jury et M. Joseph Haggège l'examineur, pour le temps et les efforts qu'ils ont consacrés.

Je remercie vivement M. Sadok Ben Daly et M. Yaacoub Fajraoui de la Société Tunisienne de l'Électricité et du Gaz (STEG) pour l'aide qu'ils m'ont apportée et les données qu'ils m'ont fournies, et qui étaient d'une très grande utilité dans cette thèse. Je remercie par la même occasion mon collègue Maher Soyah qui m'a mis en contact avec eux.

Je tiens à exprimer également toute ma gratitude à Mme Najoua Essoukri Ben Amara qui m'a accueilli dans son unité de recherche, ainsi qu'à tous mes camarades doctorants à l'École Nationale d'Ingénieurs de Sousse pour leurs encouragements, en particulier mes prédécesseurs Mohamed et Walid pour leurs précieux conseils.

Je suis reconnaissant envers tous ceux qui ont contribué de près ou de loin à ce travail même par des conseils ou des remarques, en particulier le professeur Ali Alrabti de l'université de Tripoli, et le docteur Kasim Al-Salim de l'université de Strathclyde à Glasgow, que j'ai rencontrés dans des conférences internationales.

Enfin, je remercie toute ma famille qui m'a soutenu et encouragé tout au long de ces années, en particulier mes parents envers lesquels je suis toujours reconnaissant.

RÉSUMÉ

Ce mémoire de thèse vise à présenter et à développer les différentes techniques utilisées dans le smart grid pour la prévision de la puissance électrique, demandée par les consommateurs ou générée par des sources de production renouvelables. Ce domaine de recherche implique également la prévision des conditions météorologiques qui influent sur la production renouvelable, ainsi que la prévision des prix de l'électricité, généralement associée à la demande dans les marchés dérégulés. La prévision, qui s'avère être un outil indispensable pour les gestionnaires des réseaux d'énergie, est étudiée en détails en tant qu'outil mathématique. La revue bibliographique synthétise ses différentes approches, ses modèles élaborés ainsi que ses utilisations potentielles.

La finalité de l'étude est de proposer des modèles de prédiction sophistiqués basés sur l'intelligence artificielle et les méthodes d'optimisation, pour un horizon de prévision allant d'une heure à une journée à l'avance. Les méthodes proposées sont raffinées par rapport au contexte tunisien ; en prenant en considération le comportement du consommateur, l'agencement des éoliennes dans une ferme éolienne tunisienne et ainsi de suite. Cependant, la généralisation et l'universalité des modèles ne sont pas perdues, en réalisant des comparaisons avec des réseaux d'électricité étrangers. Les modèles sont validés par des tests sur des données historiques réelles du réseau tunisien. Les résultats trouvés reflètent la performance de la démarche adoptée, qui n'est pas exigeante en termes de temps de calcul et de ressources. Des perspectives pour le réseau tunisien sont aussi proposées en matière de prévision des prix et d'estimation des coûts de production de l'électricité. Un effort particulier est attribué à la prévision de la demande durant les fêtes religieuses hégiriennes, dans le but d'adopter le modèle à la culture tunisienne et islamique en général.

ABSTRACT

This thesis aims to present and develop various techniques used in the smart grid for forecasting electric power, consumed by customers or generated by renewable generation sources. This field of research also involves predicting weather conditions affecting renewable generation and forecasting electricity prices, usually associated with demand forecast in the deregulated markets. The prediction which is an essential tool for managers of energy networks is studied in detail as a mathematical tool. The literature review summarizes its different approaches, its sophisticated models and its potential uses.

The purpose of the study is to provide sophisticated prediction models based on artificial intelligence and optimization methods for a forecast period ranging from one hour to one day in advance. The proposed methods are refined according to the Tunisian context: by taking into account consumer behavior, layout of turbines in a Tunisian wind farm and so on. However, generalization and universality of the models are not lost, by making comparisons with foreign electricity networks. The models are validated by tests on real historical data from the Tunisian network. The results reflect the performance of the approach, which does not require much computing time and resources. Prospects for the Tunisian network are also proposed for forecasting prices and estimating electricity production costs. A special attention is paid to the load forecast during Hijri religious holidays, in order to adopt the model to the Tunisian and Islamic culture in general.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vi
TABLE DES MATIÈRES	vii
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS	xiii
NOMENCLATURE	xv
LISTE DES ANNEXES	xvi
INTRODUCTION GÉNÉRALE	1
CHAPITRE 1 REVUE DE LITTÉRATURE	4
1.1 Introduction	4
1.2 La prévision : définition et généralités	5
1.2.1 Types de prévisions dans un smart grid	5
1.2.2 Relation entre les types de prévision	6
1.2.3 Horizons et pas temporels de prévision	7
1.2.4 Entrées endogènes et exogènes	8
1.3 La prévision : classifications	9
1.3.1 Approches physiques, approches statistiques	9
1.3.2 Méthodes statistiques conventionnelles, méthodes intelligentes, méthodes hybrides	10
1.4 État de l’art	13
1.4.1 Cas de la prévision de la production éolienne	13
1.4.2 Cas de la prévision de la demande	15

1.4.3	Cas de la prévision des prix	17
1.4.4	Cas de la prévision de la production photovoltaïque	18
1.4.5	Récapitulatif de l'état de l'art	19
1.5	Objectifs de la thèse	21
1.6	Développement mathématique des méthodes de prévision	22
1.6.1	Présentation mathématique du problème	22
1.6.2	Modèle de persistance	22
1.6.3	Réseaux de neurones artificiels	23
1.6.4	Machines à vecteurs de support	26
1.6.5	Forêts aléatoires	31
1.6.6	Critères d'évaluation des méthodes	35
1.7	Conclusion	37
CHAPITRE 2 PRÉVISION DE LA PRODUCTION ÉOLIENNE		38
2.1	Introduction	38
2.2	L'énergie éolienne	39
2.2.1	Progression de l'énergie éolienne en Tunisie (2007 – 2013)	39
2.2.2	Présentation de la ferme éolienne Sidi Daoud	40
2.2.3	Intégration de l'énergie éolienne et effet des erreurs de prédiction	45
2.3	Analyse de la production du parc éolien	46
2.3.1	Étude d'une seule éolienne	46
2.3.2	Étude d'un ensemble d'éoliennes	48
2.4	Conception du modèle de prévision éolienne	50
2.4.1	Choix de la méthode et des entrées	50
2.4.2	Conception du prédicteur	51
2.4.3	Données disponibles et prétraitement	53
2.5	Résultats de la prévision du vent	54
2.5.1	Résultats non-optimisés	54
2.5.2	Optimisation de la prévision	58
2.5.3	Prévision d'une à dix heures à l'avance	61
2.6	Conversion vent / puissance	63
2.7	Conclusion	66
CHAPITRE 3 PRÉVISION DE LA DEMANDE		67
3.1	Introduction	67
3.2	Importance et défis de la prévision de la demande	67
3.2.1	Importance et avantages de la prévision précise	68

3.2.2	Défis de la prévision	68
3.3	Conception du modèle de prévision de la charge	69
3.3.1	Analyse du profil de charge	69
3.3.2	Choix de la méthode et des entrées	72
3.3.3	Stratégie adoptée et sélection experte	75
3.3.4	Données disponibles et prétraitement	79
3.4	Résultats de la prévision	79
3.4.1	Résultats directs	79
3.4.2	Améliorations possibles	83
3.4.3	Comparaisons	85
3.5	Conclusion	90
CHAPITRE 4 PERSPECTIVES POUR LE RÉSEAU TUNISIEN : TARIFICATION ET PRÉ-		
VISION DES PRIX DE L'ÉLECTRICITÉ		91
4.1	Introduction	91
4.2	Tarification et prévision des prix : différences et approches utilisées	92
4.2.1	Tarification	92
4.2.2	Prévision des prix	95
4.3	Tarification de l'électricité dans le marché tunisien : approches proposées	97
4.3.1	Première approche	97
4.3.2	Deuxième approche	100
4.4	Tarification dans le marché australien dérégulé	103
4.4.1	Présentation du marché d'énergie australien	103
4.4.2	Analyse des profils	104
4.5	Prévision du prix de l'électricité dans le marché australien	106
4.5.1	Prévision d'une demi-heure à l'avance	106
4.5.2	Prévision d'une journée à l'avance	110
4.6	Conclusion	113
CONCLUSION GÉNÉRALE		114
RÉFÉRENCES		116
ANNEXES		128

LISTE DES TABLEAUX

Tableau 1.1	Divers horizons de prévision	7
Tableau 1.2	Différents types d'entrées	9
Tableau 1.3	Méthodes statistiques conventionnelles	11
Tableau 1.4	Méthodes intelligentes	11
Tableau 1.5	Métaheuristiques et techniques de traitement avancé du signal	12
Tableau 1.6	Récapitulatif de l'état de l'art : comparaison et utilisation des différentes méthodes	20
Tableau 2.1	Les différentes tranches du parc éolien	43
Tableau 2.2	Configuration de la machine SVR	54
Tableau 2.3	Critères d'évaluation de la prévision de la vitesse par la machine SVR	56
Tableau 2.4	Prévision de la moyenne spatiale des vitesses : comparatif de 4 méthodes proposées	57
Tableau 2.5	Prévision de la direction du vent : comparatif de 4 méthodes proposées	58
Tableau 2.6	Configuration de l'algorithme d'optimisation	58
Tableau 2.7	Résultats de la prévision par la machine SVR optimisée en termes de critères d'erreur	59
Tableau 2.8	Critères d'évaluation de la prévision de la puissance éolienne, janvier 2011	65
Tableau 3.1	Les entrées choisies du modèle de prévision	75
Tableau 3.2	Quantification de la performance des types d'apprentissage	77
Tableau 3.3	Erreurs de prévision	83
Tableau 3.4	Comparaison entre les différentes méthodes en termes d'erreur de prévision	89
Tableau 4.1	Tranches de production	99
Tableau 4.2	Tarification proposée en été de type RTP, première approche	100
Tableau 4.3	Tarification proposée de type RTP, deuxième approche	102

LISTE DES FIGURES

Figure 1.1	Neurone artificiel	23
Figure 1.2	Réseau de neurones à trois couches	25
Figure 1.3	Séparation des classes par hyperplan	27
Figure 1.4	Classification par SVM	29
Figure 1.5	Arbre de régression pour la prédiction de la demande	32
Figure 2.1	Évolution des installations et de la production annuelle par type de générateur	40
Figure 2.2	Parc éolien Sidi Daoud, visité le 13 novembre 2013	41
Figure 2.3	Disposition des éoliennes dans le parc, 1 ^{ère} tranche en rouge, 2 ^e tranche en vert et 3 ^e tranche en bleu	42
Figure 2.4	Potentiel éolien du parc en 2011	43
Figure 2.5	Rose des vents au parc en 2011	44
Figure 2.6	Codage des directions du vent	44
Figure 2.7	Courbe de puissance théorique	47
Figure 2.8	Points de production d'une éolienne 330 kW en 2011, en fonction de la vitesse du vent (a), la direction du vent (b), et en fonction des deux (c) . . .	48
Figure 2.9	Différentes mesures de la vitesse du vent (a), points de production du parc éolien en 2011, en fonction de la vitesse du vent (b), la direction du vent (c), et en fonction des deux (d)	49
Figure 2.10	Modèles de prévision de la vitesse (a) et de la direction (b)	52
Figure 2.11	Détection des fautes et prétraitements possibles	53
Figure 2.12	Résultats de la prévision de la vitesse (a,b) et de la direction (c,d) d'une heure à l'avance par la machine SVR, en janvier 2011	55
Figure 2.13	Organigramme d'optimisation	59
Figure 2.14	Résultats de la prévision de la vitesse (a,b) et de la direction (c,d) d'une heure à l'avance par la machine SVR optimisée, en janvier 2011	60
Figure 2.15	Modèle itératif (dynamique)	61
Figure 2.16	Résultats de la prévision de la vitesse de cinq heures à l'avance par le mo- dèle itératif de la machine SVR, de 250 heures en mars 2011	62
Figure 2.17	Évolution des critères MAE et RMSE	63
Figure 2.18	Structure du réseau de neurones artificiels	64
Figure 2.19	Résultats de la prévision de la puissance éolienne d'une heure à l'avance par la machine SVR et le modèle ANN, de 100 heures en janvier 2011 . . .	65
Figure 3.1	Évolution estimée de la production annuelle d'électricité en Tunisie	70

Figure 3.2	Profil de charge	71
Figure 3.3	Décomposition en ondelettes du signal de charge	72
Figure 3.4	Courbe d'autocorrélation du signal de charge	73
Figure 3.5	Modèle du prédicteur RF	74
Figure 3.6	Processus d'apprentissage en ligne	76
Figure 3.7	Illustration des différences entre les types d'apprentissage, avec la MAPE quotidienne	77
Figure 3.8	Organigramme décrivant la stratégie de prédiction	78
Figure 3.9	Résultats de prévision des jours réguliers par la machine RF	81
Figure 3.10	Résultats de prévision des jours fériés par la machine RF	82
Figure 3.11	Mauvaises prévisions et améliorations possibles, avec la MAPE quotidienne	84
Figure 3.12	Distribution des températures quotidiennes maximales	85
Figure 3.13	Sélection raffinée des ensembles d'apprentissage / test	86
Figure 3.14	Illustration des différences entre les méthodes proposées en termes d'erreur de prévision, du 20 au 26 janvier 2014	88
Figure 4.1	Profil du prix proposé, première approche	99
Figure 4.2	Tarification de l'électricité en Tunisie, à partir du 1 ^{er} mai 2014	101
Figure 4.3	Postes horaires	102
Figure 4.4	Profil actuel du prix en Tunisie (pour l'irrigation)	102
Figure 4.5	Profil du prix proposé, deuxième approche	103
Figure 4.6	Profil de charge et de prix, mois de mars	105
Figure 4.7	Profil de charge et de prix, mois de juillet	105
Figure 4.8	Modèle de prévision, demi-heure à l'avance	107
Figure 4.9	Prévision d'une demi-heure à l'avance, 1 à 3 mars 2013	108
Figure 4.10	Prévision d'une demi-heure à l'avance, 1 à 3 juillet 2013	109
Figure 4.11	Modèle de prévision d'une journée à l'avance	110
Figure 4.12	Prévision d'une journée à l'avance, 1 à 3 mars 2013	111
Figure 4.13	Prévision d'une journée à l'avance, 1 à 3 juillet 2013	112

LISTE DES SIGLES ET ABRÉVIATIONS

ABC	Artificial Bee Colony
ACO	Ant Colony Optimization
AEMO	Australian Energy Market Operator
ANFIS	Adaptive Neuro Fuzzy Inference System
ANN	Artificial Neural Network
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
CART	Classification And Regression Tree
CC	Cycle Combiné
CO	Cycle Ouvert
COV	Covariance
CPP	Critical Peak Pricing
DE	Differential Evolution
DLC	Direct Load Control
DR	Demand Response
DSM	Demand Side Management
EMD	Empirical Mode Decomposition
ES	Exponential Smoothing
FA	Firefly Algorithm
FIS	Fuzzy Inference System
GA	Genetic Algorithm
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GMAPE	Generalized Mean Absolute Percentage Error
ICA	Imperialist Competitive Algorithm
IPP	Independant Power Producer
KF	Kalman Filter
LM	Levenberg-Marquardt
LR	Linear Regression
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
MCP	Market Clearing Price

MIMO	Multiple Input Multiple Output
MR	Multiple Regression
MXE	Maximum Absolute Error
NEM	National Electricity Market
NMAE	Normalized Mean Absolute Error
NWP	Numerical Weather Prediction
OOBE	Out-Of-Bag Error
PER	Persistence Model
PJM	Pennsylvania-New Jersey-Maryland Interconnection
PS	Pattern Search
PSO	Particle Swarm Optimization
PV	Photovoltaïque
RBF	Radial Basis Function
RF	Random Forest
RMSE	Root Mean Squared Error
RTP	Real Time Pricing
SA	Simulated Annealing
SOM	Self Organizing Map
SRL	Slope of Regression Line
STEG	Société Tunisienne de l'Électricité et du Gaz
SVM	Support Vector Machine
SVR	Support Vector Regression
TG	Turbines à Gaz
TLC	Télécommande Centralisée
TOU	Time Of Use
TS	Tabu search
TV	Turbines à Vapeur
VAR	Variance
VI	Variable Importance
WT	Wavelet Transform

NOMENCLATURE

F	transmittance du prédicteur (Forecaster)
X	entrée du prédicteur (de taille m)
m	nombre de signaux d'entrée
Y	signal mesuré
$\hat{Y}$	signal prédit (sortie du prédicteur)
t	temps discret en heures (time)
S_n	ensemble d'apprentissage
n	nombre d'observations dans l'ensemble d'apprentissage
s	horizon de prévision en heures
β	nombre de prévisions
E	signal erreur de prévision (Error)
k	nombre total d'erreurs
v	vitesse du vent (velocity)
d	direction du vent (direction) ou jour à prédire (day) selon le contexte
h	heure (hour)
P_o	puissance électrique produite (Power)
L	puissance électrique demandée (Load)
P	prix de l'électricité (Price)

LISTE DES ANNEXES

Annexe A	L'algorithme Levenberg-Marquardt	128
Annexe B	L'apprentissage par rétropropagation	131
Annexe C	La recherche par motif	135

INTRODUCTION GÉNÉRALE

L'efficacité énergétique, l'énergie renouvelable, la libéralisation du marché et l'intelligence artificielle sont des domaines d'actualité en cette deuxième décennie du 21^e siècle. Ce mémoire de thèse présente le développement des techniques révolutionnaires de prévision dans un réseau électrique intelligent communément appelé « smart grid » qui intègre des sources d'énergie renouvelable, avec un marché de l'électricité qui devient de plus en plus déréglementé.

L'intégration des ressources d'énergie renouvelable, l'apparition des micro-réseaux et la dérégulation du marché de l'électricité dans de nombreux pays du monde, conduisent à une complexité croissante des réseaux d'énergie. De ce fait, les techniques de gestion traditionnelles ne sont plus suffisantes. Les technologies informatiques innovantes sont exploitées dans le smart grid afin d'optimiser les trois phases du cycle de vie de l'énergie électrique ; la production, la distribution et la consommation. Contrairement à un réseau traditionnel où il n'y a qu'un seul producteur, le smart grid met en relation plusieurs producteurs avec plusieurs consommateurs, ce qui crée de nouvelles contraintes de prévision et d'acheminement de l'énergie. Un smart grid a pour objectif de concilier ces différents partenaires afin d'assurer une production optimale ; l'efficacité énergétique est donc un défi majeur.

Le lissage des pics de consommation, en recourant à une prévision précise, est parmi les caractéristiques principales des technologies informatiques. Cette prévision assure entre autres la sécurité du réseau ainsi que la réduction des coûts de production. L'objectif final du smart grid est comme tout réseau électrique : assurer en temps réel l'équilibre entre la production et la demande.

Les centrales dont la production est très irrégulière, notamment solaires et éoliennes, posent un problème suite à leur comportement imprévisible. L'intégration de l'énergie renouvelable plus précisément présente plusieurs défis pour les gestionnaires de réseaux ; en particulier à cause de la variabilité des sources intermittentes, les coûts d'intégration, la réponse en fréquence, l'équilibrage du système, le stockage d'énergie supplémentaire, les contraintes de transmission, le degré de pénétration élevée de la production renouvelable, etc. L'intégration de ces ressources requiert donc une prévision de la production renouvelable aussi précise que possible.

La dérégulation du marché pose une autre problématique ; dans un environnement de concurrence libre entre les différentes entreprises génératrices d'électricité, les coûts de production varient énormément, ce qui conduit naturellement à élaborer des stratégies de prévision des prix. Cette prévision est utile pour les consommateurs comme pour les producteurs, leur permettant d'optimiser leurs affaires.

Pour faire face à tous ces défis, la prévision doit être d'une performance accrue. Beaucoup de décisions opérationnelles dans les systèmes de puissance en dépendent ; telles que la gestion des unités de production, le dispatching économique, le contrôle automatique de la production, l'évaluation de la sécurité, la planification de la maintenance et la commercialisation d'énergie. Le couplage des différents types de prévision est d'une importance cruciale, car les prévisions de la demande, de la production et des prix sont complémentaires et interdépendantes.

Les modèles de prévision sont, dans la plupart des cas, spécifiques à leurs contextes d'application. Et les gestionnaires du réseau en Tunisie ne disposent pas de modèles conçus spécialement pour le contexte tunisien. C'est pour cela que cette thèse propose des modèles spécifiques qui respectent les caractéristiques d'une ferme éolienne en Tunisie, et qui tiennent compte du comportement du consommateur, de la culture, des fêtes religieuses et de l'état du marché propres à la Tunisie.

Ce mémoire de thèse développe les différentes techniques de prévision dans un contexte de smart grid tunisien. Plusieurs méthodes de prévision sont analysées et proposées : d'abord pour la prévision de la production éolienne avec application sur le site de la ferme éolienne tunisienne de Sidi Daoud, et ensuite pour la prévision de la demande sur le réseau électrique tunisien. Enfin, la prévision des prix de l'électricité ainsi que la tarification proportionnelle aux coûts de production sont proposées comme perspectives pour le futur proche. Des procédures de tarification sont étudiées bien que cette thèse ne soit pas orientée économie, moyennant des approches très simplifiées, afin d'améliorer le rendement énergétique du smart grid.

Le mémoire est organisé comme suit. Le premier chapitre déploie la prévision en tant qu'outil mathématique telle que définie dans la littérature, la construction des modèles de prédiction, les horizons ainsi que les différentes classifications des méthodes de prévision, en passant des modèles statistiques conventionnels à l'intelligence artificielle. Une synthèse bibliographique est ensuite proposée afin d'énumérer les différentes méthodologies élaborées par les chercheurs pour chaque contexte de prévision, en présentant les différents outils, les comparaisons et les évaluations. Le deuxième chapitre s'intéresse à la prévision éolienne en Tunisie, en commençant d'abord par la présentation de l'état actuel de cette ressource renouvelable et du parc éolien sujet de l'étude. Ensuite, le comportement individuel et collectif des éoliennes ainsi que les différents facteurs influant sur la production sont étudiés en détails. L'étape suivante consiste à construire le modèle de prédiction éolienne, présenter et évaluer les résultats et proposer des optimisations. Le troisième chapitre s'attaque à la prévision de la demande en Tunisie, en commençant par poser la problématique et les défis, analyser le profil de consommation, choisir les entrées du modèle et éclaircir la stratégie de prédiction adoptée. Ensuite, plusieurs tests de prévision sont effectués dans différentes conditions, dont les résultats sont évalués et comparés à ceux couramment trouvés dans la bibliographie, avant de proposer des améliorations spécifiques au contexte et à la culture tunisienne. Des comparaisons

avec d'autres méthodes ainsi qu'avec un marché étranger sont finalement présentées. Le quatrième et dernier chapitre propose des perspectives de prévision des prix pour le marché de l'électricité dans le réseau tunisien. Il commence d'abord par proposer une procédure de tarification dynamique, corrélée avec la demande et prenant en compte les coûts de production. Par la suite, la prévision des prix dans un marché dérégulé étranger est étudiée, et plusieurs méthodes de prédiction sont appliquées tout en évaluant les résultats trouvés. Finalement, la conclusion générale du mémoire résume les différentes techniques utilisées, souligne les limitations et ouvre les perspectives pour les travaux futurs afin d'apporter des améliorations des résultats et en proposer une meilleure exploitation.

Tout au long de la thèse, différentes méthodes issues de l'intelligence artificielle ont été proposées. Ce qui relie les chapitres entre eux, à part les types des signaux qui sont tous en relation avec le smart grid, c'est le choix de la méthode de prédiction optimale pour chaque chapitre ; étant donné le contexte et les spécificités de la Tunisie. En d'autres termes, il s'agit de développer une approche exploratoire des smart grids en Tunisie, en essayant de trouver la meilleure méthode pour chaque type de prédiction. Aucune méthode n'est toujours optimale dans tous les contextes et quelles que soient les conditions ; donc le choix pertinent et les optimisations éventuelles selon le cas sont les contributions majeures de cette thèse de doctorat. L'analyse de la nature des signaux à prédire, l'étude des différents facteurs externes, le raffinement, le réglage des paramètres et surtout l'évaluation de la performance de chaque méthode sont des procédés communs et présents dans toutes les parties du mémoire.

CHAPITRE 1 REVUE DE LITTÉRATURE

1.1 Introduction

Ce chapitre consacré à l'étude de la littérature vise à donner une vue globale de tout ce qu'il faut connaître et maîtriser avant de s'élancer dans la prévision au sein du smart grid en relation avec le contexte tunisien.

Le secteur d'énergie en Tunisie, surtout l'énergie électrique, est en expansion ; des parcs éoliens installés, de nouveaux générateurs ajoutés pour couvrir la demande qui augmente continuellement, etc. Cette expansion est un appui pour le développement économique du pays. Cependant, ce secteur rencontre quelques difficultés ; notamment la coupure du service durant les pics de consommation, la subvention qui coûte énormément à l'état, le non-paiement des factures, la non-exploitation du potentiel éolien et solaire, etc. De plus, certaines lois adoptées par le gouvernement conduisent à des controverses, telles que la libéralisation de la production et de l'exportation de l'électricité produite par le secteur privé ou les services publics à partir des énergies renouvelables. Suite à tous ces problèmes, plusieurs questions se posent. Est-il possible de mieux intégrer la ressource éolienne dans le parc de production national, en concevant une méthode de prédiction empirique et adaptée aux parcs éoliens tunisiens ? Et d'autre part, le profil de consommation tunisien est-il modélisable, et facilement prédictible pour en construire un modèle de prédiction ? Est-ce que cette prédictibilité est utile et est-ce qu'elle permet d'éviter les pics de consommation et le gaspillage d'énergie ? Les méthodes de prévision actuelles sont-elles efficaces ? Quelle relation existe entre la prévision de la production et celle de la demande ? Est-il possible d'adoucir la tarification de l'électricité pour prendre en compte les coûts de production et la contribution renouvelable, et l'utiliser pour agir sur le comportement des clients ? Et finalement, est-ce que la prévision des prix de l'électricité est un problème qui se pose en Tunisie, ou c'est une approche futuriste pour une stratégie énergétique potentielle ? Cette thèse se charge de répondre à ces questions. Mais avant ceci, il est nécessaire de définir la prévision, présenter ses types en relation avec le smart grid et l'état de l'art actuel à ce niveau ainsi que les méthodes utilisées, et c'est l'objectif de ce chapitre.

Afin d'assurer une meilleure lisibilité, ce chapitre est divisé en cinq parties. La première partie définit la prévision, déploie ses différents types et horizons dans le cadre du smart grid, et définit les différentes entrées d'un modèle de prédiction. La deuxième est consacrée à la présentation des différentes méthodes ainsi que leurs classifications. Dans la troisième partie et qui est la plus importante, l'état de l'art est présenté en se basant sur une synthèse bibliographique contenant les travaux scientifiques les plus reconnus. Il s'agit d'analyser pour chaque type de prévision les

résultats, les différentes comparaisons réalisées ainsi que les conclusions auxquelles les auteurs de chaque article sont parvenus. La quatrième partie définit les objectifs et les finalités de cette thèse et recense les contributions qu'elle vise à apporter à la communauté scientifique. Finalement, la dernière partie de ce chapitre présente les outils mathématiques indispensables qui permettent d'appliquer les différentes méthodes de prédiction utilisées dans ce mémoire, ainsi que leurs critères d'évaluation.

1.2 La prévision : définition et généralités

Cette section présente tous les aspects théoriques de la prévision en tant qu'outil mathématique. Elle est basée sur une étude bibliographique détaillée. La prévision est la traduction française du terme technique forecast, qui signifie l'estimation du futur, caractérisée par un caractère aléatoire. Dans la littérature, elle est confondue avec la prédiction, qui se base sur des modèles mathématiques déterministes [1]. Dans ce mémoire, les deux notions sont confondues, et il n'y a aucune différence entre les deux. Donc la prévision (ou la prédiction) est l'estimation de l'évolution future d'une grandeur quelconque, sur un intervalle de temps bien déterminé. Cette évolution dépend en général d'un certain nombre de facteurs. C'est pourquoi la prévision est souvent une régression, c'est-à-dire une étude de l'évolution d'une variable dépendante d'un ensemble d'autres variables. Dans un smart grid, la grandeur à prédire est généralement une puissance électrique, produite ou demandée. Ce signal étant chaotique, il n'est pas possible d'en trouver une représentation mathématique en fonction du temps, ce qui fait qu'il soit considéré comme processus stochastique [2]. La prévision s'applique souvent sur un signal discret, qu'on appelle également série temporelle (time series), c'est-à-dire une succession de valeurs de ce signal équitablement espacées dans le temps. La machine qui effectue la prévision est appelée prédicteur [3].

1.2.1 Types de prévisions dans un smart grid

Dans un smart grid, deux types de prévision s'imposent : la prévision de la production d'électricité, et la prévision de la demande d'électricité, ou autrement dit : la prévision des ressources et de la demande [2].

La prévision de la production est relativement récente. Il n'était pas nécessaire de prédire la production des centrales électriques avant l'introduction des ressources renouvelables, vu que la production est connue à l'avance tant que le carburant est disponible. Mais depuis l'apparition des ressources de production intermittentes, notamment solaire et éolienne, ce n'est plus le cas. La production de ces ressources est radicalement liée aux conditions météo qu'on ne peut pas contrôler et qu'il faut absolument prédire [4].

La prévision de la demande est beaucoup plus ancienne, puisqu'elle a suivi l'évolution des installations électriques depuis des décennies. Elle arrive donc à la maturité. Néanmoins, elle devient de plus en plus difficile, à cause de la dérégulation du marché de l'électricité et l'intégration des ressources renouvelables [5]. La prévision des prix d'électricité est généralement couplée à la prévision de la demande dans un marché dérégulé [6].

Cette thèse propose alors trois types de prévision : la prévision de la production éolienne du parc Sidi Daoud en Tunisie, la prévision de la consommation d'électricité en Tunisie également, et la prévision des prix d'électricité. Puisque les prix actuels ne varient pas en fonction du temps, une procédure de tarification a été proposée au départ, avant de passer à la prévision proprement dite.

1.2.2 Relation entre les types de prévision

Relation entre la production éolienne et la demande

La production éolienne a une forte corrélation positive avec la charge électrique bien plus que la production traditionnelle. En conséquence, il est viable de traiter la production du vent comme une charge négative et d'intégrer son analyse d'incertitude dans le cadre de l'incertitude de la prévision de la demande. Par conséquent, le concept de la demande nette a été employé dans les études d'intégration de l'éolien afin d'évaluer l'impact de la variabilité de la charge et de la production éolienne sur les opérations du système de puissance. En règle générale, la demande nette a été définie comme suit :

La demande nette = charge électrique totale - production d'énergie renouvelable + échange net [5]

Cette thèse traite la prévision de la production d'énergie renouvelable, puis la prévision de la charge électrique totale. L'échange a été négligé afin de pouvoir estimer la demande nette. Il est noté cependant, que la production renouvelable n'est que très peu intégrée jusqu'à maintenant, et que sa contribution ne dépasse pas 4 %.

Relation entre la demande et le prix

Dans le smart grid, les clients sont encouragés à changer leurs habitudes de consommation d'énergie selon les prix d'électricité. En fait, dans cet environnement, le prix et la consommation de l'électricité sont fortement corrélés, et les participants au marché doivent en prendre compte dans leurs prises de décision afin de maximiser leur profit. Les grandes corrélations entre les partenaires du marché sont à l'origine du besoin d'outils fiables pour maximiser leur bénéfice, en recourant à des outils performants de prévision de la charge et du prix. D'ailleurs, plusieurs papiers traitent la prévision de la consommation et du prix en même temps [7].

Dans cette thèse, la tarification de l'électricité proposée tient compte de la prévision de charge déjà effectuée. Et le modèle de prévision des prix tient compte également de la charge actuelle, passée et future (prédite).

1.2.3 Horizons et pas temporels de prévision

La prédiction est effectuée à un intervalle de temps à l'avance (futur) appelé horizon de prévision. Selon la longueur de cet intervalle, on définit : le court terme, le moyen terme et le long terme.

Ces notations sont en fait très relatives ; le tableau 1.1 donne la signification, tirée à partir de plusieurs papiers scientifiques, de chaque terme selon le type de la prévision. Cependant, ces définitions ne sont pas uniques et elles peuvent varier selon les chercheurs.

Horizon	Prévision de la production	Prévision de la demande
Court terme	1 – 6 heures	1 heure – 1 mois
Moyen terme	6 heures – 3 jours	1 mois – 1 année
Long terme	3 – 10 jours	1 année – 30 années

Tableau 1.1 Divers horizons de prévision

La prévision de la production renouvelable est limitée dans le temps, puisqu'elle dépend des conditions météo qu'on ne peut pas prédire sur plus de dix jours à l'avance. Ainsi, le long terme pour ce type de prévision s'arrête à 10 jours [8]. Cependant, d'autres chercheurs considèrent le moyen terme : une à deux semaines à l'avance avec production quotidienne ; pour la planification et l'optimisation, et le long terme : anticipation d'une à plusieurs années à l'avance avec production mensuelle ou annuelle ; pour sélectionner les sites potentiellement intéressants et évaluer leurs ressources [9]. Par contre, la prévision de la demande ne pose pas ce problème ; le profil de consommation d'électricité peut se projeter sur plusieurs années dans le futur, c'est pourquoi la prévision à long terme de la demande s'étend parfois sur 30 ans. On trouve aussi dans la littérature le très court terme, qui est inférieur à une heure à l'avance pour les deux types de prévision [10].

L'horizon de prévision se décompose en un ou plusieurs intervalles appelés pas temporels de prévision. Selon la longueur de l'horizon, le pas temporel peut être 10 minutes, une heure, une matinée, une soirée, un jour, 3 jours, etc. Si le signal à prédire est une puissance électrique ou une condition météo (température, vent, irradiation solaire...), le prédicteur doit fournir pour chaque pas temporel soit un point, soit la moyenne, soit l'extremum ; le maximum (pointe) ou le minimum (creux).

Si le signal à prédire est une énergie, le prédicteur doit fournir l'énergie, produite ou consommée, durant chaque pas temporel (puisque l'énergie est une intégration par rapport au temps). En général,

on ne prédit des points que pour le court terme, où l'exactitude de la prévision le permet. Par contre, pour le moyen et le long terme, il est judicieux de prédire des moyennes, des extrema ou de l'énergie, d'autant plus que l'horizon de prévision s'agrandit et la qualité de prévision se dégrade [11]. Il est important de noter que le signal à prédire est déterminé point par point dans tous les travaux qui suivent. La prévision probabiliste sort du cadre de cette thèse [12, 13].

Les travaux effectués dans cette thèse sont limités au court terme, une à dix heures à l'avance pour la prévision éolienne, et une heure à 24 heures pour la prévision de la charge et du prix, ainsi que la tarification. Le pas temporel est toujours une heure (et occasionnellement une demi-heure). Le modèle de prévision prédit la moyenne horaire dans le cas de l'éolien (moyenne de 6 valeurs espacées de 10 minutes), et des points dans tous les autres cas.

1.2.4 Entrées endogènes et exogènes

On appelle désormais entrées du prédicteur, les facteurs qui influent sur le signal à prédire. Ces entrées sont endogènes (internes) ou exogènes (externes) [5].

Les entrées endogènes sont des entrées extraites de l'évolution passée du signal à prédire ; des mesures, des moyennes, des extrema, de l'énergie, etc. L'intégration des entrées endogènes se base sur la supposition suivante : l'évolution future du signal dépend de son évolution passée. Cette supposition est fondamentale dans toute cette étude, que ce soit pour prédire une puissance, une énergie, une condition météo ou un prix d'électricité. Au moins une entrée endogène est exigée pour construire un prédicteur.

Les entrées exogènes sont les facteurs externes qui ne sont pas extraits du signal à prédire. Dans le contexte du smart grid et à court terme, ces entrées sont essentiellement les conditions météo et les effets du calendrier (saison, mois, type du jour...). Il est possible cependant de construire un prédicteur sans avoir des entrées exogènes [5].

Les entrées, qu'elles soient endogènes ou exogènes, sont sous la forme de séries temporelles, donc un ensemble de valeurs espacées dans le temps. Le tableau 1.2 présente les entrées les plus communes dans un contexte de smart grid ainsi que l'utilité de chacune d'elles. Tout comme la sortie du prédicteur, les entrées peuvent être des mesures (points), des moyennes, des extrema ou de l'énergie. La période de collecte des entrées varie selon le contexte, mais en général, il est préférable qu'elle soit égale au pas temporel de prévision. Lorsque les entrées sont des mesures, les fréquences de collecte les plus communes sont par 10 minutes, par 30 minutes, par heure, par 6 heures, 3 fois par jour et par jour [5, 12, 14–16].

Pour des périodes de collecte plus larges, il convient d'utiliser des moyennes plutôt que des mesures.

Entrée	Type	Utilité
Demande	Électrique (puissance / énergie)	Les anciennes valeurs servent à prédire les valeurs futures
Production		
Prix	Monnaie	
Température	Météo	Aide à prévoir la consommation d'électricité en général (climatisation, chauffage ...)
Vitesse du vent		Aide à prévoir la production de l'énergie éolienne
Direction du vent		
Irradiation solaire		Aide surtout à prévoir la production de l'énergie photovoltaïque mais aussi la consommation d'électricité en général
Indice de couverture du ciel		
Humidité		
Pression atmosphérique		
Type du jour	Calendrier	Aide à prévoir la consommation d'électricité en général
Saison		

Tableau 1.2 Différents types d'entrées

Les entrées utilisées dans le cadre de cette thèse sont la demande, le prix, la vitesse et la direction du vent, le type du jour et la saison (plus précisément le mois). Le choix de ces entrées sera traité par la suite (chacun dans son chapitre respectif). La production n'a pas été prise comme entrée puisque la méthode abordée est indirecte (cette méthode sera clarifiée par la suite).

1.3 La prévision : classifications

Cette section présente deux critères de classification des méthodes de prévision. On classe généralement les modèles soit selon la nature de l'approche utilisée, soit selon la relation entre les entrées et les sorties.

1.3.1 Approches physiques, approches statistiques

L'approche physique est reliée aux considérations physiques telles que la température, la pression, le relief, l'agencement des éoliennes ou des panneaux dans un parc, etc. En général, elle utilise la

prévision numérique du temps (NWP : Numerical Weather Prediction) qui est une description du comportement de l'atmosphère à travers des équations mathématiques [11]. Lorsque ces équations sont résolues par ordinateur (supercalculateur en général), elles fournissent une simulation accélérée des états futurs de la météo. Les modèles NWP sont nombreux et classifiés en modèles globaux et régionaux [17]. L'approche physique s'applique à la prévision de la production renouvelable et des conditions météo exclusivement.

L'approche statistique consiste à utiliser les séries temporelles. Les anciennes valeurs du signal ainsi qu'un ensemble de facteurs externes (éventuellement) permettent de prédire ses valeurs futures. Cette approche est applicable quel que soit le type de la prévision [12].

Dans cette thèse, toutes les approches sont statistiques, puisqu'on n'a pas accès aux modèles physiques qui nécessitent des supercalculateurs.

1.3.2 Méthodes statistiques conventionnelles, méthodes intelligentes, méthodes hybrides

Cette deuxième classification concerne l'approche statistique uniquement, dans laquelle on trouve les méthodes conventionnelles, les méthodes issues de l'intelligence artificielle et l'hybridation avec les métaheuristiques [18, 19].

Méthodes statistiques conventionnelles

Les méthodes statistiques conventionnelles (tableau 1.3) sont des modèles « boîte blanche » (white-box), dans lesquelles les sorties sont explicitement liées aux entrées à travers des équations mathématiques. Ces méthodes sont simples à utiliser et bien adaptées à la prévision à court terme (quelques heures à l'avance en général). Cependant, elles ne sont pas capables d'apprendre ou de changer leurs stratégies de prédiction. C'est ce manque qui a créé le besoin pour les méthodes intelligentes [5, 11, 12].

Les méthodes ARMA et ARIMA sont appelées les méthodes Box-Jenkins. La méthode ARCH possède une variante appelée GARCH (Generalized ARCH). Les modèles autorégressifs (Box-Jenkins et ARCH) sont appelés les méthodes économétriques. Le lissage exponentiel est aussi appelé la méthode Holt–Winters [19].

Ces méthodes sont présentées afin de présenter l'évolution de la méthodologie de prévision. Néanmoins, elles ne sont pas utilisées excepté le modèle de persistance qui sert toujours comme référence pour les autres modèles.

Abrégé	Dénomination anglaise	Dénomination française
PER	Persistence Model	Modèle de persistance
LR	Linear Regression	Régression linéaire
MR	Multiple Regression	Régression multiple
ARMA	Autoregressive Moving Average	Moyenne mobile autorégressive
ARIMA	Autoregressive Integrated Moving Average	Moyenne mobile intégrée autorégressive
ARCH	Autoregressive conditional heteroskedasticity	Hétéroscédasticité conditionnelle autorégressive
ES	Exponential Smoothing	Lissage exponentiel
KF	Kalman Filter	Filtre de Kalman

Tableau 1.3 Méthodes statistiques conventionnelles

Méthodes intelligentes

Les méthodes intelligentes (tableau 1.4) sont des modèles « boîte noire » (black-box) dont la dynamique interne est inconnue. La relation entre les entrées et les sorties de ces modèles est déterminée à travers un ensemble de règles linguistiques (cas des FIS) ou un processus d'apprentissage (cas des machines apprenantes) [5, 12].

Abrégé	Dénomination anglaise	Dénomination française
FIS	Fuzzy Inference System	Système à inférence floue
ANN	Artificial Neural Network	Réseau de neurones artificiels
SVM	Support Vector Machine	Machine à vecteurs de support
CART	Classification And Regression Tree	Arbre de décision
RF	Random Forest	Forêt aléatoire
ANFIS	Adaptive Neuro Fuzzy Inference System	Système à inférence neuro-floue adaptative

Tableau 1.4 Méthodes intelligentes

Les systèmes flous s'appellent aussi les systèmes experts. L'ANN, la SVM, la CART et la RF sont appelés les machines à apprentissage automatique (learning machine). L'ANFIS est une combinaison des deux (système expert et machine apprenante). L'ANN possède une variante appelée carte auto-adaptative ou carte auto-organisatrice ou encore réseau de Kohonen (SOM : Self Organizing Map). La SVM utilisée pour les problèmes de régression s'appelle communément SVR (Support Vector Regression). La CART fait partie de la RF, qui s'appelle une méthode d'ensemble. Les

méthodes intelligentes sont pour la plupart utilisées comme classifieurs et comme prédicteurs [20]. Parmi ces méthodes, l'ANN, la SVM et la RF sont utilisées pour la prévision de la demande et des prix.

Méthodes hybrides et métaheuristiques

Les méthodes hybrides combinent les méthodes intelligentes et les métaheuristiques ou les techniques avancées du traitement du signal. Les métaheuristiques (tableau 1.5) sont des algorithmes stochastiques qui permettent de chercher une solution suffisamment bonne pour un problème d'optimisation complexe, par échantillonnage d'une fonction objective. Elles sont communément utilisées pour ajuster les paramètres des ANN et SVM, ou carrément pour faire l'apprentissage de ces derniers. D'une façon similaire, les techniques du traitement avancé du signal surtout la transformation en ondelettes (WT) sont utilisées dans les méthodes hybrides [21].

Abrégé	Dénomination anglaise	Dénomination française
GA	Genetic Algorithm	Algorithme génétique
DE	Differential Evolution	Évolution différentielle
PS	Pattern Search	Recherche par motif
PSO	Particle Swarm Optimization	Optimisation par essaims particuliers
ACO	Ant Colony Optimization	Optimisation par colonie de fourmis
ABC	Artificial Bee Colony	Colonie d'abeilles artificielle
FA	Firefly Algorithm	Algorithme des lucioles
SA	Simulated Annealing	Recuit simulé
TS	Tabu search	Recherche tabou
ICA	Imperialist Competitive Algorithm	Algorithme compétitif impérial
WT	Wavelet Transform	Transformée en ondelettes
EMD	Empirical Mode Decomposition	Décomposition en mode empirique

Tableau 1.5 Métaheuristiques et techniques de traitement avancé du signal

Le GA et la DE font partie des algorithmes évolutionnaires (ou évolutionnistes) [21]. L'EMD fait partie de la transformée de Hilbert–Huang [22]. La WT et l'EMD sont des techniques de traitement avancé du signal, tandis que toutes les autres sont des métaheuristiques [23].

La recherche par motif (PS) est utilisée dans le cadre de cette thèse dans le but d'optimiser les paramètres de la SVM utilisée pour la prévision éolienne.

1.4 État de l’art

La prévision dans le smart grid, comme indiqué au début de ce chapitre, varie selon les signaux à prédire. Chaque type de prévision possède ses propres intérêts. Les chercheurs se consacrent en général à la prédiction de la production éolienne et photovoltaïque, de la demande et des prix. La prévision des conditions météo n’est qu’une partie de la prévision de la production. Et parmi les ressources renouvelables, il n’y a que l’éolien et le solaire qui sont intermittents et qu’il faut prédire. Par conséquent, les prévisions de la production éolienne et photovoltaïque, de la demande et des prix sont alors présentées une par une dans cette section, pour en présenter l’état de l’art.

1.4.1 Cas de la prévision de la production éolienne

L’état de l’art de la littérature montre une certaine maturité déjà atteinte dans le domaine de prévision éolienne. Plusieurs références récapitulent les différentes méthodes existantes [12, 13, 17, 21, 24], et elles conduisent toutes vers deux conclusions évidentes : 1. Le vent est une des grandeurs météorologiques les plus difficiles à prédire, et 2. Aucune stratégie n’est idéale ; chaque méthode possède ses propres avantages et inconvénients, et elle peut être adaptée à un contexte et non adaptée à un autre. Les méthodes sont généralement classifiées en approches physiques et approches statistiques, comme indiqué déjà dans la section précédente. Cependant, certains chercheurs classent ces méthodes en quatre familles : méthodes physiques, méthodes statistiques, méthodes de l’intelligence artificielle et méthodes combinées [8]. D’autres considèrent la corrélation spatiale comme un groupe de méthodes [24]. Ces classifications diverses résultent du fait qu’il existe plusieurs critères permettant de catégoriser les différentes méthodes existantes. Considérant par exemple le critère du type des données à prédire, les méthodes peuvent être directes ou indirectes. Les méthodes indirectes prédisent la vitesse du vent (et éventuellement sa direction) et utilisent la courbe de puissance pour en déduire la production. Les méthodes directes prédisent directement la puissance éolienne produite [8, 25]. On trouve dans la littérature les avantages et les inconvénients de chaque approche aussi bien que le contexte approprié pour l’appliquer [21]. On trouve également l’intérêt de combiner plusieurs méthodes et la comparaison des modèles les plus communs parus jusqu’à la fin de 2013 [21]. On montre aussi que l’évaluation des performances de prévision varie selon l’horizon de prédiction, du très court au long terme [26]. La classification des différents intérêts de la prévision éolienne est également d’une grande importance. Ces intérêts varient du contrôle des aérogénérateurs et la planification du dispatching à court terme, à la planification de la maintenance et l’évaluation des potentiels des sites pour des horizons plus larges [25]. Ces différentes tâches mentionnées montrent l’intérêt d’avoir une prévision précise.

L’approche physique contient généralement des modèles NWP qui peuvent être utilisés directement

ou avec post-traitement pour raffiner les résultats. L'utilisation directe consiste à évaluer l'exactitude des sorties du modèle NWP sans aucun traitement postérieur [27]. Le post-traitement permet de raffiner davantage ces sorties, par un apprentissage effectué sur l'historique des mesures [28], par un filtre de Kalman [29,30], ou c'est le modèle NWP qui contient lui-même des procédures de post-traitement, comme le système Skiron [31]. Cependant, cette approche est considérée comme non adaptée au court terme et aux petites surfaces, en plus de la nécessité d'un temps de simulation et des ressources excessifs [8, 21, 32], ce qui n'est pas le cas de l'approche statistique qui se présente comme une alternative. Parmi les méthodes statistiques conventionnelles, on trouve les modèles ARMA et ARIMA qui sont souvent utilisés. La méthode ARMA est parmi les plus anciennes. Elle nécessite un prétraitement ; la transformation et la standardisation des séries temporelles [33]. Le modèle ARIMA, qui est une généralisation de ARMA, est utilisé pour modéliser les processus non stationnaires et donc mieux représenter la vitesse du vent qui est très chaotique [34]. Le modèle de persistance est généralement utilisé pour évaluer la performance des autres ; un modèle quelconque est un bon prédicteur s'il fournit une erreur inférieure à celle de la persistance [35]. Contrairement à la NWP, les méthodes conventionnelles sont bien adaptées au court terme, mais elles sont incapables d'apprendre et de changer leur stratégie de prédiction, puisqu'elles établissent des relations mathématiques strictes entre les entrées et les sorties [8]. Les méthodes intelligentes, surtout les machines d'apprentissage automatique, sont créées pour surmonter ce handicap. Les réseaux de neurones artificiels représentent la méthode la plus utilisée en littérature avec des architectures différentes ; architecture simple [36], des réseaux d'ordre supérieur entraînés par le filtre de Kalman [14], combinés avec la décomposition en ondelettes [37], entraînés par la propagation résiliente [38] ou utilisés avec une mixture de distributions gaussiennes [39], comme ils peuvent être combinés avec la corrélation spatiale [40, 41]. La longue période d'exploitation de cette technique pour la prévision éolienne, allant de 1997 [36] à 2016 [39], a permis d'avoir toutes ces dérivées et combinaisons, et a fait du réseau neuronal une des méthodes les plus performantes et les plus avancées. La machine à vecteurs de support est également exploitée pour la prédiction du vent, dans sa forme la plus simple [42], optimisée par des algorithmes évolutionnaires [43], ou combinée avec l'EMD qui sert à décomposer les séries temporelles complexes [44, 45]. La différence entre l'ANN et la SVM, leurs réglages et leurs limitations seront abordés dans la section 1.4.2 consacrée à la prévision de la demande afin d'éviter la répétition.

Il est à noter que tous les papiers cités ci-dessus ont utilisé des méthodes indirectes. Quant aux méthodes directes, on trouve également l'approche physique [46, 47] et l'approche statistique [48–51]. La méthode directe peut améliorer les résultats en évitant la conversion vent / puissance qui peut induire des fautes, mais diminue l'interprétabilité de la prévision vu que la relation entre le vent et l'énergie est perdue. Il serait donc plus difficile d'expliquer l'éventuel mauvais comportement du prédicteur. Cette méthode n'est pas présentée en détails puisqu'elle n'est pas utilisée dans ce

mémoire.

Même si la production éolienne est reliée essentiellement à la vitesse du vent, d'autres facteurs atmosphériques peuvent aussi avoir leur impact, à savoir la direction du vent, la pression, la température et l'humidité relative [52]. La direction du vent par exemple a été prédite par la méthode des moindres carrés [53] ainsi qu'avec le modèle ARMA [54] en parallèle avec la vitesse. L'effet de la direction du vent est démontrée dans le cas de conversion vent / puissance par réseau de neurones [55], ce qui est utilisé par la suite dans le chapitre 2. L'effet de lissage spatial sera également mis en évidence afin de lisser la vitesse du vent prédite. Le lissage spatial veut dire que la prévision de la production éolienne de plusieurs sites combinés est meilleure que celle de chaque site tout seul, car les erreurs se compensent [12].

En conséquence de tout ce qui a été présenté dans cette section, l'approche physique ne sera pas adoptée puisqu'on n'a pas accès à un modèle NWP, mais aussi parce qu'elle n'est pas adaptée à la prévision de moins de 6 heures à l'avance. Les méthodes statistiques conventionnelles non plus, vu qu'elles sont incapables de traiter la non-linéarité. Les méthodes intelligentes sont insuffisantes dans ce cas aussi, donc on va recourir à une méthode hybride. La méthode choisie consiste à optimiser la machine SVM par l'algorithme PS. La méthode est indirecte ; on prévoit d'abord la vitesse et la direction du vent et on en déduit la production éolienne. De plus, l'effet de la direction du vent et du lissage spatial seront repris en détails.

1.4.2 Cas de la prévision de la demande

Les méthodes utilisées pour la prévision de la demande ne contiennent pas l'approche physique, puisque cette dernière est réservée aux conditions météo.

La famille des méthodes statistiques conventionnelles utilisées pour prédire la demande inclut la régression linéaire simple, la régression multiple, les modèles Box-Jenkins ; ARMA et ARIMA, le lissage exponentiel et le filtre de Kalman. La régression linéaire n'est pas capable de prédire le signal non linéaire, une raison pour laquelle elle est utilisée avec des procédures de classification [56]. La régression multiple par contre, est utilisée toute seule et sans changement pendant une longue durée, allant de 1997 [57] à 2007 [58]. La méthode ARMA donne généralement de meilleurs résultats [59], et ARIMA davantage encore [60] surtout lorsqu'elle est combinée avec la WT [61]. Néanmoins, on a montré que le lissage exponentiel donne le meilleur résultat à très court terme [10]. Le filtre de Kalman est une méthode issue de l'automatique qu'on utilise généralement avec la représentation d'état et l'identification des systèmes [62]. Les méthodes conventionnelles sont simples à implémenter et bien adaptées au court terme, mais incapables de manipuler la non-linéarité existante dans le signal de charge (demande), d'où le besoin pour les méthodes intelligentes.

La famille des méthodes intelligentes, dans le cas de prévision de la demande, contient trois méthodes très utilisées : les systèmes experts (FIS), les réseaux de neurones (ANN) et les machines à vecteurs de support (SVM). À part ces trois méthodes principales, peu d'attention a été accordée à la méthode des forêts aléatoires (RF). Les papiers qui traitent le réseau de neurones et la logique floue sont les plus anciens ; réseau de neurones simple pour la prévision d'une journée à l'avance (1996) [63] et avec prétraitement des données (1997) [64], prévision d'une heure à l'avance (2002) [65], prévision des prix en parallèle avec la demande 2005 [66], prédiction d'une journée à l'avance par les systèmes flous (1999) [67] et prédiction des extrema (2000) [68]. De nos jours, ces deux méthodes (ANN et FIS) ne sont plus utilisées toutes seules, mais moyennant des combinaisons et des optimisations. La machine à vecteurs de support est relativement plus récente, par exemple : prévision par SVM qui tient en compte les effets économiques [69] et prévision SVM par zones géographiques [70]. Cette méthode également est de nos jours utilisée avec hybridation. Cependant, la forêt aléatoire qui est exploitée récemment, demeure l'unique méthode toujours utilisée sans optimisation ni hybridation, grâce à ses propriétés intrinsèques et sa faible sensibilité aux changements des paramètres [71, 72].

Les méthodes intelligentes possèdent l'énorme avantage d'estimation non linéaire. Un réseau de neurones à trois couches est capable d'estimer n'importe quelle fonction continue [73]. Cependant, il a des problèmes de sous ajustement et de surajustement, et il peut tomber sur des optimums locaux (solution optimale, mais localement). La SVM utilise le principe de minimisation empirique du risque pour surmonter les problèmes de l'ANN [73, 74], et c'est ce qui fait toute sa puissance. Les systèmes experts ont l'avantage de fournir une bonne interprétabilité du système [75], tandis que la performance de la RF réside dans sa faible sensibilité aux valeurs de ses paramètres [71]. Tous ces aspects font de ces méthodes des prédicteurs très puissants, mais ont quand même leurs limitations, telles que la détermination d'une architecture optimale (cas des ANN par exemple) et le choix des paramètres. Ces limitations sont dans la plupart des cas franchies par hybridation.

En général, presque toutes les techniques de prévision récentes sont des combinaisons des trois approches principales (FIS, ANN et SVM) et de leurs dérivées, ou des hybridations avec des métaheuristiques ou des techniques de traitement de signal. Par exemple, quelques dérivées des ANN sont les réseaux de neurones « spiking » [76], « abductive » [77], structurés [78], « Elman » récurrents [79], et généralisés [80]. Des hybridations des ANN sont possibles avec le lissage exponentiel [81], la théorie des systèmes gris [82], la transformation en ondelettes [83, 84] et les algorithmes évolutionnaires [85]. Le réseau de Kohonen (SOM) est aussi élaboré pour la prévision [75]. Cette large gamme d'applications différentes fait de sorte que ANN soit la méthode de prévision la plus utilisée. Mais ça ne veut pas dire qu'elle est toujours la plus performante. De même, la SVM peut être utilisée en parallèle avec les algorithmes génétiques [86], la colonie de fourmis [87], la colonie d'abeilles [7], la PSO [88], l'ANFIS [89] et la WT [7, 90]. Les systèmes flous peuvent être égale-

ment combinés avec les algorithmes évolutionnaires [91–93]. Le lissage exponentiel, bien qu’il ne soit pas parmi les méthodes intelligentes, peut être combiné avec la WT et donner des résultats aussi précis que ceux des autres méthodes [94]. Les méthodes hybrides qui combinent l’apprentissage automatique et l’optimisation métaheuristique sont les modèles les plus performants et les plus précis selon plusieurs chercheurs. Néanmoins, les limitations existent toujours, telles que l’exigence du temps de simulation et des ressources, la performance variable en fonction du contexte et des données disponibles, etc.

Toutes les méthodes présentées jusqu’à maintenant pour la prévision de la demande ont été élaborées pour le court terme. Pour les horizons de prévision plus larges, la littérature est beaucoup moins riche. La plupart des chercheurs utilisent l’ANN hybride que ce soit pour le court terme [74,95–97] ou le long terme [98,99]. Dans de tels cas, les facteurs qui pilotent la consommation d’électricité sont significativement différents, comme la valeur industrielle ajoutée, la tarification d’électricité, les prix du pétrole, le nombre d’employés, etc [100].

En conséquence, la méthode utilisée dans cette thèse pour la prévision de la demande est la forêt aléatoire, puisque c’est une méthode peu connue et prometteuse, et surtout parce qu’elle est peu sensible à la variation de ses paramètres, donc elle ne nécessite pas une optimisation. Ce n’était pas possible de l’utiliser pour la prévision éolienne, parce que le vent est un signal beaucoup plus chaotique pour lequel il faut absolument une optimisation. La prévision sera une journée à l’avance par pas horaires (donc court terme), pour la Tunisie entière et non pas par zones, étant donné la faible surface du pays. Les conditions météo utilisées sont celles de Tunis, la capitale et la ville la plus gourmande en énergie.

1.4.3 Cas de la prévision des prix

Cette section est étroitement liée à la section précédente, puisque plusieurs papiers récents traitent la prévision des prix en parallèle avec la prévision de la demande. L’état de l’art n’est alors pas très détaillé ; seuls les aspects qui font la différence sont traités.

Avec la dérégulation du marché d’énergie, le prix d’électricité devient une caractéristique clé. L’objectif fondamental de la dérégulation est de maximiser la génération effective tout en réduisant les coûts. La réponse à la demande effective (DR : demand response) dépend de la capacité de prévision de la charge / production renouvelable côté fournisseurs, et de la prévision du prix côté marché / clients. Avec l’introduction de la tarification en temps réel, cette prévision doit se faire sur des intervalles beaucoup plus courts, d’une heure, demi-heure ou même 5 minutes, puisque des interactions sont introduites entre le prix et la demande d’électricité et le degré de pénétration des énergies renouvelables. Ceci fait appel aussi à des outils qui fournissent des intervalles de confiance au lieu d’estimer un seul point, afin de prendre en compte les incertitudes croissantes. Les

prix d'électricité ont une nature fortement saisonnière, et sont corrélés à la charge demandée [101]. En général, la prévision de la charge et du prix dans les marchés de gros sont des activités mutuellement entrelacées. Les producteurs comme les consommateurs comptent sur la prévision du prix pour préparer leurs stratégies. La prévision de charge dépend essentiellement des facteurs météorologiques, alors que la prévision du prix est pilotée essentiellement par la demande. Les courbes de charge quotidiennes ont des motifs similaires, tandis que le prix montre une très grande volatilité. Il peut sauter à des dizaines ou même des centaines de fois sa valeur normale, comme il peut tomber à zéro ou même à des valeurs négatives. En Australie par exemple, les générateurs offrent au marché des quantités d'électricité à des prix particuliers. Les offres sont soumises chaque 5 minutes. Parmi toutes les offres, le système de gestion du marché sélectionne les générateurs dont il a besoin. Les prix sont calculés chaque 5 minutes, et chaque 6 prix consécutifs sont moyennés chaque demi-heure pour déterminer le prix commercial comptant de chaque région. Donc la prévision des prix est importante pour les producteurs comme pour les consommateurs, vu qu'elle les aide à optimiser leurs affaires [66].

Plusieurs autres facteurs du marché influent aussi sur le prix ; tels que la croissance économique, les conditions météo, et les prix du pétrole. La haute fréquence, la non-stationnarité, la saisonnalité multiple, l'effet du calendrier, la volatilité élevée, le pourcentage élevé des prix inhabituels et le comportement non linéaire font de la prévision précise un vrai défi. Plusieurs techniques ont été développées pour prédire le prix d'électricité ; les réseaux de neurones et leurs modèles étendus, les machines à vecteurs de support, les modèles de Markov cachés, l'ARIMA, la GARCH, la régression linéaire, le filtre de Kalman étendu et l'arbre de régression [102, 103].

Les prix d'électricité en Tunisie ne varient pas en fonction du temps, vu qu'on ne dispose encore pas de compteurs intelligents. Par conséquent, une procédure de tarification en temps quasi-réel (horaire) est proposée dans cette thèse en premier lieu, qui tient compte des coûts de production et de la prévision de la demande. Ensuite, la prévision des prix sera testée dans un marché australien dérégulé dans lequel les prix sont très volatiles, en utilisant l'ANN, la SVM et la RF, de manière similaire à la prévision de la production et de la charge.

1.4.4 Cas de la prévision de la production photovoltaïque

Tout comme la prévision éolienne, il est possible d'utiliser les sorties d'un modèle de prévision numérique de la météo NWP (approche physique). Il est également possible de prédire l'irradiation solaire et d'en déduire la production (méthode indirecte), ou de prédire la production directement (méthode directe). Les fluctuations de l'énergie solaire peuvent causer un déséquilibre entre la production et la demande d'électricité et donc menacer la stabilité du réseau. Deux stratégies sont possibles ; soit prédire la production de chaque unité et faire la somme, soit prédire la production de

toute la zone cible en totalité. Pour modéliser directement la production, des méthodes de régression sont utilisées pour apprendre le lien entre des facteurs météorologiques bien déterminés (sorties d'un modèle NWP) et la production électrique observée. La prévision est requise pour des heures à des journées à l'avance et doit fournir des informations sur la production estimée et le degré d'incertitude pour les périodes volatiles. La variabilité spatiale peut ne pas être négligeable pour des champs solaires de dizaines de kilomètres carrés. Les aspects généraux à considérer dans ce cas sont la régulation des réserves, le suivi de la demande et la planification des tâches. La production des champs photovoltaïques est intrinsèquement variable à cause de la trajectoire du soleil, mais aussi à cause des nuages et d'autres conditions météorologiques qui sont généralement très rapides.

Plusieurs méthodes peuvent être utilisées pour prédire la production des sources d'énergie solaire à des échelles temporelles différentes. Dans les méthodes de prévision traditionnelles, un réseau de neurones utilise les données de jours similaires pour apprendre la tendance. Cependant, apprendre toutes les données de jours similaires n'est pas suffisant si les conditions météorologiques changent brusquement. Le temps qu'un nuage passant peut prendre pour ombrager un système PV dépend de la taille de la ferme PV, de la vitesse du nuage et d'autres facteurs qui peuvent être prédits autrement. Les autres modèles qui peuvent être utilisés sont : le modèle de persistance, les modèles linéaires, l'arbre de régression, le bagging et le boosting, les forêts aléatoires et les machines à vecteurs de support [9, 16].

Cette prévision n'est pas développée dans ce mémoire vu que le réseau tunisien n'intègre encore pas l'énergie photovoltaïque dans son parc de puissance, contrairement à l'énergie éolienne qui est en cours d'exploitation. Pour cette raison, l'état de l'art présenté était très concis.

1.4.5 Récapitulatif de l'état de l'art

Le tableau 1.6 résume tout ce qu'il faut retenir à partir de l'état de l'art qui a été présenté. Il récapitule en gros les avantages et les inconvénients de chaque méthode, ainsi que son utilisation dans ce mémoire.

Méthode	Avantages	Inconvénients	Utilisation dans cette thèse
NWP	Description complète de l'atmosphère Prise en compte des grandeurs physiques Elle fournit les meilleures prévisions pour un horizon de 6 heures et plus	Modèle très compliqué Non adaptée au court terme et aux petites surfaces Temps de simulation et ressources excessifs	Non utilisée
Méthodes statistiques classiques	Simple à utiliser Minimum de paramètres Bien adaptées au court terme	Incapables de modifier leur stratégie de prédiction (ne peuvent pas apprendre) Incapables de traiter la non-linéarité	Non utilisées
ANN	Capable d'apprendre Adaptée à la plupart des problèmes de prévision Plusieurs architectures et combinaisons possibles	Problème de minimum local (problème de convergence) Peut nécessiter une optimisation Nombre de paramètres élevés	Conversion vent / puissance éolienne Prévision des prix (demi-heure à l'avance)
SVM	Capable d'apprendre Pas de minimum local	Peut nécessiter une optimisation Très sensible à certains de ses paramètres	Prévision de la production éolienne (heure à l'avance)
RF	Capable d'apprendre Pas de minimum local Peu sensible aux variations de ses paramètres Ne nécessite pas une optimisation	Non adaptée aux signaux aigus (non lisses)	Prévision de la demande (journée à l'avance) Prévision des prix (journée à l'avance)
Méthodes hybrides (SVM+PS)	Très performantes et optimisées Capables de dépasser la plupart des autres méthodes en termes d'exactitude	Parfois trop compliquées Posent des problèmes de choix des paramètres à optimiser et du meilleur algorithme d'optimisation	Prévision de la production éolienne (heure à l'avance)

Tableau 1.6 Récapitulatif de l'état de l'art : comparaison et utilisation des différentes méthodes

1.5 Objectifs de la thèse

Après avoir présenté l'état de l'art des différentes méthodes de prévision dans le smart grid, il est nécessaire de clarifier les objectifs de la thèse. Les gestionnaires du réseau tunisien utilisent des approches basées sur les réseaux de neurones pour la prévision de la demande. Concernant la prévision éolienne, c'est l'entreprise qui a fabriqué les éoliennes qui s'en occupe ; en utilisant ses propres modèles NWP. Les chercheurs tunisiens ont contribué à leur tour dans le domaine, en proposant des approches probabilistes afin de modéliser la courbe de charge [104], ou pour cerner la variabilité de prédictibilité du pic quotidien [105]. Cette thèse vise à pousser encore loin les méthodes, pas en approchant la courbe de charge par un modèle mathématique, mais plutôt en extrapolant son évolution par des modèles empiriques issus de l'intelligence artificielle, dans le but d'améliorer les résultats de prévision. Les efforts des chercheurs tunisiens dans la prévision éolienne ont conduit à concevoir un prédicteur de deux heures à l'avance dans la ferme éolienne Sidi Daoud, en se basant sur les réseaux de neurones et en considérant des facteurs météorologiques différents [106]. Afin de surmonter les limitations des réseaux de neurones détaillées dans les sections précédentes, cette thèse propose d'utiliser la machine à vecteurs de support et d'optimiser ses paramètres pour le même contexte de prévision. D'autres travaux ont été également réalisés et se sont adaptés au contexte tunisien, pour assurer la bonne intégration des parcs éoliens et estimer les réserves de puissance [107], le traitement et la correction des ensembles de données relatives au parc éolien [108] et la prévision de la production photovoltaïque [109].

La ferme éolienne a besoin d'un modèle de prévision spécifique à son site, son potentiel éolien, l'agencement et les caractéristiques de ses aérogénérateurs, etc. Ce mémoire propose donc dans son deuxième chapitre un modèle de prévision intelligent et optimisé, et qui respecte les caractéristiques du parc éolien Sidi Daoud en Tunisie. D'autre part, la culture tunisienne joue un rôle important dans la consommation d'électricité, ainsi que les traditions et le climat. C'est alors le troisième chapitre qui présente un modèle de prévision de la demande, raffiné par rapport aux fêtes religieuses, le climat chaud, l'utilisation de la climatisation, les caractéristiques du marché et ainsi de suite. L'infrastructure du réseau électrique évolue, et l'utilisation des compteurs intelligents est inévitable dans le futur proche. C'est pour cela que la quatrième chapitre étudie la prévision des prix de l'électricité et propose une procédure de tarification dynamique qui reflète mieux les coûts de production. Toutes les prévisions sont évaluées et comparées à celles de la littérature pour bien se situer par rapport à l'existant.

1.6 Développement mathématique des méthodes de prévision

1.6.1 Présentation mathématique du problème

Le problème de prévision dans cette thèse est présenté comme suit. Soit F un modèle de prédiction où $X = (X^1, \dots, X^m)$ sont ses m entrées et $\hat{Y}$ est sa sortie.

$$\hat{Y} = F(X), X \in \mathbb{R}^m, \hat{Y} \in \mathbb{R} \quad (1.1)$$

On note Y la valeur réelle (mesurée) du signal à prédire, $Y \in \mathbb{R}$. X , Y et $\hat{Y}$ sont en fonction du temps t qui est discret. Le pas temporel est une heure, donc prédire 24 heures à l'avance par exemple revient à déterminer $\hat{Y}(t+1)$ jusqu'à $\hat{Y}(t+24)$.

Soit $S_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$ un ensemble d'apprentissage qui sert à construire le modèle F à partir des n observations (X_i, Y_i) . Notons que les X_i sont des échantillons passés de X . Il est de même pour les Y_i . Soit s le nombre de pas temporels à prédire, autrement dit, la longueur en heures de l'horizon de prédiction. La prédiction revient à déterminer $\hat{Y}(t+1)$ à $\hat{Y}(t+s)$ à partir de $X(t+1)$ à $X(t+s)$. Dans toutes les méthodes qui suivent, on s'intéresse à la détermination de $\hat{Y}(t+1)$ uniquement, car les autres sont déterminés de la même façon. Donc à la fin de chaque méthode, on doit avoir une relation sous la forme :

$$\hat{Y}(t+1) = F(X(t+1)) \quad (1.2)$$

1.6.2 Modèle de persistance

Le modèle de persistance, également appelé approche naïve, est le modèle statistique le plus simple et le plus évident. Les valeurs prédites par un modèle de persistance sur tout l'horizon de prévision sont littéralement égales aux valeurs passées. Par exemple, pour un horizon d'une heure : le modèle prédirait que, dans une heure, la valeur future du signal serait égale à la valeur actuelle. C'est en quelque sorte le décalage parfait ; où la courbe prédite est une version décalée (retardée) de la courbe mesurée. L'intervalle de retard est égal à l'horizon de prévision.

Le modèle de persistance est utilisé comme référence surtout pour la prévision à court terme, où il assure déjà une prévision assez performante. Le défi pour les autres modèles serait toujours de fournir une erreur de prédiction bien plus faible [11]. Mathématiquement, la persistance implique que $\hat{Y}(t+1) = Y(t)$, indépendamment de $X(t+1)$.

1.6.3 Réseaux de neurones artificiels

Le réseau de neurones artificiels est abrégé ANN (Artificial Neural Network). C'est la méthode la plus utilisée en littérature pour la prévision, et elle est parmi les modèles les plus robustes et les plus performants. Seuls les aspects des ANN utilisés dans ce mémoire sont présentés par la suite. Dans le cadre de cette thèse, l'ANN est essentiellement utilisée comme référence pour faire des comparaisons avec d'autres méthodes.

Le neurone formel

Les neurones artificiels sont inspirés du système nerveux en mimant les neurones biologiques ; proposés pour la première fois par Mac Culloch et Pitts en 1943. Ils sont conçus pour reproduire certaines caractéristiques des mémoires biologiques ; à savoir l'apprentissage, le stockage et le traitement des informations incomplètes.

Un neurone artificiel est l'élément de base d'un réseau de neurones. Il réalise une fonction $f(y)$ d'une sommation pondérée y de $n_l + 1$ signaux $x_0, x_1, \dots, x_{n_l}$ qui lui parviennent (figure 1.1).

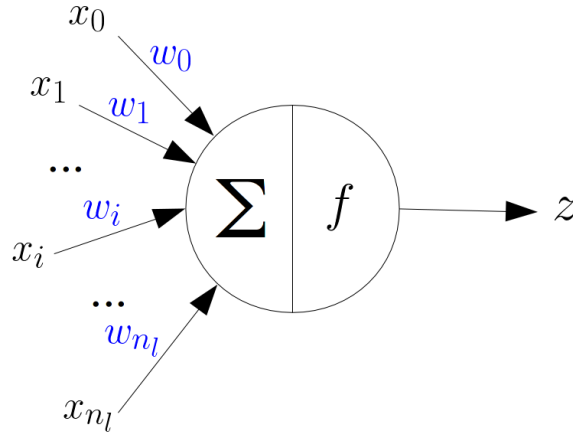


Figure 1.1 Neurone artificiel

$$y = \sum_{i=0}^{n_l} w_i x_i \quad (1.3)$$

$$z = f(y) \quad (1.4)$$

Les fonctions de pondération w_i s'appellent les poids synaptiques. En général, l'entrée x_0 est appelée biais, et son coefficient w_0 est négatif. La fonction f est appelée fonction d'activation [110]. Les fonctions d'activation les plus communes sont :

- L'échelon unipolaire : $f(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases}$
- L'échelon bipolaire : $f(x) = \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x < 0 \end{cases}$
- La sigmoïde unipolaire : $f(x) = \frac{1}{(1 + e^{-x})}$
- La sigmoïde bipolaire (tangente hyperbolique) : $f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$
- La gaussienne (radiale de base) : $f(x) = \exp\left(\frac{-\|x - \mu\|^2}{\sigma^2}\right)$

Où μ est l'espérance mathématique et σ est l'écart type [5].

Le réseau multicouche

On appelle perceptron multicouche un réseau qui contient au moins trois couches de neurones artificiels : une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. L'information circule de l'entrée vers la sortie à travers les couches cachées.

On note $z_i^{(l)}$ la sortie du $i^{\text{ème}}$ neurone de la couche l à n_l neurones, $w_{ij}^{(l)}$ le poids liant le $j^{\text{ème}}$ neurone de la couche $l - 1$ au $i^{\text{ème}}$ neurone de la couche l , et $f_i^{(l)}$ la fonction d'activation du $i^{\text{ème}}$ neurone de la couche l .

$$z^{(l)} = [z_1^{(l)}, z_2^{(l)}, \dots, z_{n_l}^{(l)}]^T \quad (1.5)$$

$$y^{(l)} = [y_1^{(l)}, y_2^{(l)}, \dots, y_{n_l}^{(l)}]^T \quad (1.6)$$

$$z_i^{(1)} = x_i \quad (1.7)$$

$$x = [x_1, x_2, \dots, x_{n_1}]^T \quad (1.8)$$

Le perceptron est représenté par les équations :

$$y_i^{(l+1)} = \sum_{j=1}^{n_l} w_{ij}^{(l+1)} z_j^{(l)} + w_{i0}^{(l+1)} \quad (1.9)$$

$$z_i^{(l)} = f_i^{(l)}(y_i^{(l)}) \quad (1.10)$$

En notant :

$$W^{(l)} = \{w_{ij}^{(l)}\}, j = 1, \dots, n_{l-1}, i = 1, \dots, n_l \quad (1.11)$$

$$w_0^{(l)} = [w_{10}^{(l)}, \dots, w_{n_l 0}^{(l)}]^T \quad (1.12)$$

$$f^{(l)}(y^{(l)}) = [f_1^{(l)}(y_1^{(l)}), \dots, f_{n_l}^{(l)}(y_{n_l}^{(l)})]^T \quad (1.13)$$

On obtient la forme matricielle :

$$z^{(l)} = f^{(l)}(y^{(l)}) \quad (1.14)$$

$$y^{(l+1)} = W^{(l+1)} z^{(l)} + w_0^{(l+1)} \quad (1.15)$$

Cette forme facilite la mise en équations [110].

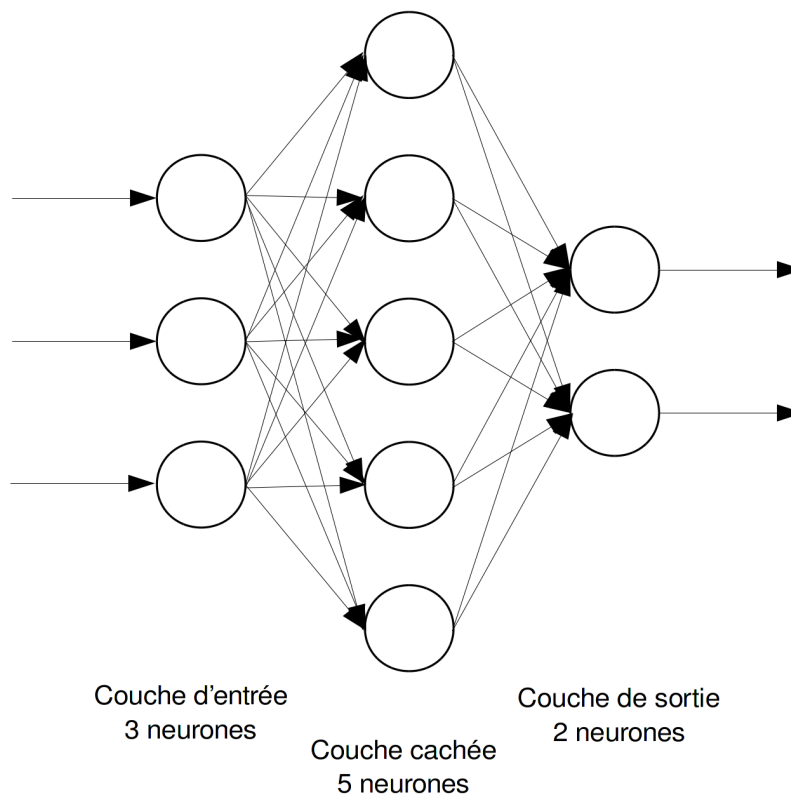


Figure 1.2 Réseau de neurones à trois couches

Le réseau feedforward

Un réseau de neurones à action directe (feedforward en anglais) est un réseau neuronal artificiel où les connexions entre les unités ne forment pas une boucle ou un cycle dirigé, contrairement aux réseaux de neurones récurrents. Il est le premier et le plus simple type de réseau neuronal artificiel conçu. Dans ce réseau, l'information se déplace dans une seule direction, à partir des

nœuds d'entrée, par l'intermédiaire des nœuds cachés (le cas échéant) jusqu'aux nœuds de sortie (figure 1.2) [110].

Types d'apprentissage d'un réseau de neurones

Dans le cas d'un apprentissage non supervisé, des prototypes ou des « patrons » sont présentés au réseau qu'on laisse s'auto-organiser à l'aide des lois locales. Ce mode est aussi appelé apprentissage par compétition.

Dans le cas d'un apprentissage supervisé, on essaye d'imposer au réseau un fonctionnement donné. Ce fonctionnement s'obtient en forçant les sorties du réseau à prendre des valeurs données face aux entrées qui lui sont appliquées, en modifiant les poids synaptiques. Le réseau se comporte alors comme un filtre dont les paramètres de transfert sont ajustés à partir des couples entrées/sorties présents. L'adaptation des paramètres du réseau s'effectue à partir d'un algorithme d'optimisation, l'initialisation des poids synaptiques étant le plus souvent aléatoires [110].

Apprentissage supervisé par rétropropagation

Les couples entrées/sorties imposés sont les observations de l'ensemble S_n . La rétropropagation, donnée par les annexes A et B, sert à déterminer les poids synaptiques w_{ij} à partir de S_n . Autrement dit, elle permet de construire le modèle F . Pour le problème de prédiction présenté dans cette thèse, il n'y a qu'une seule sortie, donc un seul neurone z_1 dans la couche de sortie. Dans le cas d'un réseau à M couches, $\hat{Y}(t+1)$ serait alors $z_1^{(M)}$, et $x = X(t+1)$, avec $n_1 = m$.

1.6.4 Machines à vecteurs de support

La machine à vecteurs de support est une autre méthode issue de l'intelligence artificielle. Elle permet de surmonter quelques limitations des ANN, telles que les optimums locaux. Elle peut être optimisée plus facilement qu'un réseau de neurones, et elle possède moins de paramètres à régler. Elle est introduite dans cette section pour la classification (contexte original) puis pour les problèmes de régression, y compris la prédiction. Dans le reste du mémoire, elle est utilisée pour la prévision de l'énergie éolienne.

Classification

La machine à vecteurs de support est abrégée SVM (Support Vector Machine). Introduite pour la première fois par Vapnik en 1992 [111], elle se base sur deux principes fondamentaux : trouver le meilleur hyperplan séparateur entre deux classes différentes (ou plus) en maximisant la marge entre

les échantillons différents les plus proches, et projeter les échantillons vers une autre dimension lorsqu'ils ne sont pas linéairement séparables, en utilisant une fonction noyau (figure 1.3).

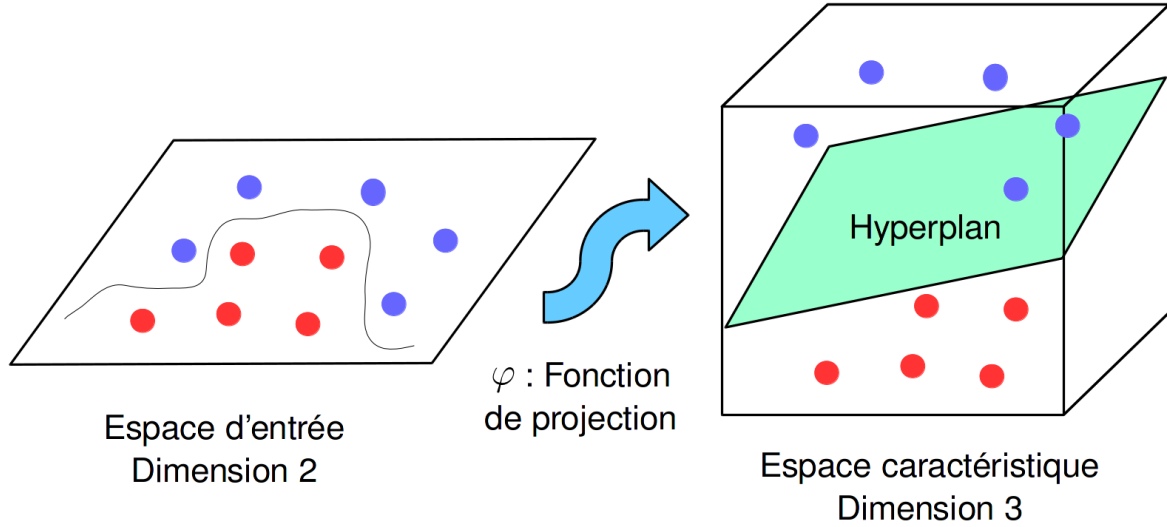


Figure 1.3 Séparation des classes par hyperplan

Dans cette partie dédiée à la classification, $Y \in \{-1, 1\}$. Y sera à nouveau dans $\mathbb{R}$ dans la section suivante consacrée à la régression. (La même chose pour $\hat{Y}$)

Dans un problème de classification binaire (2 classes), il s'agit de séparer un ensemble de vecteurs S_n (ensemble d'apprentissage) :

$$S_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}, X \in \mathbb{R}^m, Y \in \{-1, 1\} \quad (1.16)$$

Par l'hyperplan :

$$\langle w, X \rangle + b = 0 \quad (1.17)$$

Où w est le vecteur de poids, b est un scalaire (biais) et $\langle w, X \rangle$ est le produit scalaire. L'équation $f(X) = \langle w, X \rangle + b$ est celle d'un classifieur linéaire. Mais généralement, les échantillons ne sont pas linéairement séparables, et ce qui conduit à l'utilisation d'un classifieur non linéaire :

$$f(X) = \langle w, \varphi(X) \rangle + b \quad (1.18)$$

Où φ est une fonction utilisée pour projeter X à un espace de dimension plus élevée.

La SVM cherche à résoudre le problème d'optimisation suivant [111] :

$$\min_{w,b,\xi} \left(\frac{1}{2} \langle w, w \rangle + C \sum_{i=1}^n \xi_i \right) \quad (1.19)$$

Sous les contraintes :

$$\left\{ \begin{array}{l} Y_i(\langle w, \varphi(X_i) \rangle + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{array} \right\} \quad (1.20)$$

Où C est le paramètre de pénalité du terme d'erreur, appelé aussi paramètre de régularisation ($C > 0$), ξ_i est une variable qui mesure la mauvaise classification de X_i , et φ est la fonction de projection. Les échantillons situés sur la marge ($\xi = 0$) sont appelés les vecteurs de support (figure 1.4). La projection est définie par une autre fonction K appelée noyau :

$$K(X_i, X_j) = \langle \varphi(X_i), \varphi(X_j) \rangle \quad (1.21)$$

Les fonctions noyaux les plus communes sont les suivantes, où γ , r et d sont des paramètres réglables :

- Linéaire : $K(X_i, X_j) = \langle X_i, X_j \rangle$
- Polynomial : $K(X_i, X_j) = (\gamma \langle X_i, X_j \rangle + r)^d$, $\gamma > 0$
- Radial de base : $K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2)$, $\gamma > 0$
- Sigmoidé : $K(X_i, X_j) = \tanh(\gamma \langle X_i, X_j \rangle + r)$

Le choix d'un noyau ou de l'autre dépend généralement de l'application. Le noyau linéaire n'est utilisé que si les échantillons sont linéairement séparables. Le noyau radial est parmi les plus utilisés et les plus performants. La résolution du problème d'optimisation précédent est équivalente à la résolution du problème dual, défini par :

$$\min_{\alpha} \left(\frac{1}{2} \alpha^T Q \alpha - e^T \alpha \right) \quad (1.22)$$

Sous les contraintes :

$$\left\{ \begin{array}{l} \langle Y, \alpha \rangle = 0 \\ 0 \leq \alpha_i \leq C \end{array} \right\} \quad (1.23)$$

Où $\alpha = [\alpha_1, \dots, \alpha_n]^T$ sont les multiplicateurs de Lagrange, $e = [1, \dots, 1]^T$ est le vecteur des 1 (de

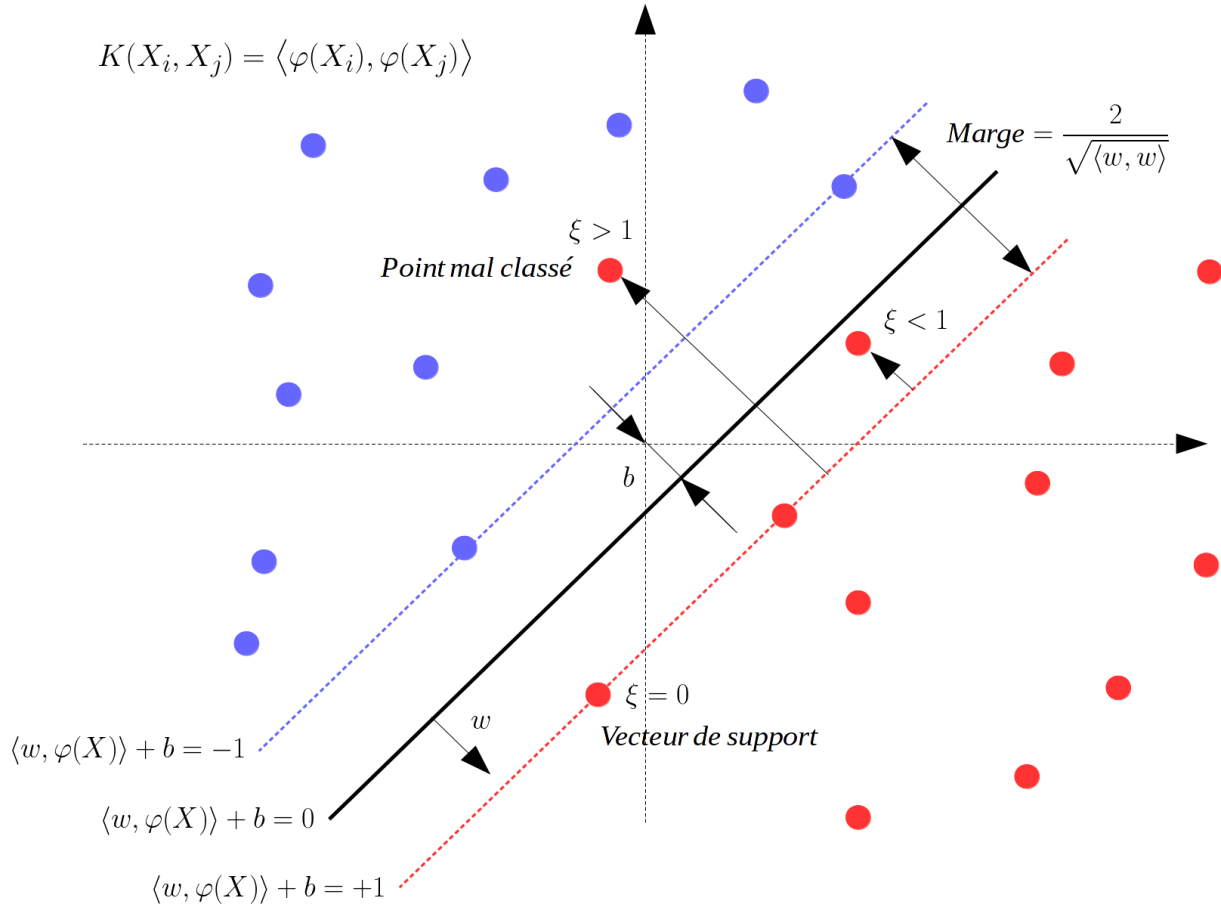


Figure 1.4 Classification par SVM

taille n), et Q est une matrice semi-définie positive telle que $Q_{i,j} = Y_i Y_j K(X_i, X_j)$. Lorsque le problème est résolu, w vérifie :

$$w = \sum_{i=1}^n Y_i \alpha_i \varphi(X_i) \quad (1.24)$$

Le classifieur ainsi construit s'écrit alors sous la forme :

$$f(X) = \langle w, \varphi(X) \rangle + b = \sum_{i=1}^n Y_i \alpha_i K(X_i, X) + b \quad (1.25)$$

La construction du classifieur met fin à l'étape d'apprentissage de la machine SVM. Tout nouveau vecteur X peut désormais être classifié en utilisant $\hat{Y} = \text{sign}(f(x))$ qui prédit sa classe, 1 ou -1 [112].

Régression

La machine SVM possède une variante pour les problèmes de régression, abrégée SVR (Support Vector Regression). Introduite en 1996 [113], cette variante étend la plage de variation de la sortie Y à toute la droite réelle $\mathbb{R}$. Le nouvel ensemble d'apprentissage S_n est alors défini par :

$$S_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}, \quad X \in \mathbb{R}^m, \quad Y \in \mathbb{R} \quad (1.26)$$

Le problème ici s'agit d'approcher l'ensemble S_n par une fonction non linéaire de la forme :

$$f(X) = \langle w, X \rangle + b \quad (1.27)$$

Dans ce cas, le problème d'optimisation devient :

$$\min_{w, b, \xi, \xi^*} \left(\frac{1}{2} \langle w, w \rangle + C \sum_{i=1}^n \xi_i + \xi_i^* \right) \quad (1.28)$$

Sous les contraintes :

$$\left\{ \begin{array}{l} Y_i - \langle w, \varphi(X_i) \rangle - b \leq \varepsilon + \xi_i \\ \langle w, \varphi(X_i) \rangle + b - Y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{array} \right\} \quad (1.29)$$

Les paramètres ξ_i et ξ_i^* sont des variables d'écart représentant les contraintes supérieure et inférieure, et ε est le terme d'erreur ($\varepsilon > 0$). Le problème dual est défini par :

$$\min_{\alpha, \alpha^*} \frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n Y_i (\alpha_i - \alpha_i^*) \quad (1.30)$$

Sous les contraintes :

$$\left\{ \begin{array}{l} \langle e, (\alpha - \alpha^*) \rangle = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq C \end{array} \right\} \quad (1.31)$$

Les variables α et α^* sont les multiplicateurs de Lagrange, e est le vecteur des 1 (de taille n) et $Q_{i,j} = K(X_i, X_j)$.

La résolution de ce problème donne la fonction de régression suivante :

$$f(X) = \sum_{i=1}^n (-\alpha_i + \alpha_i^*) K(X_i, X) + b \quad (1.32)$$

La détermination de la fonction de régression met fin au processus d'apprentissage. Cette fonction permet désormais de prédire la sortie $\hat{Y}$ de n'importe quel nouveau vecteur d'entrée X .

$$\hat{Y}(t+1) = f(X(t+1)) \quad (1.33)$$

Ce type de régression s'appelle la régression ε . Il existe un autre type, appelé régression ν , qui introduit une nouvelle variable ν afin de contrôler le nombre des vecteurs de support [112].

1.6.5 Forêts aléatoires

La forêt aléatoire est une méthode de l'intelligence artificielle tout comme ANN et SVM. C'est une méthode d'ensemble qui combine plusieurs arbres de décisions à travers un algorithme appelé bagging. Beaucoup moins utilisée que les autres méthodes, elle représente l'apport principal de cette thèse. Elle est utilisée pour la prévision de la demande d'électricité, et elle est raffinée pour prendre en compte les spécificités du consommateur tunisien.

Arbres de décision

Un arbre de décision, également appelé arbre de classification et de régression, est abrégé CART (Classification And Regression Tree). C'est un modèle statistique introduit par Breiman en 1984 [114]. Il décrit les différentes classes ou valeurs qu'une sortie peut prendre en fonction d'un ensemble de caractéristiques d'entrée. De façon générale, un arbre est un ensemble de nœuds et de branches organisés selon une hiérarchie sans boucles. Un arbre de décision est un arbre dont les nœuds enregistrent une fonction de test à appliquer aux entrées qui parviennent. Les nœuds terminaux sont appelés feuilles de l'arbre. Chaque feuille stocke le résultat de test final. L'arbre est binaire si chaque nœud interne contient exactement deux branches sortantes ; appelées fils de droite et fils de gauche. La figure 1.5 montre un arbre de régression servant à prédire la demande d'électricité de la Tunisie en 2013, en supposant que la demande est en fonction de la température maximale, le type du jour et la saison. L'arbre de décision est robuste, peu vulnérable face aux entrées non significatives et fournit une bonne interprétabilité.

Pour des raisons de concision, la suite du développement mathématique des forêts aléatoires se restreint aux problèmes de régression. En adoptant la même notation mathématique, soient X le vecteur d'entrée contenant m caractéristiques (composantes), Y la sortie scalaire (ou étiquette) et S_n un ensemble d'apprentissage contenant n observations (X_i, Y_i) .

$$S_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}, X \in \mathbb{R}^m, Y \in \mathbb{R} \quad (1.34)$$

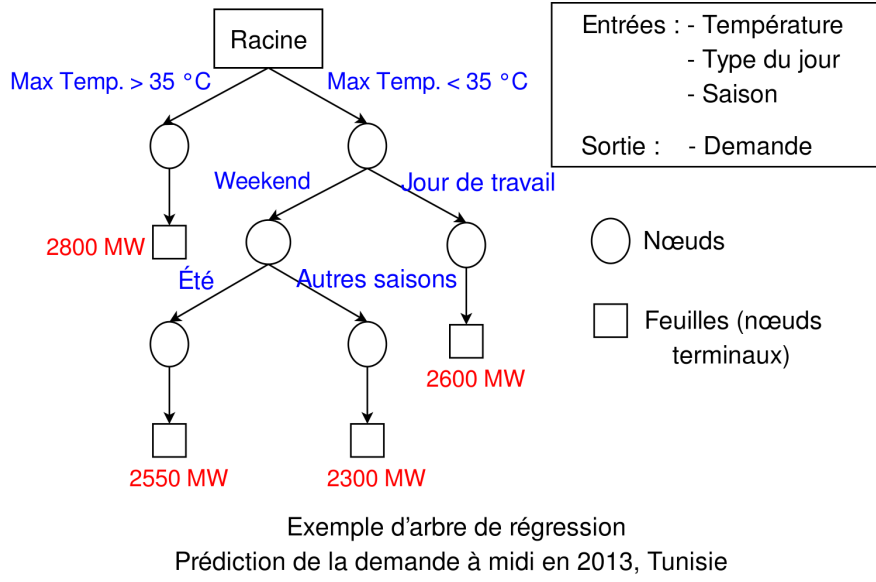


Figure 1.5 Arbre de régression pour la prédiction de la demande

Durant l'apprentissage, un algorithme pilote la scission des entrées à chaque nœud, de sorte que les paramètres des fonctions de découpe soient optimisés pour s'adapter à l'ensemble S_n . Le principe consiste à découper récursivement l'espace d'entrée X en cherchant les sous-partitions optimales. Plus précisément, la première étape de l'algorithme CART consiste à découper au mieux la racine en deux nœuds fils différents selon :

$$\{X^j < d\} \cup \{X^j > d\} \quad (1.35)$$

Où $j \in \{1, \dots, m\}$ et $d \in \mathbb{R}$. Pour sélectionner la meilleure découpe, le couple (j, d) doit minimiser une fonction coût, qui est généralement la variance des nœuds fils. La variance d'un nœud p est définie par :

$$VAR(p) = \sum_{i: X_i \in p} (Y_i - \bar{Y}_p)^2 \quad (1.36)$$

Où $\bar{Y}_p$ est la moyenne des observations Y_i présentes au nœud p . Ensuite, les nœuds fils sont aussi découpés selon la même procédure. Le développement de l'arbre est stoppé par un critère de terminaison. Il est commun d'arrêter l'arbre lorsqu'un nombre maximal de niveaux est atteint, ou lorsqu'un nœud contient moins d'un nombre prédéfini d'observations. À la fin de la phase d'apprentissage, une fonction de prédiction $\hat{h}(X, S_n)$ est construite sur S_n .

La phase de test consiste à déterminer une estimation $\hat{Y}$ de la sortie Y correspondante à une nou-

velle entrée X (nouvelle veut dire non contenue dans S_n) :

$$\hat{Y} = \hat{h}(X, S_n) \quad (1.37)$$

À partir de la racine, chaque nœud applique sa propre fonction de découpe à X . Selon le résultat du test binaire, les données sont envoyées au fils de droite ou de gauche. Ce processus se répète jusqu'à ce que les données atteignent une feuille (un nœud terminal) [114].

Bagging et forêts aléatoires

La forêt aléatoire, abrégée RF (Random Forest) est une méthode d'ensemble qui combine la prédiction de plusieurs arbres de décision [115]. Le principe de base est appelé bagging (bootstrap aggregation) ; où un échantillon de taille n pris de l'ensemble d'apprentissage S_n est sélectionné aléatoirement et adapté à un arbre de régression. Cet échantillon est appelé bootstrap, et il est choisi avec remplacement, c'est-à-dire les mêmes observations (X_i, Y_i) peuvent apparaître plusieurs fois.

Un échantillon bootstrap est obtenu en sélectionnant aléatoirement n observations avec remplacement de S_n , où chaque observation a une probabilité de $1/n$ d'être sélectionnée. Les variables aléatoires indépendantes identiquement distribuées Θ_l représentent cette sélection aléatoire. L'algorithme de bagging sélectionne plusieurs échantillons bootstrap $(S_n^{\Theta_1}, \dots, S_n^{\Theta_q})$, applique l'algorithme CART sur ces échantillons pour obtenir une collection de q arbres de prédiction $(\hat{h}(X, S_n^{\Theta_1}), \dots, \hat{h}(X, S_n^{\Theta_q}))$, et ensuite effectue l'agrégation des sorties de tous ces prédicteurs.

En plus du bagging, la forêt aléatoire sélectionne un nombre prédéfini appelé $mtry$ parmi les m caractéristiques pour la découpe au niveau de chaque nœud. L'algorithme RF essaye de trouver la meilleure découpe uniquement parmi les $mtry$ caractéristiques sélectionnées. La sélection dans chaque nœud est uniforme, donc chaque caractéristique possède la probabilité de $1/m$ d'être choisie. Le nombre $mtry$ est le même pour tous les arbres de prédiction, et il est recommandé qu'il soit égal à la racine carrée du nombre de caractéristiques m .

$$mtry = \lfloor \sqrt{m} \rfloor \quad (1.38)$$

Où l'opérateur $\lfloor x \rfloor$ représente la partie entière de x . Le reste de l'algorithme est semblable à celui du CART, le meilleur couple de découpe (j, d) est obtenu en minimisant une fonction coût, et la procédure continue jusqu'au développement complet de tous les arbres.

L'agrégation est effectuée en moyennant les sorties de tous les arbres. Par conséquent, l'estimation $\hat{Y}$ de la sortie d'un nouveau vecteur d'entrée X s'écrit comme suit :

$$\hat{Y} = \frac{1}{q} \sum_{l=1}^q \hat{h}(X, S_n^{\Theta_l}) \quad (1.39)$$

En l'adoptant pour le problème de prédiction de cette thèse, cette équation devient :

$$\hat{Y}(t+1) = \frac{1}{q} \sum_{l=1}^q \hat{h}(X(t+1), S_n^{\Theta_l}) \quad (1.40)$$

L'avantage principal de l'agrégation bootstrap est l'immunité au bruit, puisqu'elle génère des arbres non corrélés à travers différents ensembles d'apprentissage. Un prédicteur faible (un seul arbre de régression) peut être sensible au bruit, tandis que la moyenne de plusieurs arbres décorrélés ne l'est pas. La sélection d'un sous-ensemble aléatoire de caractéristiques à chaque découpe a le même but de décorrélation des arbres.

Deux caractéristiques principales distinguent les forêts aléatoires : l'erreur out-of-bag (*OUBE*) (erreur en dehors du bootstrap) et la mesure d'importance des variables. La (*OUBE*), également appelée erreur de généralisation, est une sorte de validation croisée interne ; c'est l'erreur de prédiction moyenne des observations vues pour la première fois, c'est-à-dire en utilisant uniquement les arbres qui ne les ont pas vues pendant l'apprentissage. Plus explicitement, pour chaque observation (X_i, Y_i) de S_n , une estimation $\hat{Y}_i$ de Y_i est réalisée en agrégeant uniquement les arbres construits sur des échantillons bootstrap non contenant (X_i, Y_i) . La (*OUBE*) est très utile pour estimer la capacité de généralisation du modèle construit.

$$OUBE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1.41)$$

La mesure d'importance des variables est calculée en permutant une caractéristique et en moyennant la différence dans *OUBE* avant et après la permutation sur tous les arbres. Soient un échantillon bootstrap $S_n^{\Theta_l}$ et son OBB_l associé, c'est-à-dire l'ensemble des observations non incluses dans $S_n^{\Theta_l}$. Pour un j fixe parmi les m caractéristiques, les valeurs de la $j^{\text{ème}}$ variable sont permutées aléatoirement sur OBB_l pour obtenir un échantillon perturbé appelé $\widetilde{OBB}_l$. La nouvelle $\widetilde{OUBE}_l$ de l'échantillon perturbé est alors calculée. Ces opérations se répètent pour tout échantillon bootstrap. L'importance de la $j^{\text{ème}}$ variable, notée $VI(X^j)$, est définie par la différence entre les erreurs moyennes des OBB_l original et perturbé [116].

$$VI(X^j) = \frac{1}{q} \sum_{l=1}^q (\widetilde{OUBE}_l - OBBE_l) \quad (1.42)$$

Si les permutations aléatoires de la $j^{\text{ème}}$ variable génèrent une augmentation de l'erreur, la variable

est importante. Plus le score $VI(X^j)$ est élevé, plus la caractéristique X^j est importante [115]. Pour des raisons de lisibilité, le nombre d'arbres q est renommé *ntree* dans le reste du mémoire.

1.6.6 Critères d'évaluation des méthodes

Les critères d'évaluation des méthodes sont des quantifications de l'erreur de prévision. Ils servent à estimer par différentes approches, la distance entre Y et $\hat{Y}$. On note $E = \hat{Y} - Y$ le signal erreur, $E(t)$ est alors l'erreur à l'heure t . On suppose avoir effectué β prévisions, et qu'à chaque prévision on obtient s erreurs (On rappelle que s est la longueur de l'horizon de prévision en heures). Le nombre total d'erreurs trouvées est alors $k = \beta.s$. On définit alors l'erreur absolue moyenne (MAE : Mean Absolute Error), la racine de l'erreur carrée moyenne (RMSE : Root Mean Squared Error), l'erreur absolue moyenne en pourcent (MAPE : Mean Absolute Percentage Error) et l'erreur absolue moyenne graduée (MASE : Mean Absolute Scaled Error) par :

$$MAE = \frac{1}{k} \sum_{t=1}^k |E(t)| \quad (1.43)$$

$$RMSE = \sqrt{\frac{1}{k} \sum_{t=1}^k E^2(t)} \quad (1.44)$$

$$MAPE = \frac{1}{k} \sum_{t=1}^k \frac{|E(t)|}{Y(t)} \times 100 \quad (1.45)$$

$$MASE = \frac{1}{k} \sum_{t=1}^k \left(\frac{|E(t)|}{\frac{1}{k-1} \sum_{i=2}^k |Y(i) - Y(i-1)|} \right) \quad (1.46)$$

Ces critères sont les plus utilisés dans la littérature [12]. L'erreur absolue moyenne est le critère le plus naturel, car c'est simplement la moyenne de toutes les erreurs trouvées, moyennant la valeur absolue pour éviter que les erreurs ne se compensent. La racine de l'erreur carrée moyenne fait la même chose, sauf qu'elle évite le signe de l'erreur non pas par la valeur absolue, mais par l'élévation au carré, la sommation et finalement la racine carrée, de manière similaire à la valeur efficace d'un courant électrique. L'erreur RMSE est généralement supérieure à la MAE. Ces deux critères permettent d'évaluer la performance de la prévision, mais ils donnent des quantifications brutes, c'est-à-dire non normalisées. Par exemple, si le signal à prédire est une puissance, ces critères seraient en watts. Et alors une erreur de 2 MW n'a pas la même signification pour un parc de 50 MW que pour un autre de 1000 MW. Une des solutions c'est d'utiliser le critère MAPE, qui permet de diviser l'erreur à chaque fois par la valeur réelle du signal avant de moyennner, afin d'avoir un pourcentage. Ce critère s'appelle aussi par abus de langage l'erreur relative. Néanmoins,

ce critère possède aussi ses inconvénients, surtout lorsque la valeur réelle tend vers zéro. On aura alors un critère qui tend vers l'infini, ce qui n'est pas réaliste, et qui est notamment le cas de la production intermittente (renouvelable). Pour les séries temporelles intermittentes en particulier, le critère MASE est proposé. Ce critère qui est sans unité, reflète une prédiction précise lorsqu'il est proche de zéro. Sa particularité réside dans le fait qu'il ne donne jamais des valeurs infinies sauf au cas où toutes les mesures $Y(i)$ sont identiques (ce qui n'a pas de sens). Par contre, il ne représente pas un pourcentage, contrairement à l'erreur MAPE.

Cette thèse utilise aussi d'autres critères : l'erreur absolue maximale (MXE : Maximum Absolute Error), l'erreur MAPE généralisée (GMAPE : Generalized Mean Absolute Percentage Error) et la pente de la ligne de régression (SRL : Slope of Regression Line) :

$$MXE = \max_{1 \leq t \leq k} |E(t)| \quad (1.47)$$

$$GMAPE = \frac{1}{k} \sum_{t=1}^k \frac{|E(t)|}{Y_{installée}} \times 100 \quad (1.48)$$

$$SRL = \frac{COV(Y, \hat{Y})}{VAR(Y)} \quad (1.49)$$

L'erreur MXE détermine l'erreur la plus élevée en valeur absolue sur toute la période à prédire. Elle est surtout utile pour détecter les erreurs les plus graves, et sert comme fonction objectif lorsqu'il s'agit de minimiser les pics par une méthode d'optimisation. Le critère GMAPE est une généralisation de la MAPE utilisée par les ingénieurs de la Société Tunisienne de l'Électricité et du Gaz (STEG) [117]. Elle consiste à diviser l'erreur par toute la puissance installée, au lieu de la diviser par la valeur réelle du signal. Elle n'est donc utilisable que lorsque le signal à prédire est une puissance. Elle évite par définition la division par des valeurs proches de zéro, et donc surmonte le problème majeur de l'erreur MAPE. Elle diffère du critère MASE (qui évite également la division par zéro) dans le sens où elle représente toujours un pourcentage. La pente de la ligne de régression SRL est le seul critère qui ne se débarrasse pas du signe de l'erreur, et qui permet donc de détecter le problème de sur-prédiction ou de sous-prédiction. Si ce coefficient est inférieur à 1, c'est que le prédicteur tend à sous-estimer le signal réel, et vice versa. Les notations MXE, SRL et GMAPE sont des notations propres à cette thèse, qu'on ne trouve pas dans la littérature. Le critère GMAPE est parfois appelé l'erreur absolue moyenne normalisée (NMAE : Normalized Mean Absolute Error).

En conclusion, si le signal réel est toujours très loin de zéro comme le cas de la prévision de charge, alors le critère MAPE est largement suffisant, car il donne une erreur en pourcentage bien significative. Et si le signal s'approche parfois de zéro comme le cas de la production éolienne, alors on recourt aux autres critères. Il est cependant suffisant d'utiliser le critère GMAPE lorsque le signal

à prédire est une puissance. Mais si le signal est une condition météo (vitesse du vent, irradiation solaire, température), ce critère n'est plus utilisable tel qu'il est défini. D'autres définitions peuvent être appliquées de manière similaire, qui consistent à diviser l'erreur $E(t)$ par la vitesse nominale du vent de 15 m/s dans le cas de la prévision éolienne par exemple. Néanmoins, dans cette thèse, le critère MASE est largement suffisant.

1.7 Conclusion

La prévision, ses types, ses horizons ainsi que les différentes classifications de ses méthodes ont été présentés au début de ce chapitre, tout en montrant les différences entre l'approche physique et l'approche statistique, et les principales évolutions à partir des méthodes statistiques conventionnelles vers les méthodes intelligentes et les métaheuristiques. La partie suivante est la synthèse bibliographique, qui a résumé les contributions scientifiques dans le domaine de la prévision, que ce soit de la production, de la demande ou du prix. Les principales démarches et comparaisons abordées par la littérature sont explicitement détaillées. Au prochain chapitre, on s'intéresse à la prédiction de la production éolienne et aux contributions de cette thèse à ce niveau. Il est intéressant de prédire la contribution de l'énergie éolienne dans la production avant de prédire la demande, afin de cerner la volatilité et l'incertitude de cette ressource intermittente ainsi que son degré de pénétration dans le réseau.

CHAPITRE 2 PRÉVISION DE LA PRODUCTION ÉOLIENNE

2.1 Introduction

Après avoir présenté en détails la prévision, ses différentes méthodes et utilisations en littérature, ce nouveau chapitre s'intéresse à la prévision de la production éolienne en utilisant les données historiques réelles d'un parc éolien en Tunisie, dans le but de trouver la méthode de prédiction optimale. On rappelle que cette prévision est très importante pour l'équilibre entre la production et la demande d'électricité en temps réel, et qu'elle est davantage plus intéressante lorsque la pénétration de l'éolien dans le parc de production est élevée. Et pour cette même raison la ressource photovoltaïque a été ignorée dans ce chapitre, vu qu'elle n'est pas intégrée au parc de production. Les autres ressources, le gaz naturel en particulier, ne sont pas intermittentes et ne nécessitent donc pas une prévision. La moindre amélioration dans la prévision éolienne intéresse énormément le gestionnaire du réseau surtout dans des cas de dispatching en temps réel.

D'une façon générale, il est important de faire face aux défis liés à l'intégration de l'énergie éolienne dans un système de production [118]. Le problème principal généralement rencontré est l'intermittence du vent, qui induit une instabilité dans le réseau et peut conduire soit à un manque, soit au gaspillage d'énergie. En général, une étude complète sur la variabilité du vent dans un site doit être faite avant d'y installer une ferme éolienne [119]. Mais ce n'est pas tout, car il n'est jamais possible de contrôler les rafales de vent. Cependant, il est possible, voire indispensable, de prédire leur comportement, et c'est ce qui donne à la prévision éolienne toute son utilité. Les chercheurs tunisiens ainsi que les gestionnaires du réseau utilisent en général les réseaux de neurones artificiels dans leurs modèles, comme indiqué au chapitre précédent. Mais ce n'est pas suffisant en général, car plusieurs facteurs entrent en jeu, tels que la disposition des éoliennes dans le parc, l'effet de lissage spatial, la direction du vent, etc. Ce chapitre essaye de prendre en compte toutes ces contraintes et de proposer la méthode qui semble être la plus adéquate avec ses optimisations potentielles.

Le chapitre vise à élaborer la prédiction d'énergie éolienne sur le court terme allant d'une à dix heures à l'avance en utilisant les séries temporelles. Les questions qui se posent à ce stade sont : quel effet la direction du vent a-t-elle sur la production éolienne, et quelle est sa relation avec la disposition des aérogénérateurs dans le parc ? Comment peut-on profiter du lissage spatial et éviter en même temps le sillage éolien ? La SVM est-elle adéquate pour la prédiction de la vitesse et de la direction du vent, et quelles sont ses optimisations ? Comment la qualité de la prévision est-elle affectée en fonction de la longueur de l'horizon, et selon quels critères d'évaluation ? Et finalement,

comment déduit-on l'énergie éolienne totale à partir des données de vent prédites ?

2.2 L'énergie éolienne

L'énergie éolienne a attiré l'attention depuis la fin du vingtième siècle, comme une ressource d'énergie renouvelable très évolutive et très prometteuse. Déjà installée et mise en production dans de nombreux pays du monde, elle intéresse surtout les états voulant réduire leurs dépendances des produits pétroliers. Pour les chercheurs, le défi est de tirer le plus de profit de cette ressource peu coûteuse sans générer des perturbations ou des risques qui pourraient affecter les installations. Selon le « Global Wind Report » [120], la capacité éolienne mondiale installée est de 318 105 MW à la fin de 2013, et elle est en expansion continue depuis plus de 20 ans. La Chine, le leader mondial, dispose de 91 412 MW de capacité installée, dont 16 088 MW installée uniquement en 2013. Cette énorme évolution reflète la tendance mondiale vers l'exploitation des ressources d'énergie non conventionnelles.

2.2.1 Progression de l'énergie éolienne en Tunisie (2007 – 2013)

La problématique énergétique en Tunisie est identique à celle du monde entier : comment augmenter la contribution de l'énergie renouvelable dans la production totale. Bien que ce soit un objectif déjà atteint dans quelques états, notamment les pays nordiques scandinaves et l'Islande (parfois la production conventionnelle est complètement abandonnée), ces cas sont encore des exceptions de pays de très faible population et donc très faible demande d'énergie.

En Tunisie, la part des énergies renouvelables dans la capacité de production d'électricité a augmenté de 2,4 % en 2007 à environ 3,6 % en 2009 (soit 116 MW par rapport à une puissance totale installée de 3312 MW). La consommation annuelle nationale d'électricité en 2007 a été de 11 249 GWh réalisant une progression de 3,5 % par rapport à l'année 2006. La pointe de consommation électrique a été de 2416 MW, valeur enregistrée en saison estivale. La consommation de combustible nécessaire à la production d'électricité a atteint 2541 ktep (kilotonne d'équivalent pétrole) en 2007 dont 2200 ktep de gaz naturel, principal combustible pour les centrales thermiques (turbines à vapeur et turbines à gaz), 339 ktep de fuel lourd utilisé comme combustible de secours des turbines à vapeur et 2 ktep de gasoil, utilisé comme combustible de secours pour les turbines à gaz. Cette lourde dépendance des produits pétroliers ainsi que l'augmentation annuelle de la demande ont poussé à élargir davantage la part des énergies renouvelables.

En 2013, sur 4425 MW de puissance installée, 306 MW (7 % de l'installation totale) sont assurés par des ressources renouvelables, ce qui constitue un pourcentage encore assez faible. Sur ces 306 MW d'énergie renouvelable, 62 MW proviennent de l'hydroélectricité et 244 MW de l'énergie

éolienne, dont 104 MW uniquement étaient en cours d'exploitation. L'énergie éolienne a fourni 357,8 GWh en 2013 ; soit 2 % de la production totale de 17 056 GWh. Cette contribution mineure est censée augmenter dans quelques années. Par contre, l'énergie solaire photovoltaïque n'est pas encore significativement intégrée au réseau tunisien. La figure 2.1 donne une idée sur l'évolution du parc de production national, ainsi que sur la production annuelle par type de générateur [121].

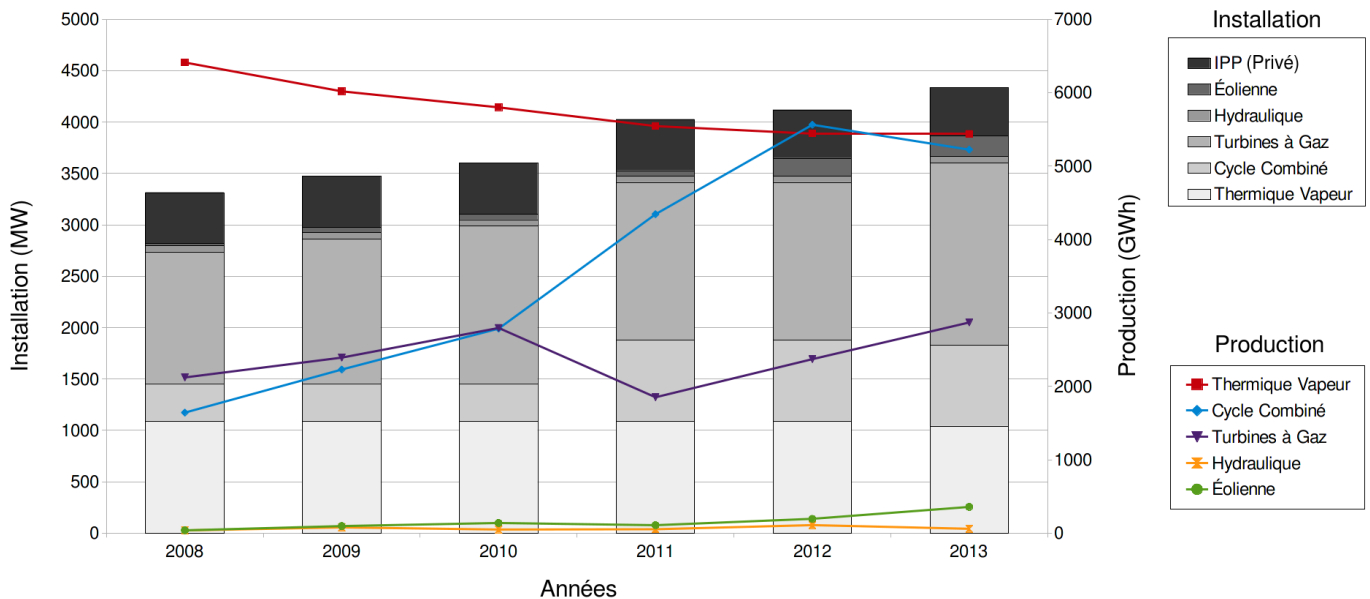


Figure 2.1 Évolution des installations et de la production annuelle par type de générateur

2.2.2 Présentation de la ferme éolienne Sidi Daoud

Le dernier projet d'extension de la centrale éolienne côtière de Sidi Daoud (figure 2.2), première centrale éolienne en Tunisie et sujet de cette étude, a permis d'augmenter la capacité de production d'électricité installée de ce champ éolien de 19,28 MW à 53,5 MW. Le tableau 2.1 et la figure 2.3 identifient les différentes tranches installées ainsi que leurs capacités respectives.



Figure 2.2 Parc éolien Sidi Daoud, visité le 13 novembre 2013

L'altitude des sommets des montagnes « Djebel El Hammam » et « Djebel Ghormane », sur lesquels sont implantés les aérogénérateurs (de la 1^{ère} et de la 2^e tranche), varie respectivement de 50 et de 100 m par rapport au niveau de la mer. Les aérogénérateurs de la 3^e tranche sont situés à quelques centaines de mètres des côtes marines. Les coordonnées géographiques du site sont 37°02' (latitude) et 10°56' (longitude). La centrale dans sa totalité avec ses trois tranches de machines s'étend sur superficie de 6 km de longueur et 5 km de largeur au nord-est du village de Sidi Daoud (environ 3500 hectares).

La décision du choix du site de Sidi Daoud pour l'implantation d'une centrale éolienne s'est appuyée sur une étude de faisabilité et de prospection des sites potentiels réalisée par « US Wind Power – 1991 ». Le potentiel éolien du parc en 2011 est représenté à la figure 2.4, sous forme de distribution des mesures de la vitesse du vent en 2011, échantillonnés à 10 minutes. L'allure de l'histogramme ressemble à une gaussienne centrée sur 5 m/s, qui est la vitesse dominante. La

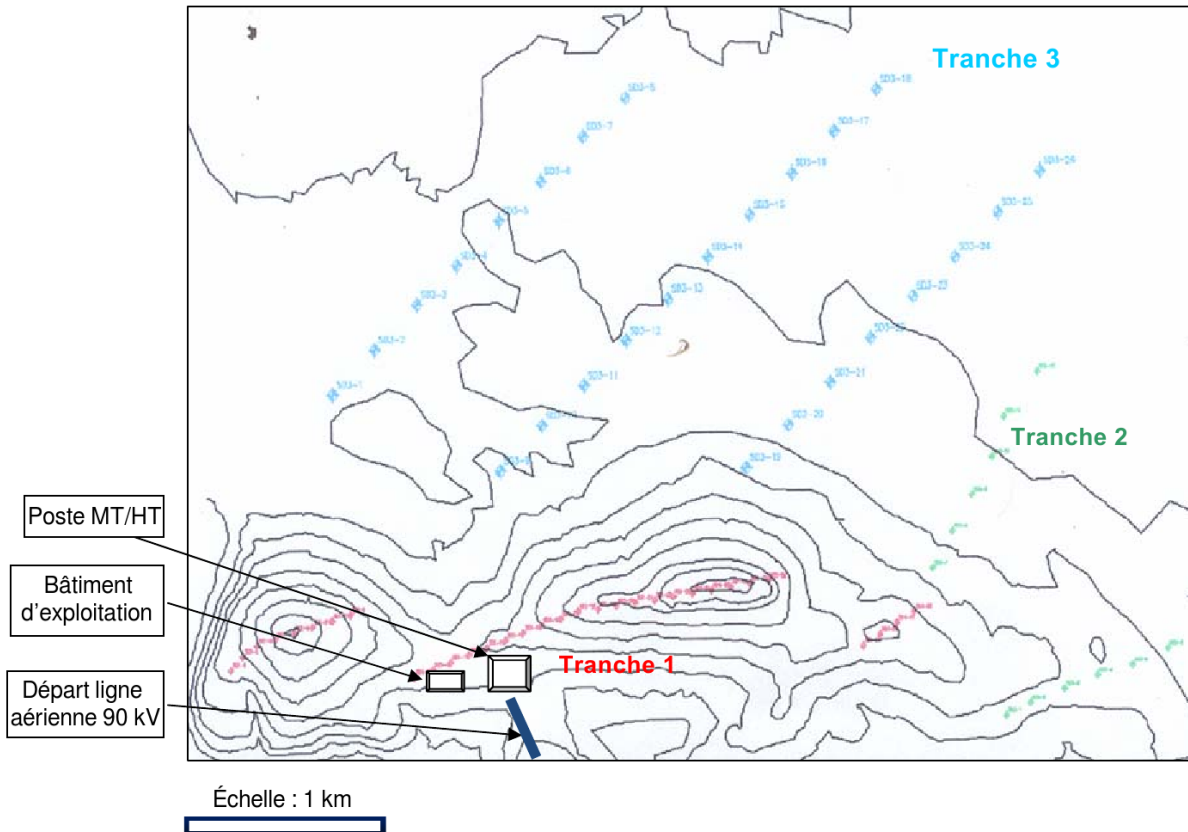


Figure 2.3 Disposition des éoliennes dans le parc, 1^{ère} tranche en rouge, 2^e tranche en vert et 3^e tranche en bleu

même distribution mais sous forme d'un nuage de points en coordonnées polaires est utilisée pour construire la figure 2.5, appelée rose des vents. Le module et l'argument de chaque point complexe sur cette figure représentent la vitesse et la direction du vent, respectivement. La rose des vents indique que la direction du vent dominante sur ce site varie de l'ouest à l'ouest nord-ouest. Le codage des directions du vent est circulaire formant 360° à partir du nord, tel que montré par la figure 2.6.

On remarque que les éoliennes des trois tranches sont disposées perpendiculairement à la direction du vent dominante (nord-ouest). Les espacements varient selon la taille des pales. Par exemple, les éoliennes de la tranche 3 (1320 kW) sont beaucoup plus espacées que celles de la tranche 1 (330 kW). Cette disposition qu'on trouve généralement dans les parcs éoliens sert à limiter l'effet de sillage, qui sera étudié en détails dans la section 2.3.2.

Le site comporte aussi un bâtiment administratif à deux étages qui abritent les bureaux du personnel et la salle de commande centralisée d'exploitation des 3 tranches. Un poste de transformation électrique 30/90 kV est situé à quelques mètres du bâtiment administratif, de superficie 500 m^2 en-

	1 ^{ère} tranche	2 ^e tranche	3 ^e tranche
Puissance électrique installée	10 560 kW	8720 kW	34 320 kW
Nombre d'aérogénérateurs	32	12	26
Puissance unitaire des machines	330 kW	10 aérogénérateurs de 660 kW 1 aérogénérateur de 800 kW 1 aérogénérateur de 1320 kW	1320 kW
Énergie produite estimée	> 25 GWh / an	> 20 GWh / an	> 100 GWh / an
Mise en service industriel	août 2000	septembre 2003	février 2009

Tableau 2.1 Les différentes tranches du parc éolien

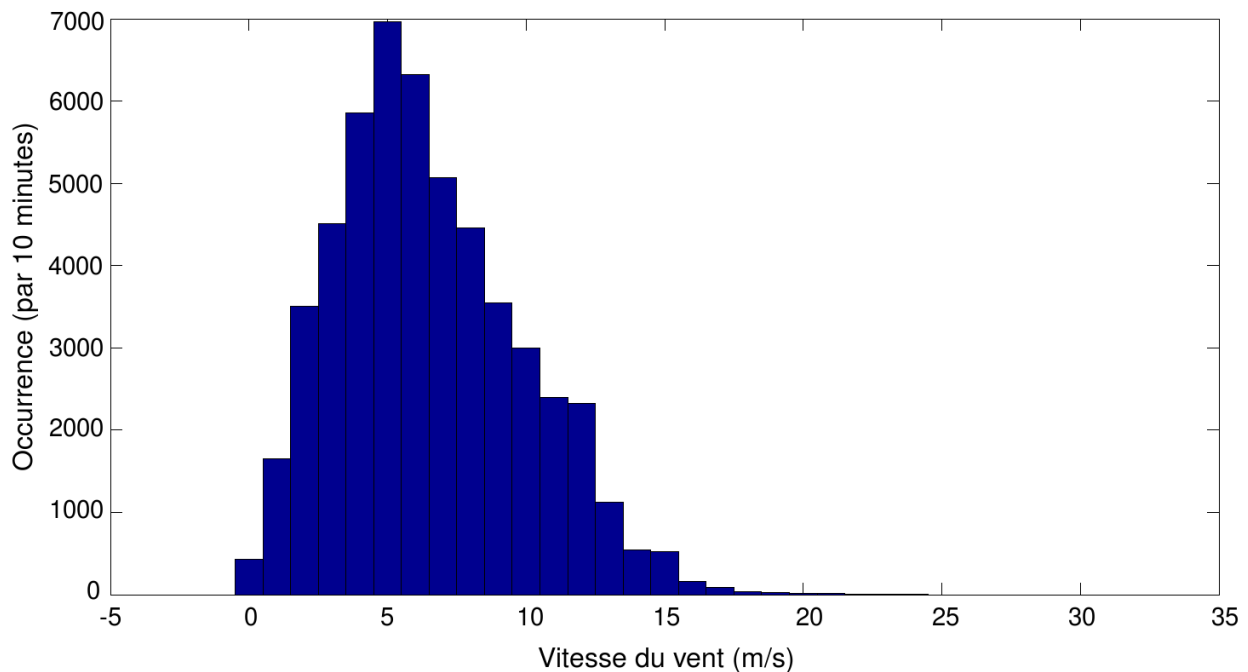


Figure 2.4 Potentiel éolien du parc en 2011

viron (constitué de 4 transformateurs de puissance avec leurs lignes et les jeux de barres) et clôturé avec un petit local de commande. Des câbles électriques souterrains MT (30 kV) relient les tours des aérogénérateurs au poste de transformation. Un mât météorologique fournit la vitesse du vent

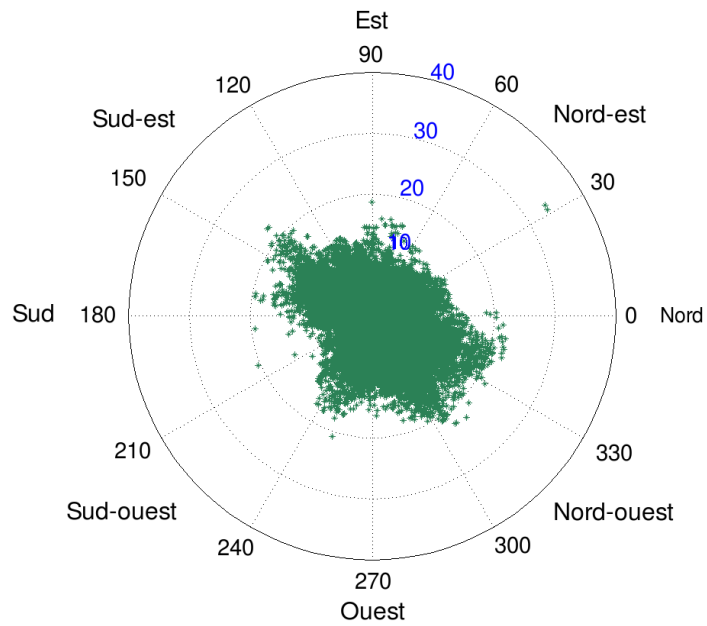


Figure 2.5 Rose des vents au parc en 2011

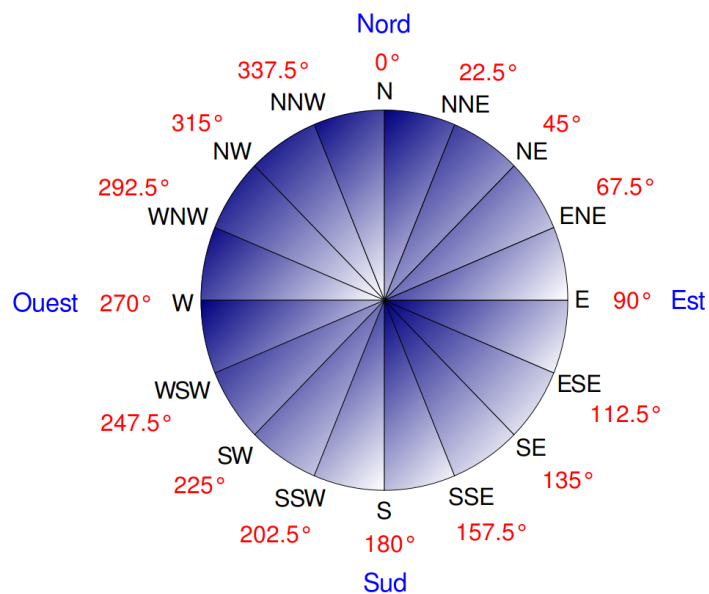


Figure 2.6 Codage des directions du vent

mesurée à deux altitudes différentes, la direction du vent, la température et l'humidité relative au sein du parc.

Un aérogénérateur est une tour tubulaire tronconique porteuse d'une nacelle contenant tous les équipements électromécaniques de production d'électricité (alternateur), de régulation de tension

et de fréquence, de contrôle et de sécurité. Chaque générateur gère ses propres protections et sa sécurité de façon autonome, de telle manière qu'en cas de foudre ou de vent extrême l'aérogénérateur se déconnecte automatiquement du réseau en se mettant en position de « stand by ». Il ne redémarrera que lorsque les données transmises à son ordinateur auront satisfait les conditions de sécurité de la machine.

La nacelle de l'aérogénérateur est équipée d'un anémomètre et d'une girouette permettant de déterminer à tout instant la vitesse et la direction du vent. Les aérogénérateurs sont productifs d'électricité à partir d'une vitesse de vent de 4 m/s. Dès que la vitesse du vent dépasse la limite prédéfinie de 25 m/s, le système de contrôle intégré stoppe automatiquement la rotation des pales pour respecter la limite de résistance mécanique des rotors.

L'énergie électrique générée sous basse tension (690 V alternative triphasée) est transmise par 3 câbles électriques au pied de la tour vers un transformateur élévateur de tension (690 V/30 kV) pour être évacuée par câbles souterrains vers le poste électrique MT/HT (30 kV/90 kV), où elle est injectée dans le réseau de transport aérien [122].

2.2.3 Intégration de l'énergie éolienne et effet des erreurs de prédiction

L'intégration de l'énergie éolienne en Tunisie a ses effets sur le réseau de distribution, puisqu'elle est vue comme une ressource distribuée (faibles puissances divisées sur plusieurs parcs éoliens). La production distribuée (décentralisée) a ses avantages et ses inconvénients. Parmi les avantages, on compte la diminution de la charge alimentée par les ressources centralisées, la diminution des pertes joule (diminution des distances) et l'amélioration du profil de tension. Les effets nuisibles peuvent se manifester sous la forme de fluctuations de tension, distorsions harmoniques, surintensités, dépassement des limites de capacité du réseau et diminution de la fiabilité. Pour profiter des avantages et éviter les inconvénients, la ressource distribuée doit être fiable, répartie, de taille adaptée et localisée aux endroits appropriés. Mais si la source est intermittente comme le cas de l'éolien, ces conditions sont difficiles à respecter. Le respect des contraintes d'intégration est d'autant plus critique que la puissance de la ressource est plus importante [123].

Concernant l'effet des erreurs de prédictions, on parlera de l'impact sur les coûts de régulation. On distingue alors trois cas. Au cas où la production correspond parfaitement à la consommation, aucune régulation n'est nécessaire ; si la production n'est pas suffisante pour répondre à la charge, on a une situation de régulation à la hausse, pour laquelle il est nécessaire d'augmenter la production (ou diminuer la consommation) ; et si la consommation dépasse la production, on fait face à une situation de régulation à la baisse, dans laquelle la production doit être réduite (ou la consommation doit augmenter). Dans des cas pareils, l'exactitude de la prévision est d'une importance capitale. On a démontré que l'utilisation d'une méthode d'intelligence artificielle, à savoir les réseaux de

neurones flous, peut diminuer les coûts de régulation à la hausse de 38 % et les coûts de régulation à la baisse de 34 % par rapport au modèle de persistance [124]. En Tunisie cependant, la capacité éolienne limitée à une centaine de MW fait que la performance de la prévision éolienne ne soit pas aussi importante que celle de la demande, qui dépasse parfois les 3500 MW (étudiée au chapitre 3). Par conséquent, les erreurs de prévision de la demande seront beaucoup plus graves que celles de l'éolien, et induiront des coûts plus élevés et des problèmes de stabilité du réseau plus importants.

2.3 Analyse de la production du parc éolien

Cette section sert à déterminer les facteurs influant sur la production du parc éolien, en examinant le comportement d'une seule éolienne, puis d'un ensemble d'éoliennes. Ces facteurs permettent de bien choisir les entrées du prédicteur et la stratégie de prédiction.

2.3.1 Étude d'une seule éolienne

La courbe de puissance est une caractéristique intrinsèque de l'éolienne donnée par son constructeur. Théoriquement, elle est définie comme une fonction de transfert non linéaire reliant la vitesse du vent à la puissance produite, telle que montrée par la figure 2.7. Selon la fonction de transfert du générateur disponible, la conversion de l'énergie éolienne cinétique (qui est proportionnelle au cube de la vitesse du vent) en puissance récupérable varie non linéairement, avec une production nulle en dessous d'un seuil de vitesse minimum (de l'ordre de 3 m/s), une croissance rapide de la puissance jusqu'à ce que la machine atteigne sa puissance nominale (environ 15 m/s), et une sortie constante au-dessus de ce niveau jusqu'à ce que la vitesse de coupure soit atteinte (environ 25 m/s).

La puissance théorique capturée par le rotor d'une éolienne est calculée comme suit :

$$P_r = \frac{1}{2} \rho \pi R^2 C_p(\lambda, \eta) v^3 \quad (2.1)$$

Où ρ est la densité de l'air, R est le rayon du rotor et v la vitesse du vent, en supposant qu'elle est uniforme sur toute la surface balayée. Le paramètre C_p est appelé coefficient de puissance du rotor et il dépend de l'angle de tangage des pales η (pitch angle) et du rapport pointe de vitesse λ (tip-speed ratio) [125].

Cependant, la courbe réelle peut différer, car l'éolienne fait dévier le vent, et par conséquent la vitesse en amont et en aval du rotor n'est pas la même. De plus, la densité de l'air ρ varie avec l'altitude, la température et l'humidité relative. La figure 2.8(a) montre le nuage de points de fonctionnement d'une éolienne 330 kW du parc Sidi Daoud, où la puissance fournie est en fonction de la vitesse du vent mesurée localement (anémomètre au-dessus de la nacelle, à une altitude de 30

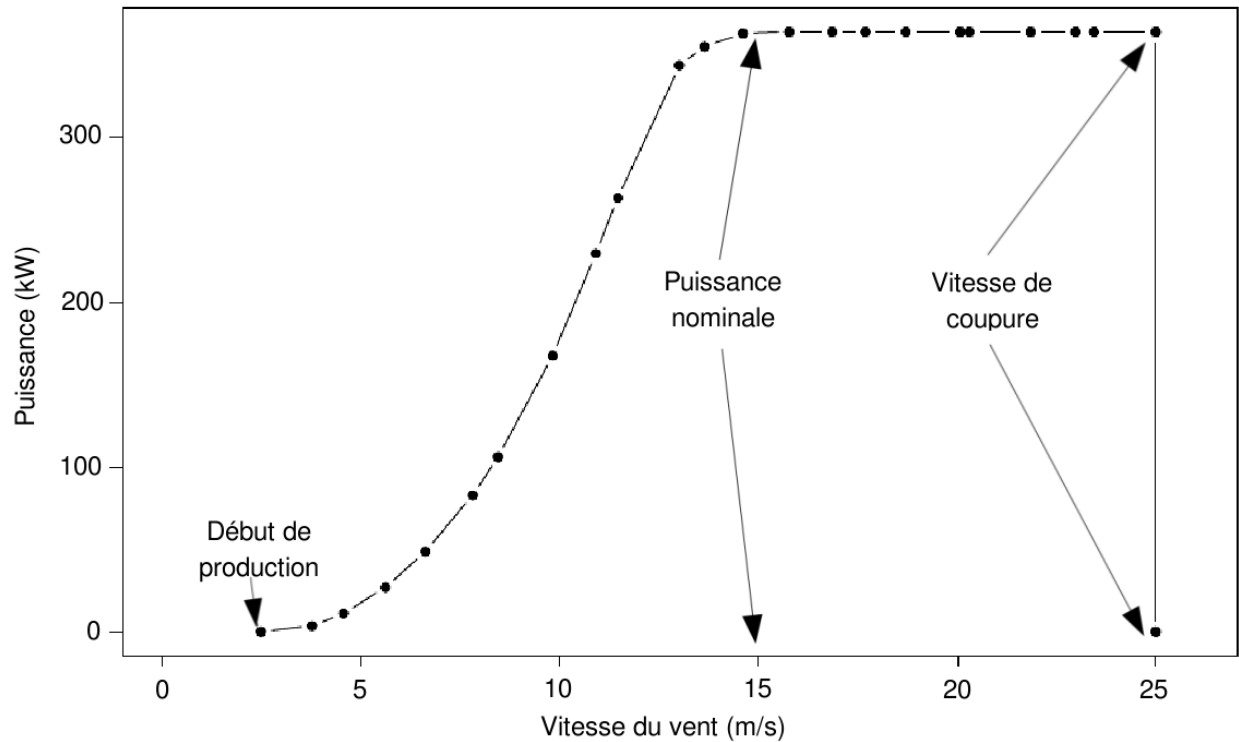


Figure 2.7 Courbe de puissance théorique

m). Une plage de valeurs de 270 à 380 kW correspond à la même vitesse de 15 m/s. Ce qui prouve expérimentalement que la courbe de puissance théorique ne suffit pas pour déduire la puissance réellement produite.

La direction du vent a aussi son effet sur la courbe de puissance réelle, même si les nacelles sont auto-orientables, c'est-à-dire capables de suivre la direction du vent. Cet effet est illustré par la figure 2.8(b), où la puissance est en fonction de la direction mesurée par le mât. La production est maximale dans la direction NW (direction dominante) ainsi que pour la direction opposée SE, alors qu'elle est minimale pour les directions perpendiculaires SW et NE. La figure 2.8(c) avec l'extrapolation d'une surface 3d illustre encore mieux les crêtes et les creux de puissance en fonction de la direction.

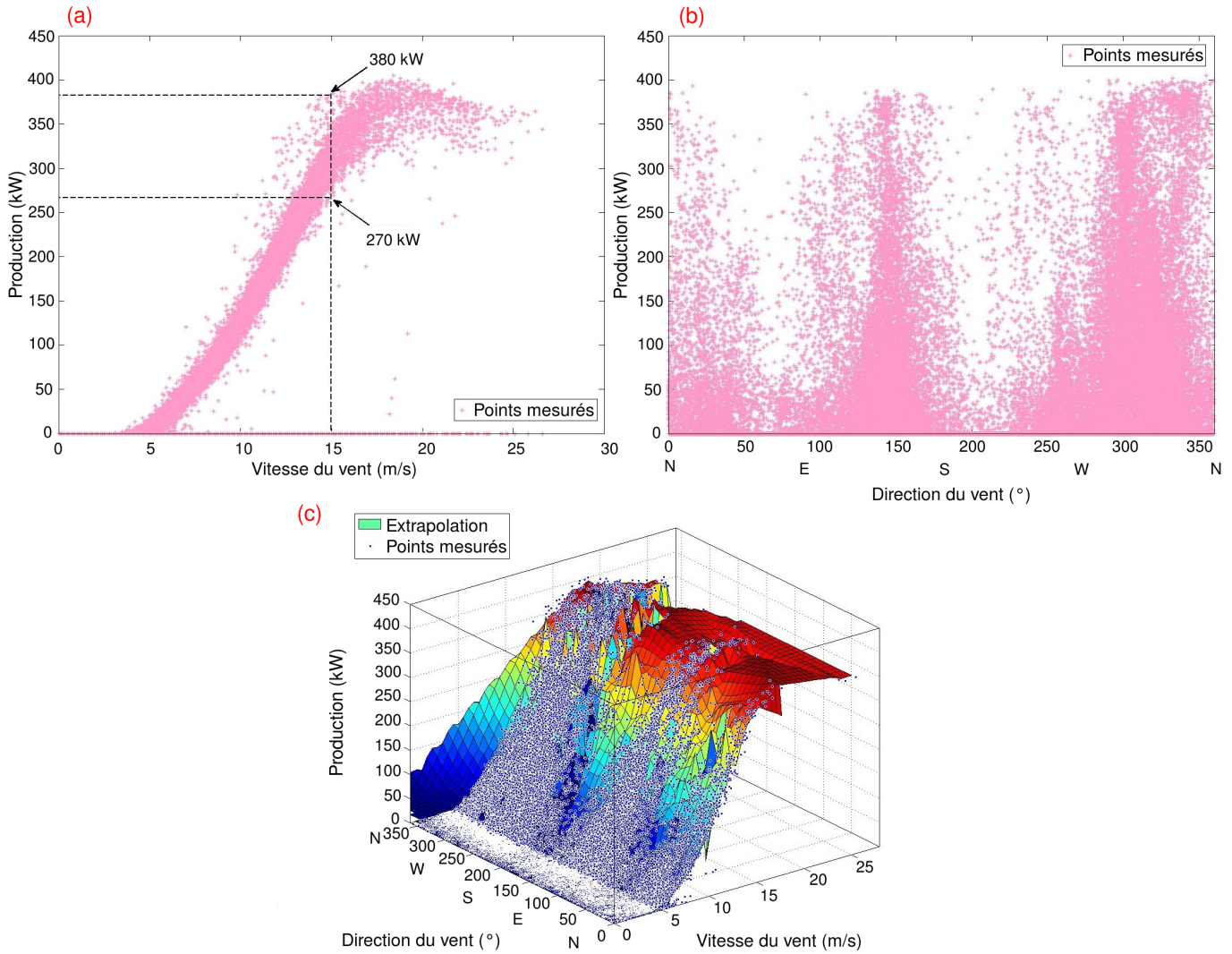


Figure 2.8 Points de production d'une éolienne 330 kW en 2011, en fonction de la vitesse du vent (a), la direction du vent (b), et en fonction des deux (c)

2.3.2 Étude d'un ensemble d'éoliennes

La vitesse du vent n'est pas identique en tout point du parc éolien malgré sa petite surface de 6×5 km². La figure 2.9(a) illustre cette différence sur les signaux des anémomètres de 5 éoliennes déjà très proches, la direction étant mesurée par le mât. Vu cette différence, bien visible pour certaines directions (notamment SSE), la question qui se pose est : quelle vitesse doit-on prédire ? Afin de répondre à cette question et afin d'éviter la multi-prévision, une solution simple consiste à prédire la moyenne spatiale des vitesses, définie par la moyenne instantanée des mesures de tous les anémomètres. Ce choix s'inspire de l'effet de lissage spatial, connu pour réduire les erreurs de pré-

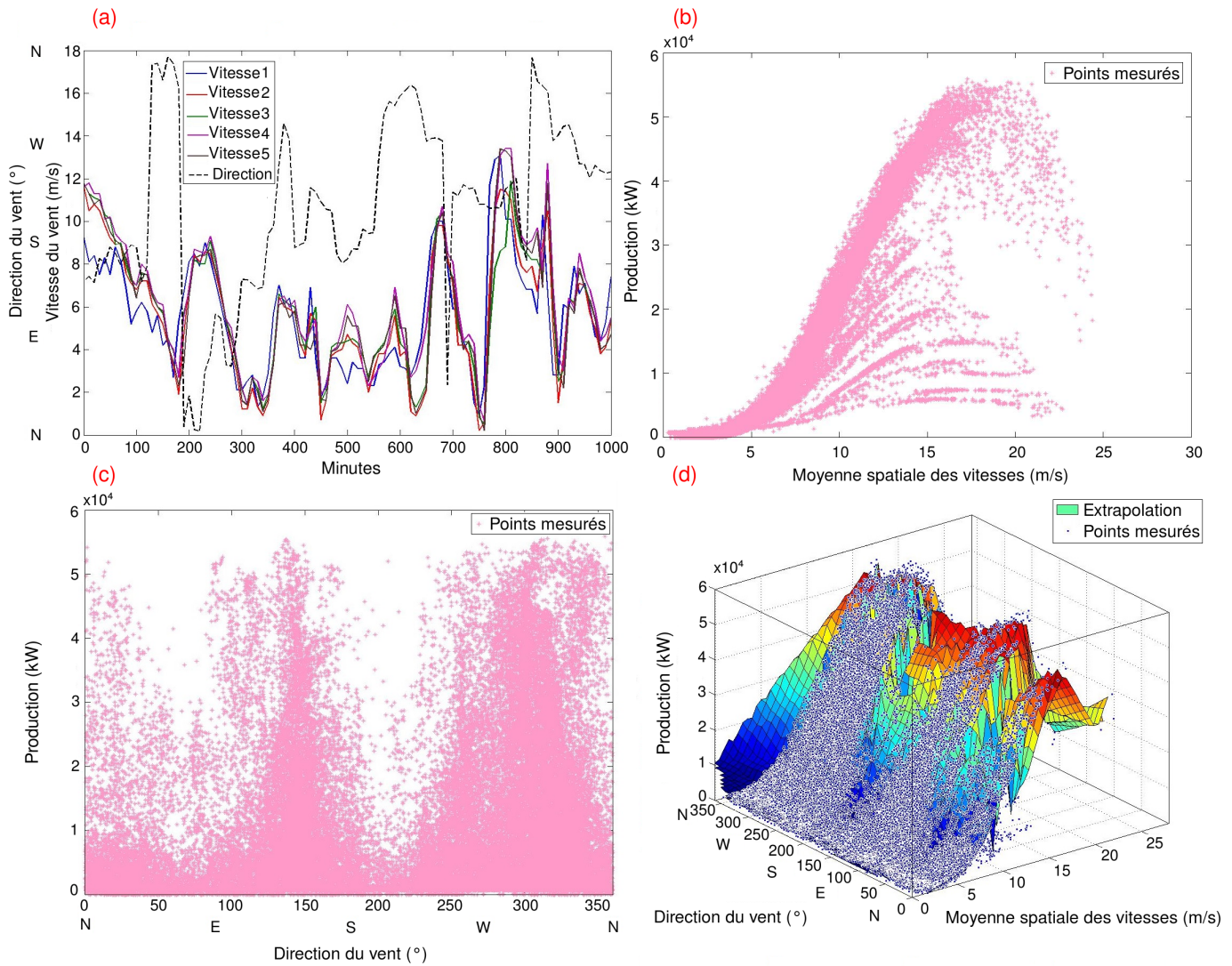


Figure 2.9 Différentes mesures de la vitesse du vent (a), points de production du parc éolien en 2011, en fonction de la vitesse du vent (b), la direction du vent (c), et en fonction des deux (d)

vision [52, 126]. Cet effet se révèle par une courbe de puissance globale plus lisse et tendant vers un profil bien régulier, d'autant plus que le nombre d'éoliennes est grand et couvre une surface plus large. La somme des puissances produites par les 70 aérogénérateurs du parc en fonction de la moyenne spatiale des vitesses donne le nuage de points de la figure 2.9(b). Le profil est bel et bien similaire à celui d'une seule éolienne, mais avec une puissance nominale de 53,4 MW (puissance totale du parc), avec d'autres profils similaires plus petits et moins épais. Ce profil bien lisse et régulier n'apparaît qu'en fonction de la moyenne spatiale des vitesses. Les petits profils apparaissent bien évidemment lorsque le parc éolien ne fonctionne pas à pleine capacité, c'est-à-dire lorsque

certaines éoliennes sont à l'arrêt. Si la vitesse du vent sur la figure 2.9(b) était mesurée en un seul point du parc, quel que soit ce point, le nuage de points aurait été beaucoup plus épais, et il serait impossible d'établir une relation vitesse-puissance.

De même, la direction a son effet sur la puissance produite. Sur la figure 2.9(c), la puissance totale produite par le parc éolien est représentée en fonction de la direction du vent mesurée par le mât, sous forme d'un nuage de points. Encore, la production est maximale uniquement pour certaines directions, et minimale pour d'autres. La figure 2.9(d) avec la surface d'extrapolation montre davantage l'effet de la direction, où les creux et les crêtes de puissance sont plus visibles que ceux de la figure 2.8(c). Il n'est pas utile de faire la moyenne spatiale des directions, à cause des erreurs liées au codage circulaire (voir la section 2.4.2).

La disposition des aérogénérateurs dans le parc est d'une importance capitale. Les éoliennes sont largement espacées dans la direction du vent dominante (3 à 9 fois le diamètre du rotor), alors que les espaces sont moins importants dans la direction perpendiculaire (3 à 5 fois le diamètre du rotor). Cet espacement sert à diminuer l'effet de la turbulence qui naît derrière les pales, connue sous le nom de sillage éolien. Mais en contrepartie, il provoque la diminution de la production selon les directions peu espacées. Évidemment, ce problème se pose même si les nacelles s'orientent selon la direction du vent, qui est le cas de la ferme éolienne Sidi Daoud. Cependant, il est à noter que la direction du vent n'a cet effet que pour les éoliennes à axe horizontale.

2.4 Conception du modèle de prévision éolienne

2.4.1 Choix de la méthode et des entrées

La méthode choisie pour la prévision éolienne est une méthode statistique indirecte à court terme, qui prédit la production éolienne sur un horizon d'une heure à l'avance ($s = 1$) par le modèle SVR. Les entrées choisies sont la vitesse et la direction du vent. On procède à la conception de deux modèles, un pour la prédiction de la vitesse, et l'autre pour la prédiction de la direction. Ces deux modèles permettent de prédire le signal de vent de l'heure qui suit, en fonction des mesures de vent de l'heure actuelle. Ensuite, ces prédictions sont converties en puissance éolienne moyennant un réseau de neurones. Les justificatifs de tous ces choix sont élaborés par la suite, un par un (et pour les comprendre, il faut revenir à la section 1.4 du chapitre 1).

L'approche utilisée est statistique car l'approche physique n'est pas accessible (nécessité d'un modèle de l'atmosphère et d'un supercalculateur). Elle est indirecte pour profiter de l'information de la direction du vent et mettre en évidence son effet sur la production, mais aussi pour ne pas perdre la interprétabilité de la prévision (La section 2.3 a bel et bien montré l'effet de la direction du vent sur la production éolienne, et sa relation avec la disposition des aérogénérateurs, même

lorsque ces derniers sont capables de s'auto-orienter). La méthode indirecte permettra aussi par la suite d'établir deux processus d'optimisation (un pour la vitesse et l'autre pour la direction), tandis que l'approche directe ne supporte qu'un seul processus (optimisation de la production tout court). L'horizon est une heure à l'avance puisqu'on ne peut aller trop à l'avant moyennant une méthode statistique pour un signal aussi chaotique comme le vent. Et le plus important, c'est le choix de la méthode SVR. La SVR est choisie car c'est une méthode intelligente capable de capturer la non-linéarité du signal de vent, contrairement aux méthodes conventionnelles. Et ce qui la distingue des autres méthodes intelligentes, c'est qu'elle peut être optimisée facilement. L'ANN par exemple est difficile à optimiser vu le nombre important de paramètres à régler (nombre de couches, nombre de neurones, fonctions d'activations, algorithme d'apprentissage). La CART et la RF ne peuvent pas être optimisées, puisqu'elles sont peu sensibles à leurs paramètres.

La principale limitation de la méthodologie proposée est qu'elle ne peut pas prédire les rafales de vent, puisque la vitesse du vent est considérée constante et égale à sa valeur moyenne durant le pas temporel futur. Les rafales de vent sont parfois dangereuses et peuvent obliger les gestionnaires du parc éolien à freiner mécaniquement les pales. Mais il n'est possible de prédire ces événements qu'avec des modèles physiques, les méthodes statistiques en étant incapables [39]. De plus, les méthodes statistiques sont plus orientées vers la prévision de la production, plutôt que la prédiction des dangers potentiels. Une autre limitation de la méthodologie est que sa performance se dégrade avec l'élargissement du pas temporel futur, ce qui la restreint à un horizon très court.

2.4.2 Conception du prédicteur

En général, prédire l'énergie éolienne revient à prédire la puissance que le parc éolien est censé produire en fonctionnant sous un vent constant pendant une période donnée [125]. Donc, afin de prédire la production d'une heure à l'avance, le vent moyen d'une heure à l'avance doit être prédit et ensuite convertit en puissance. La méthode consiste à exposer un nombre m de pas temporels passés (mesures passées du vent) à un prédicteur SVR, afin d'estimer le vent moyen d'un pas temporel futur. Le vent moyen dans ce cas est moyenné temporellement, et non pas spatialement. Pour faire la différence, on parlera de vitesse moyenne pour désigner la moyenne temporelle, et de moyenne spatiale des vitesses pour désigner la moyenne spatiale abordée dans la section 2.3.2.

La figure 2.10 montre le modèle construit dans le but de prédire la vitesse et la direction du vent, où le nombre m d'échantillons passés est égal à 6. Ces échantillons sont espacés de 10 minutes, et appartiennent tous à l'heure actuelle (qui précède l'heure à prédire). La sortie est la moyenne estimée de six valeurs également espacées de 10 minutes. Il est important de mentionner que, durant la phase d'apprentissage, uniquement la vitesse du vent est moyennée temporellement à la sortie. Quant à la direction, une valeur au milieu du pas temporel futur est choisie et considérée

comme direction dominante durant cet intervalle de temps. Le problème avec le moyennage de la direction est qu'il n'a pas de sens. Si par exemple le vent soufflait du nord (0°) après avoir soufflé du sud (180°) pendant quelque temps, la moyenne temporelle serait l'est (90°) puisque le codage des directions est circulaire (figure 2.6), et c'est complètement faux comme résultat. La même chose s'applique au moyennage spatial.

En utilisant la notation mathématique du chapitre 1, on aura donc : $X(t+1) = (X^1(t+1), \dots, X^6(t+1))^T$ et $\hat{Y}(t+1) = F(X(t+1))$, où F est la transmittance du prédicteur SVR, t est l'heure actuelle et $t+1$ est l'heure à prédire. Notons v la vitesse et d la direction du vent, $v(t)$ et $d(t)$ sont deux vecteurs contenant chacun 6 valeurs consécutives prises à l'heure t et espacées de 10 minutes. Dans le cas de la vitesse, on dispose de $X(t+1) = v(t)$ et $Y(t) = \bar{v}(t)$, et le prédicteur SVR doit calculer $\hat{Y}(t+1) = F(X(t+1)) = \hat{\bar{v}}(t+1)$, où $\bar{v}$ est la moyenne du vecteur v et $\hat{\bar{v}}$ est l'estimation de la moyenne de v . Et dans le cas de la direction, on dispose de $X(t+1) = d(t)$ et $Y(t) = \tilde{d}(t)$, et le prédicteur SVR doit calculer $\hat{Y}(t+1) = F(X(t+1)) = \hat{\tilde{d}}(t+1)$, où $\tilde{d}$ est la composante dominante du vecteur d et $\hat{\tilde{d}}$ est l'estimation de la composante dominante de d .

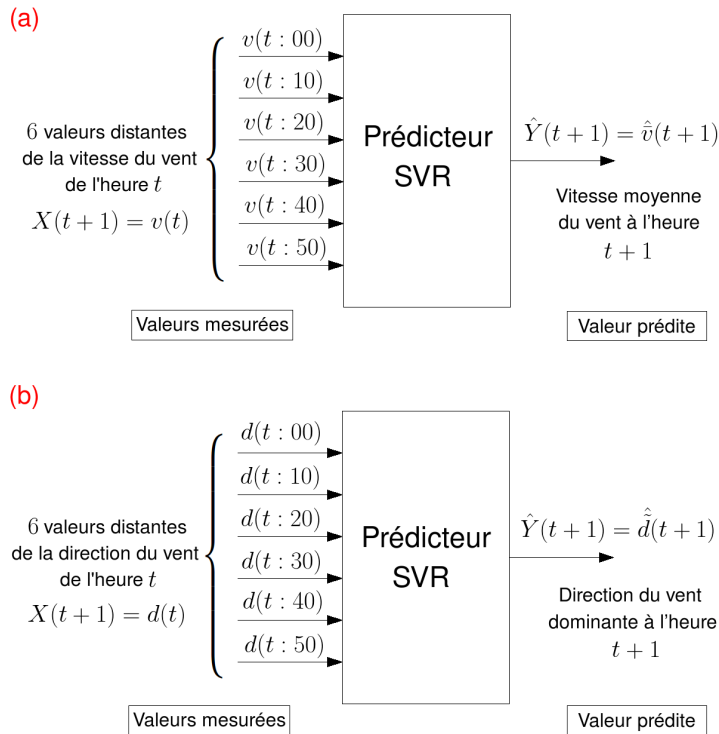


Figure 2.10 Modèles de prévision de la vitesse (a) et de la direction (b)

2.4.3 Données disponibles et prétraitement

Les données disponibles dans cette étude sont fournies par la Société Tunisienne de l'Électricité et du Gaz (STEG) par l'intermédiaire du Centre National des Mouvements d'Énergie à Radès [117] pour des buts de recherche scientifique exclusivement. Les données sont celles du parc éolien Sidi Daoud. Elles sont brutes et sans aucun traitement postérieur, telles qu'elles sont enregistrées par les capteurs. Pour les années 2010 et 2011, on dispose pour chacun des aérogénérateurs de la vitesse du vent mesurée localement en mètres par seconde (par l'anémomètre au-dessus de la nacelle) et de la puissance produite en kW, les deux par une précision d'un chiffre après la virgule, et par échantillonnage de 10 minutes. On dispose également des enregistrements du mât météorologique, qui fournit de sa part pour les deux années 2010 et 2011 et par pas de 10 minutes, la vitesse du vent mesurée à deux altitudes différentes en m/s, la direction du vent en degrés, la température de l'air en degré Celsius et la pression atmosphérique en hPa.

Date et heure	Vitesse du vent (m/s)	Direction du vent (°)	Température (° C)
01/01/2010 00:00:00	6,8	113	15,8
01/01/2010 00:10:00	6,8	113	15,3
01/01/2010 00:20:00	6,5	113	14,6
01/01/2010 00:30:00	6,4	113	14,6
01/01/2010 00:40:00	6,1	113	14
01/01/2010 00:50:00	6,1	113	13
01/01/2010 00:50:00	6,1	113	13
01/01/2010 00:50:00	6,1	113	13
01/01/2010 01:00:00	5,9	50,9	12,2
01/01/2010 01:10:00	5,9	23,4	12,2
01/01/2010 01:20:00	5,8	44	12
01/01/2010 01:30:00	4,8	22,1	
01/01/2010 01:40:00	4,8	5,1	12,3
01/01/2010 01:50:00	5,1	48,6	12,6
01/01/2010 02:00:00	4,7	18	13

Redondances : valeurs en excès supprimées

Moyennage : remplacer 6 valeurs par leur moyenne (si applicable)

Défaillance capteur : période évitée

Valeur manquante : remplacée par la précédente (12)

Figure 2.11 Détection des fautes et prétraitements possibles

L'énorme quantité de données contient nécessairement des fautes dues aux défaillances des capteurs, ce qui nécessite un prétraitement. Les données sont traitées par des logiciels de calcul numérique, tels que Matlab. Lorsqu'une donnée manquante est détectée, elle est signalée par le logiciel. Dans ce cas, elle est remplacée par la valeur précédente. De même, lorsqu'il y a des données redondantes, elles sont supprimées. Les redondances sont détectées par le nombre de lignes du tableau

contenant les données, qui doit contenir nécessairement $365 \times 24 \times 6 = 52\,560$ lignes pour chaque année. S'il y a des lignes de plus ou de moins, c'est qu'il y a des données en excès ou manquantes. On note cependant que ces données sont très rares et n'ont aucune influence sur la qualité de la prévision. Le traitement ne s'arrête pas là, car parfois les capteurs se bloquent sur une valeur pendant une très longue durée. Ces durées de blocages sont simplement évitées. Des informations supplémentaires peuvent être calculées, telles que la puissance totale produite par le parc (somme de la production de tous les aérogénérateurs), la moyenne spatiale des vitesses (la moyenne des enregistrements des 70 anémomètres), et la moyenne temporelle des vitesses (qui consiste à remplacer chaque 6 valeurs consécutives par leur moyenne). La figure 2.11 récapitule les prétraitements les plus fréquents.

2.5 Résultats de la prévision du vent

2.5.1 Résultats non-optimisés

La configuration de la machine SVR utilisée pour la prévision est donnée dans le tableau 2.2, et elle est commune pour la vitesse et la direction (revenir aux équations (1.19) à (1.25) pour comprendre l'utilité du noyau et la signification des paramètres C et γ). Les paramètres de configuration sont soit empiriques (choisis après plusieurs tests de simulation), soit recommandés par le guide SVM [127].

L'ensemble d'apprentissage contient 4 mois considérés comme suffisants pour contenir toutes les valeurs possibles. Cette période contient 3672 heures, qui servent à construire 3671 observations ($n = 3671$), puisque la première heure est non prédictible (les entrées pour la prédire sont en dehors de l'ensemble d'apprentissage). Lorsque le modèle est établi, le test est effectué sur les données de l'ensemble de test. Cet ensemble contient 744 heures, soit 743 prédictions possibles ($\beta = 743$).

Paramètre	Valeur
Ensemble d'apprentissage	Août à décembre 2010
Ensemble de test	Janvier 2011
Type de régression	Epsilon ε
Noyau	Fonction radiale de base RBF
Paramètre de régularisation C	100
Coefficient du noyau γ	0,0004

Tableau 2.2 Configuration de la machine SVR

La moyenne spatiale des vitesses est prédite pour une heure à l'avance, par la machine SVR proposée. Les résultats apparaissent sur la figure 2.12(a), qui montre les 100 premières heures de janvier

2011. La nature chaotique du vent est évidente à travers la courbe, ce qui pose des difficultés pour le prédicteur, surtout lorsqu'il s'agit de variations brusques et aiguës. La figure 2.12(b) montre le nuage de points de l'ensemble de test en totalité, soulignant la corrélation entre les valeurs prédites et mesurées. La droite de régression est calculée par la méthode des moindres carrés. Bien évidemment, la prévision est de plus en plus précise au fur et à mesure que la droite de régression ainsi que les points soient plus proches de la première bissectrice, représentée par une ligne discontinue.

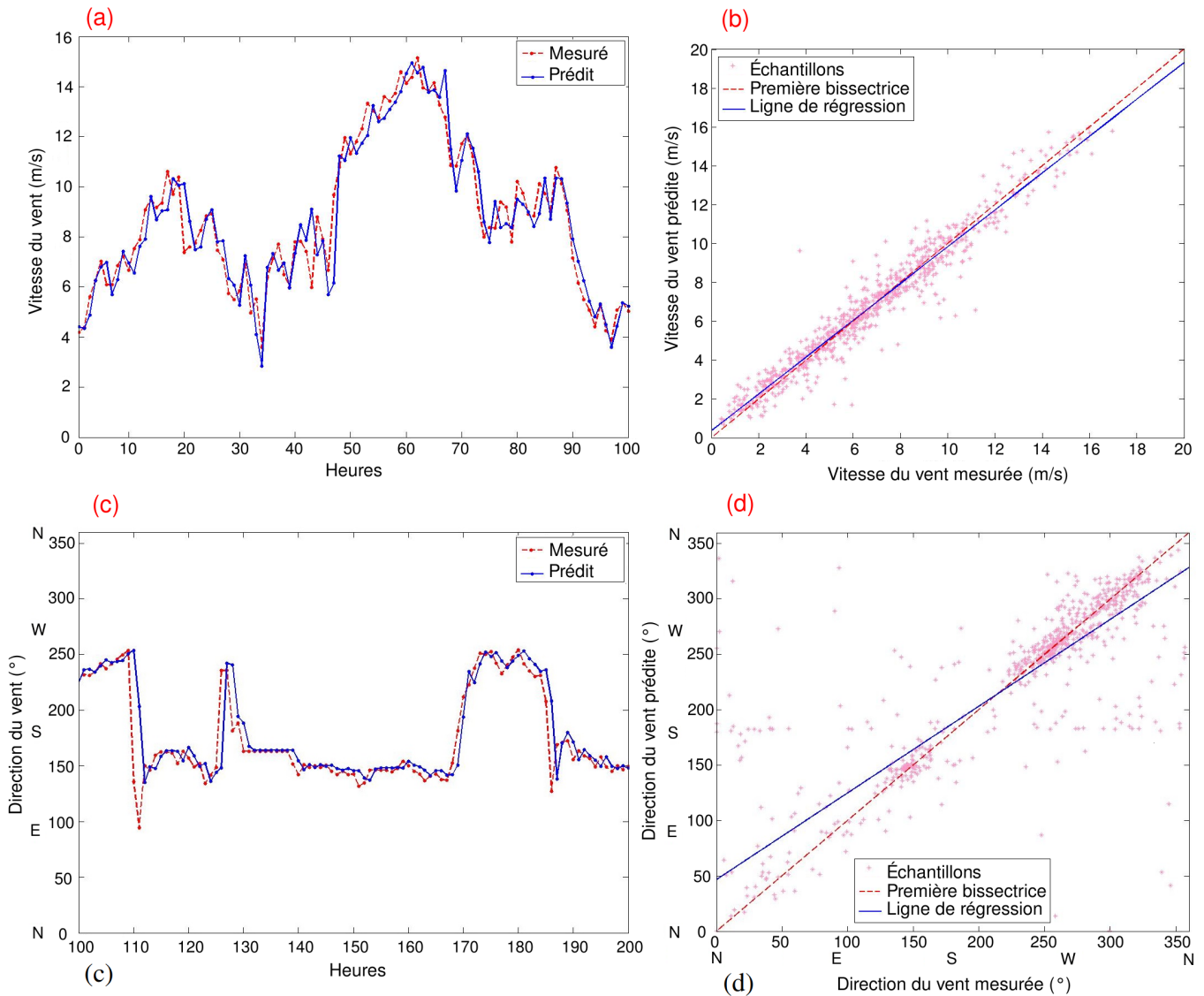


Figure 2.12 Résultats de la prévision de la vitesse (a,b) et de la direction (c,d) d'une heure à l'avance par la machine SVR, en janvier 2011

Un autre test de prévision est effectué sur la vitesse du vent mesurée par le mât météorologique, afin de la comparer à la moyenne spatiale des vitesses. Les résultats de comparaison donnés par

le tableau 2.3, contiennent l'erreur absolue moyenne (MAE), la racine de l'erreur carrée moyenne (RMSE), l'erreur absolue moyenne en pourcent (MAPE), l'erreur absolue moyenne graduée (MASE), l'erreur absolue maximale (MXE), et la pente de la ligne de régression SRL (voir les équations (1.43) à (1.49)). Le nombre total d'erreurs est $k = \beta.s = 743 \times 1 = 743$.

À titre de rappel, les critères MAE et RMSE sont deux quantifications différentes de l'erreur moyenne utilisant l'unité originale de la vitesse, qui est le mètre par seconde. Le critère MAPE quantifie aussi l'erreur moyenne, mais sous la forme d'un pourcentage qu'on appelle parfois l'erreur relative. Le critère MASE fait la même chose, sauf qu'il est plus adapté aux séries intermittentes (et il n'est pas exprimé en %). Le critère MXE donne l'erreur maximale en mètre par seconde, alors que la pente SRL indique, sans unité, une performance accrue tant qu'elle se rapproche de 1. Tous ces critères (excepté SRL) quantifient l'erreur de prévision d'une manière à négliger le signe et à se concentrer uniquement sur la quantité ; donc rien n'indique si l'erreur est une surestimation ou une sous-estimation.

Critère	Moyenne spatiale des vitesses		Vitesse mesurée par le mât	
	PER	SVR	PER	SVR
MAE (m/s)	0,77	0,55	0,73	0,62
RMSE (m/s)	1,11	0,81	1,22	1,04
MAPE (%)	16,54	11,80	-	-
MASE	1,00	0,72	1,00	0,85
MXE (m/s)	7,42	5,88	11,83	11,36
SRL	0,94	0,94	0,89	0,90

Tableau 2.3 Critères d'évaluation de la prévision de la vitesse par la machine SVR

Le tableau 2.3 figure la diminution de tous les types d'erreur jusqu'à 28 % en utilisant le modèle SVR par rapport au modèle de persistance (on rappelle que la persistance est la référence). De plus, la prédiction de la moyenne spatiale des vitesses possède des avantages par rapport à la vitesse mesurée à un seul endroit, notamment la réduction importante de l'erreur maximale MXE et des pics d'erreur en général ; du fait qu'elle soit plus lisse (effet du lissage spatial) donc facilement prévisible. Le coefficient SRL s'est amélioré à son tour. Un autre aspect impressionnant de la moyenne spatiale des vitesses : elle n'est nulle à aucun moment sur les 52 560 échantillons mesurés (soit une année avec un pas de 10 minutes), alors que la vitesse mesurée à un seul endroit s'annule parfois. Ceci prouve que la vitesse du vent ne peut être jamais nulle en tout point du parc éolien, même si elle l'est dans quelques endroits. Cet aspect a permis le calcul de l'erreur MAPE uniquement pour la moyenne spatiale, alors que ce n'est pas possible pour la vitesse mesurée par le mât, à cause de la division par zéro (voir l'équation (1.45)).

Le tableau 2.4 présente des comparatifs entre quatre méthodes, à savoir le réseau de neurones

ANN et la forêt aléatoire RF en plus de la SVR et la persistance déjà testées. Les résultats affichés concernent la moyenne spatiale des vitesses uniquement. L'ANN contient une seule couche cachée de 20 neurones, avec la fonction d'activation sigmoïde unipolaire (voir la section 1.6.3) et la rétropropagation Levenberg-Marquardt (Annexe B). La RF est caractérisée par $n_{tree} = 500$ et $m_{try} = 4$ (voir la section 1.6.5). Presque tous les critères d'évaluation achevés par la prédiction SVR sont les meilleurs. On remarque cependant que le réseau neuronal conduit à des résultats très proches. Par contre, la RF n'est pas adaptée à ce contexte de prévision.

	PER	SVR	RF	ANN
MAE (m/s)	0,77	0,55	0,59	0,55
RMSE (m/s)	1,11	0,81	0,86	0,81
MAPE (%)	16,54	11,80	13,76	12,45
MASE	1,00	0,72	0,76	0,72
MXE (m/s)	7,42	5,88	6,97	5,85
SRL	0,94	0,94	0,93	0,93

Tableau 2.4 Prédiction de la moyenne spatiale des vitesses : comparatif de 4 méthodes proposées

La même procédure a été suivie pour prédire la direction du vent d'une heure à l'avance. Les données sont celles mesurées par le mât, vu que le moyennage dans ce cas fausse les résultats (voir la section 2.4.2). De plus, on suppose qu'elle est identique en tout point du parc. La figure 2.12(c) montre les résultats de prévision de 100 heures. La première impression est que la direction n'est pas aussi aléatoire que la vitesse, vu qu'elle peut se stabiliser sur de longues périodes. Cependant, les changements sont très aigus lorsqu'ils se produisent, ce qui provoque des pics d'erreur importants. Le nuage de points de l'ensemble de test entier à la figure 2.12(d) montre des échantillons plus éparpillés que ceux de la prévision de la vitesse (figure 2.12(b)).

Les résultats sont donnés dans le tableau 2.5, contenant les mêmes critères d'erreur. Puisque les directions sont exprimées par des angles, l'erreur maximale ne peut logiquement pas dépasser 180° , ce qui veut dire littéralement la direction opposée. Donc, 360° sont ajoutés à toute erreur inférieure à -180° , et 360° sont soustraits de toute erreur supérieure à 180° avant de calculer les critères d'évaluation.

Comme c'est montré dans le tableau 2.5, l'utilisation de la même configuration de la machine SVR ne convient pas à tous les cas, et peut même conduire à des résultats pires que ceux du modèle de persistance. L'erreur MAE passe même de $15,73^\circ$ à $22,17^\circ$, ce qui veut dire une augmentation de 40 %. Afin de remédier à ce problème, deux processus d'optimisation différents peuvent être appliqués ; un pour la vitesse, et l'autre pour la direction.

	PER	SVR	RF	ANN
MAE (°)	15,73	22,17	29,71	24,81
RMSE (°)	30,30	42,17	48,81	38,76
MAPE (%)	-	-	-	-
MASE	0,75	1,06	1,42	1,18
MXE (°)	175,26	179,88	178,77	174,03
SRL	0,87	0,78	0,63	0,75

Tableau 2.5 Pr vision de la direction du vent : comparatif de 4 m thodes propos es

2.5.2 Optimisation de la pr vision

Le but de cette partie est de d terminer le choix optimal des param tres C et γ par une optimisation contrainte. Contrairement   ANN qui poss de plusieurs param tres   ajuster tels que le nombre de couches et de neurones, l'algorithme d'apprentissage et la fonction d'activation, la performance de la pr vision SVR d pend presque totalement du choix du couple (C, γ) , ce qui facilite  norm ment l'optimisation. L'algorithme de recherche par motif (PS) est choisi pour minimiser l'erreur MXE ou MAE selon le cas, en raison de sa convergence rapide (voir l'annexe C). Cet algorithme recherche le meilleur couple (C, γ)   partir des valeurs utilis es dans la section 2.5.1, comme le montre le tableau 2.6. L'optimisation suit l'organigramme de la figure 2.13, et il se termine lorsqu'une des conditions d'arr t est rencontr e. L'arr t se produit g n ralement lorsque la diff rence entre deux valeurs successives de la fonction objective (MXE ou MAE) ou du couple (C, γ) devient inf rieure   une valeur pr d finie, ou lorsque le nombre maximal d'it rations est atteint.

M�thode d'optimisation	Recherche par motif
Fonction � minimiser	$MXE = f(C, \gamma)$ ou $MAE = f(C, \gamma)$
Contraintes	$1 \leq C \leq 200$ $0,00001 \leq \gamma \leq 1$
Point de d�part	$C = 100$ $\gamma = 0,0004$

Tableau 2.6 Configuration de l'algorithme d'optimisation

Ce processus d'optimisation est appliqu    la pr vision de la moyenne spatiale des vitesses en minimisant le crit re MXE. Le choix d'un type d'erreur ou de l'autre d pend de l'am lioration apport e par celui-ci ou celui-l . Et dans le cas de la vitesse, il est pertinent de limiter l'erreur maximale au lieu de l'erreur moyenne qui est d j  suffisamment faible. La figure 2.14(a,b) rel ve une l g re am lioration dans les r sultats de pr vision, et le tableau 2.7 souligne la diminution de tous les types d'erreur, y compris MXE, gr ce au couple (C, γ) trouv .

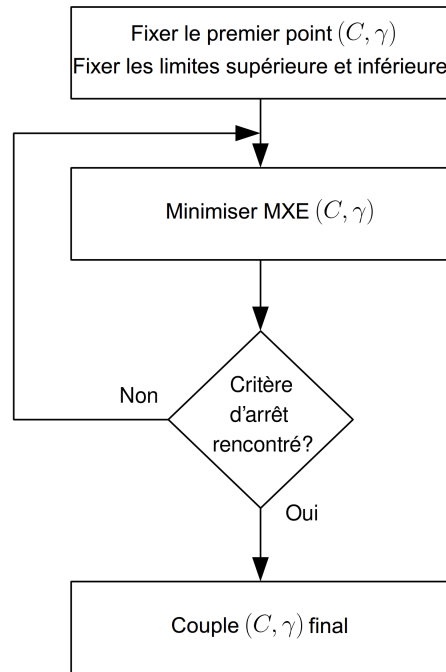


Figure 2.13 Organigramme d'optimisation

Concernant la direction du vent, le même processus d'optimisation est appliqué, excepté la fonction objective à minimiser qui est MAE. En fait, minimiser le critère MXE n'apporte pratiquement aucune amélioration, à cause des variations aiguës décrites dans la section 2.5.1. Par contre, en minimisant MAE, l'amélioration est évidente à travers la figure 2.14(c,d), montrant des courbes plus proches et des points moins éparpillés dans le graphe de régression. Le tableau 2.7 décrit une diminution significative des erreurs par rapport au modèle de persistance.

Paramètre / Critère	Moyenne spatiale des vitesses	Direction du vent
C	195,48	199,96
γ	0,01909	$2,043.10^{-5}$
MAE	0,55 m/s	12,72 °
RMSE	0,80 m/s	27,85 °
MAPE	11,58 %	-
MASE	0,71	0,60
MXE	4,52 m/s	174,97 °
SRL	0,95	0,88

Tableau 2.7 Résultats de la prévision par la machine SVR optimisée en termes de critères d'erreur

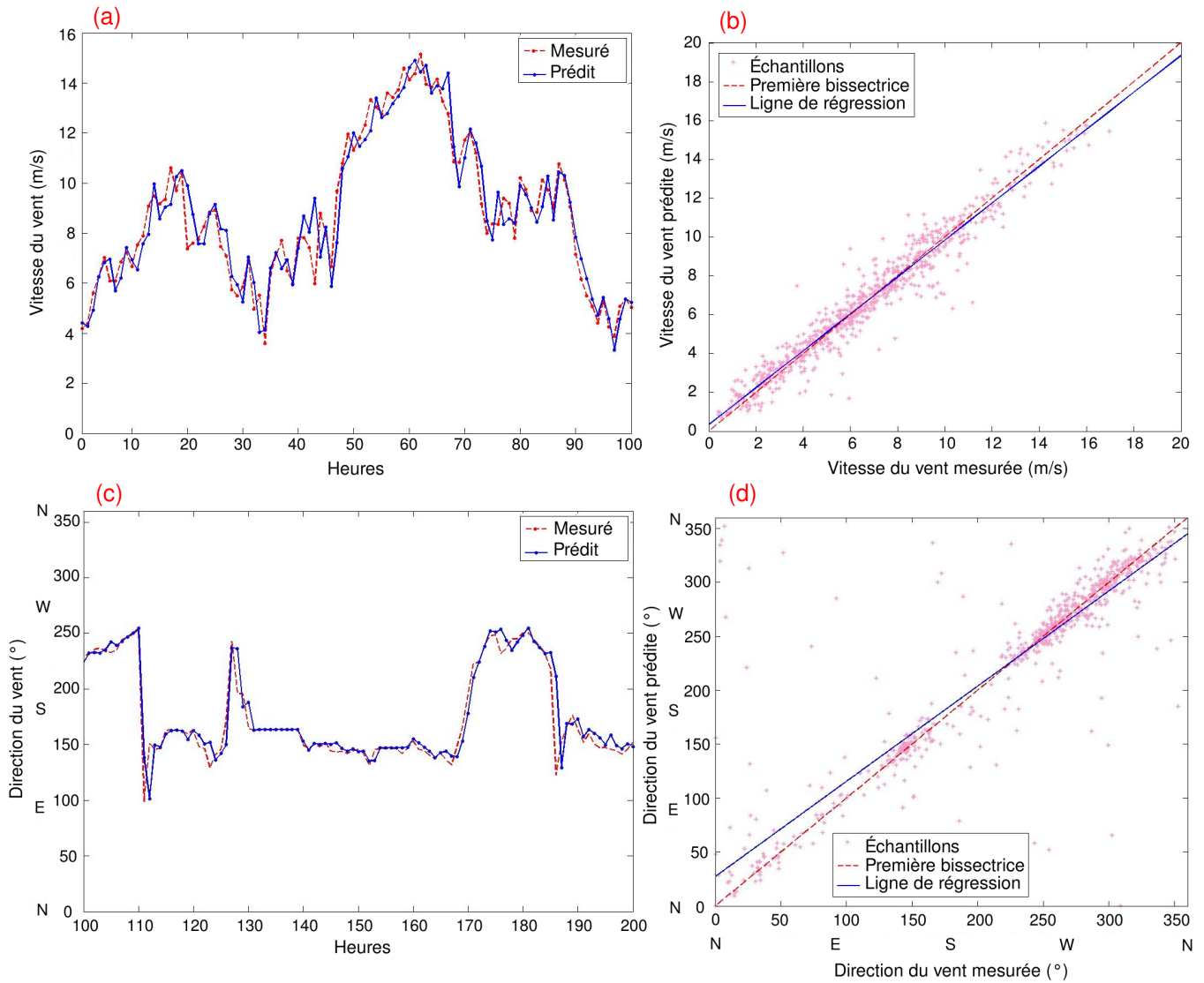


Figure 2.14 Résultats de la prévision de la vitesse (a,b) et de la direction (c,d) d'une heure à l'avance par la machine SVR optimisée, en janvier 2011

Il est difficile d'effectuer une comparaison équitable de ces résultats avec d'autres de la bibliographie, puisque les modèles de prévision sont généralement spécifiques aux sites sur lesquels ils s'appliquent et liés à l'horizon de prévision [21]. Cependant, il est connu que le critère MAPE tel qu'il est défini est relativement élevé, pouvant atteindre jusqu'à 50 % que ce soit pour la vitesse ou pour la direction. Le critère GMAPE sera abordé plus tard (section 2.6), afin de suggérer une information plus significative que celle donnée par MAPE. Le critère RMSE est jugé acceptable s'il ne dépasse pas 1,5 m/s, alors que l'erreur MAE doit être encore plus faible [14, 28, 33, 34, 41, 128].

La machine SVR optimisée a ses limitations, notamment le fait d'exiger la connaissance des valeurs mesurées pour pouvoir avancer et calculer les paramètres optimaux. Une solution possible consiste à fixer quotidiennement les meilleurs paramètres, et à les appliquer sur la prévision du jour suivant, vu qu'ils ne sont pas censés trop varier d'un jour à l'autre. Un autre problème qui se pose est qu'il n'y a aucune règle pour choisir ni la taille de l'ensemble d'apprentissage, ni la fonction objective et les contraintes. Tous ces choix restent empiriques, donc dépendant des tests de simulation.

2.5.3 Prévision d'une à dix heures à l'avance

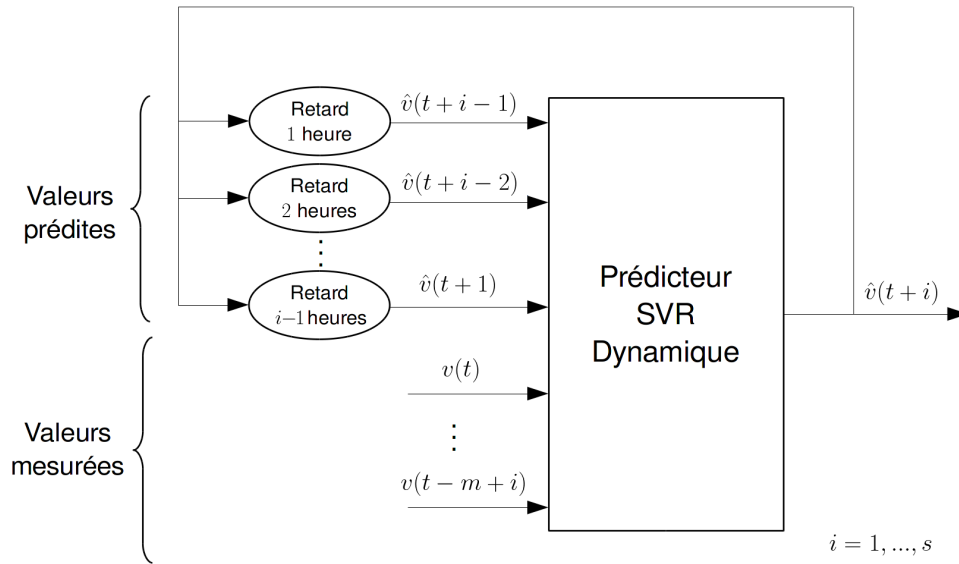


Figure 2.15 Modèle itératif (dynamique)

Le processus de prévision d'une heure à l'avance peut être appelé itérativement afin de prédire plusieurs heures à l'avance. Cette méthode, bien qu'utilisée dans la littérature [40], reste limitée et conduit dans la plupart des cas à une erreur qui grandit au fur et à mesure que l'horizon de prévision s'élargit. Il ne faut pas utiliser trop d'itérations, pour ne pas franchir les limites des séries temporelles statistiques et casser la supposition de dépendance entre le comportement futur et passé.

Un test simple est effectué en utilisant le modèle itératif, qu'on appelle aussi modèle dynamique, afin de prédire la vitesse du vent allant jusqu'à 10 heures à l'avance. Le modèle de la machine est celui de la figure 2.15. Pour prédire s heures à l'avance, on commence par prédire une heure en utilisant comme entrées les mesures des m dernières heures. Ensuite, on prédit l'heure qui suit en utilisant $m-1$ mesures et la valeur prédite précédemment et ainsi de suite jusqu'à ce qu'on termine les s heures. Bien évidemment, il faut que $m \geq s$ pour que la prévision ait un sens. Les échantillons

d'entrée ne sont pas espacés de 10 minutes comme ça été le cas dans la section 2.4.2, et ceci dans le but d'aborder plusieurs cas envisageables. Les paramètres (C, γ) sont choisis manuellement pour la simplicité. Néanmoins, il est toujours possible d'appliquer le processus d'optimisation.

Mathématiquement parlant, il s'agit de modifier le modèle de cette façon :

$$X(t+i) = (\hat{Y}(t+i-1), \dots, \hat{Y}(t+1), Y(t), Y(t-1), \dots, Y(t-m+i))^T, \quad i = 1, \dots, s \quad (2.2)$$

$Y(t) = v(t)$ ou $Y(t) = d(t)$, où $v(t)$ et $d(t)$ sont des scalaires (et non plus des vecteurs) représentant la vitesse et la direction du vent à l'heure t , respectivement. L'équation de prédiction s'écrit sous la forme :

$$\hat{Y}(t+i) = F(X(t+i)), \quad i = 1, \dots, s \quad (2.3)$$

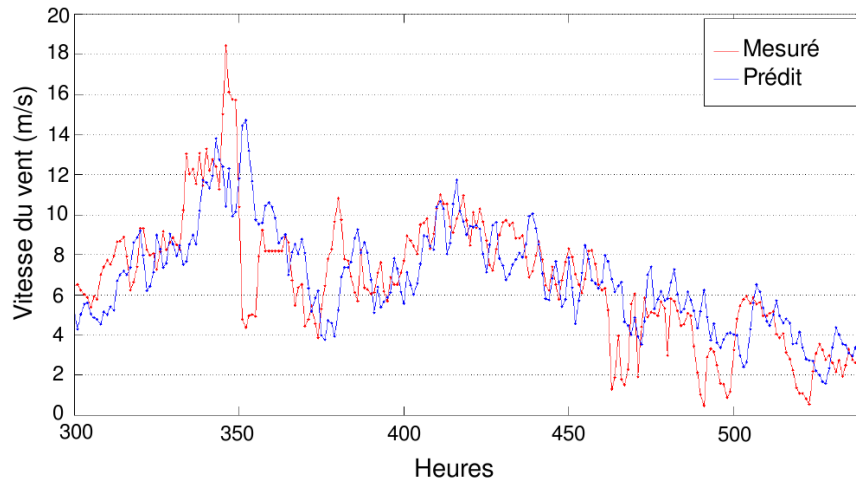


Figure 2.16 Résultats de la prédiction de la vitesse de cinq heures à l'avance par le modèle itératif de la machine SVR, de 250 heures en mars 2011

Les résultats trouvés après un apprentissage utilisant les données de janvier et février et un test sur les données de mars apparaissent sur la figure 2.16, avec $s = 5$ et $m = 10$ (soit 5 heures à l'avance en utilisant 10 échantillons passés). La vitesse du vent dans ce cas est mesurée à un seul endroit, et n'est pas la moyenne spatiale des vitesses. Les résultats étant plus ou moins précis, la courbe prédite trouve du mal parfois à suivre la courbe mesurée. La figure 2.17 montre l'évolution des critères d'erreur MAE et RMSE en fonction de l'horizon de prédiction. On voit bien qu'ils sont en hausse continue, ce qui fait que la performance se dégrade en élargissant la période à prédire. La prédiction de la direction par ce modèle est aussi possible, bien que les résultats sont souvent

médiocres. Il est souvent conseillé de corriger les résultats de ce modèle par les sorties d'un modèle NWP, afin de dépasser les limites des séries statistiques et d'aboutir à de meilleurs résultats.

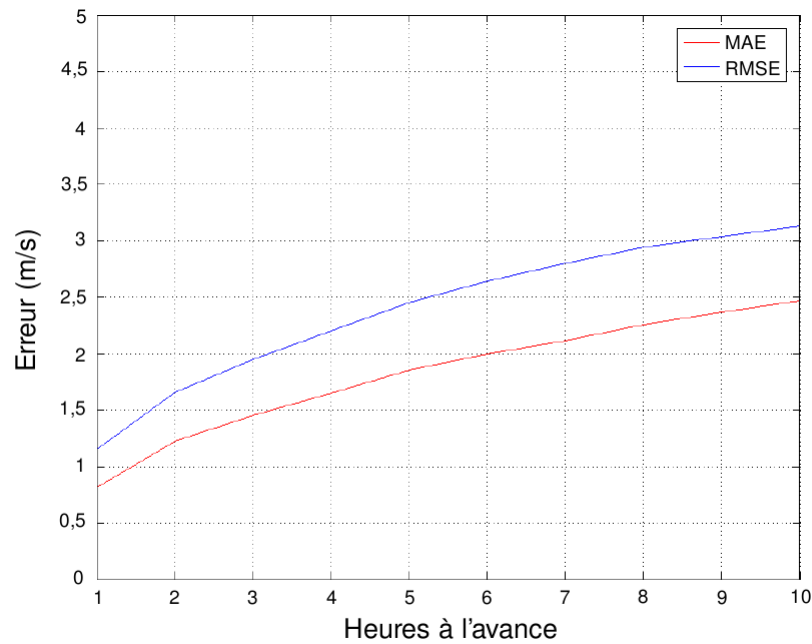


Figure 2.17 Évolution des critères MAE et RMSE

2.6 Conversion vent / puissance

Une étape très importante consiste à convertir les données de vent prédites en puissance éolienne. Puisqu'on a démontré l'effet de la direction du vent, il fallait trouver une méthode pour l'intégrer dans la conversion, au lieu d'utiliser la courbe de puissance classique qui n'exploite que la vitesse du vent. Une méthode très simple consiste à utiliser un réseau de neurones artificiels (ANN), afin de déterminer la puissance générée par le parc, étant donné la moyenne spatiale des vitesses et la direction. Le choix de ANN à la place de SVM dans ce cas se justifie par la relation vent / puissance qui est simple à apprendre (relativement, par rapport à la prévision) et qui ne requiert pas une optimisation. Et dans de telles conditions, le modèle ANN est plus adapté, où ses paramètres par défaut sont largement suffisants et permettent en général d'obtenir des résultats supérieurs à ceux des autres techniques intelligentes.

Le modèle ANN choisi est donné par la figure 2.18. Les entrées utilisées sont les résultats de prévision de la section 2.5.2. Les ensembles d'apprentissage et de test ainsi que l'horizon de prévision sont les mêmes (août à décembre pour l'apprentissage, janvier pour le test, horizon d'une heure à l'avance $s = 1$). Les résultats de la conversion sont donnés par la figure 2.19, où la courbe prédite

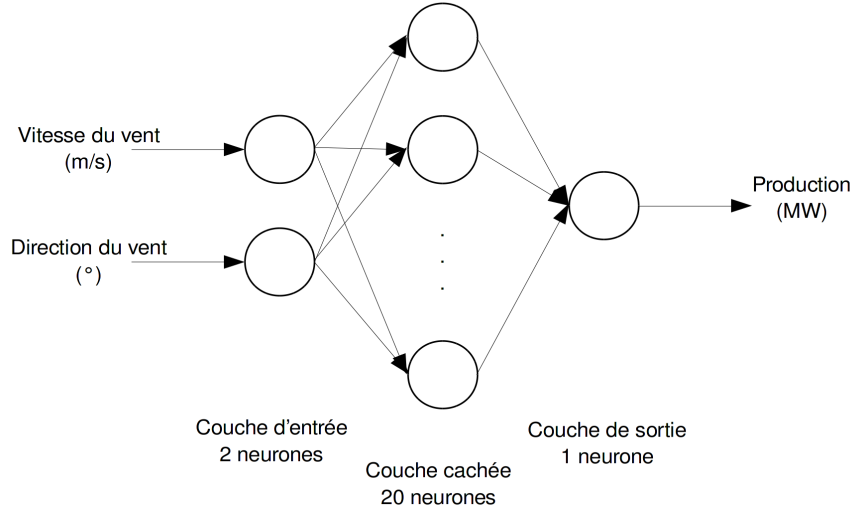


Figure 2.18 Structure du réseau de neurones artificiels

est celle de la puissance convertie à partir des données de vent prédites. La courbe réelle correspond à la moyenne horaire de la puissance totale (moyenne de chaque 6 valeurs consécutives). Afin de mieux interpréter les résultats, on va utiliser GMAPE (voir l'équation (1.48)). Ce nouveau critère sert à mieux quantifier l'erreur de prévision sous la forme d'un pourcentage (erreur relative) et donc à surpasser les limitations de la définition mathématique abstraite. La GMAPE est calculée en divisant par la puissance totale installée du parc, et non pas par la valeur mesurée de la puissance. Définie ainsi, il n'était pas possible d'en parler lorsqu'il s'agissait de la prévision de la vitesse ou de la direction du vent.

$$MAPE = \frac{1}{k} \sum_{t=1}^k \frac{|\hat{P}_o(t) - P_o(t)|}{P_o(t)} \times 100 \quad (2.4)$$

$$GMAPE = \frac{1}{k} \sum_{t=1}^k \frac{|\hat{P}_o(t) - P_o(t)|}{P_{o_{installée}}} \times 100 \quad (2.5)$$

$P_o(t)$ et $\hat{P}_o(t)$ sont la puissance mesurée et la puissance prédite à l'heure t , respectivement (On utilise la notation P_o pour la puissance au lieu de P pour ne pas la confondre avec le prix). $P_{o_{installée}}$ est la puissance totale installée du parc, et k est le nombre d'heures prédites ($k = \beta.s = 743 \times 1 = 743$). Afin de mieux comprendre l'utilité de GMAPE, prenons l'exemple d'un parc éolien de 50 MW de puissance installée. Et supposons que le prédicteur ait prédit 1 MW, alors que la puissance réelle fournie était de 2 MW (cas d'un vent très faible). Le critère MAPE donnerait une erreur de

50 %, alors que le critère GMAPE donnerait 2 %. On voit bien que GMAPE reflète mieux l'exactitude de la prévision, alors que MAPE tend parfois vers des valeurs très élevées. Tous les critères d'évaluation de la prévision sont groupés dans le tableau 2.8 pour la méthode SVR + ANN, et sont comparés à ceux du modèle de persistance.

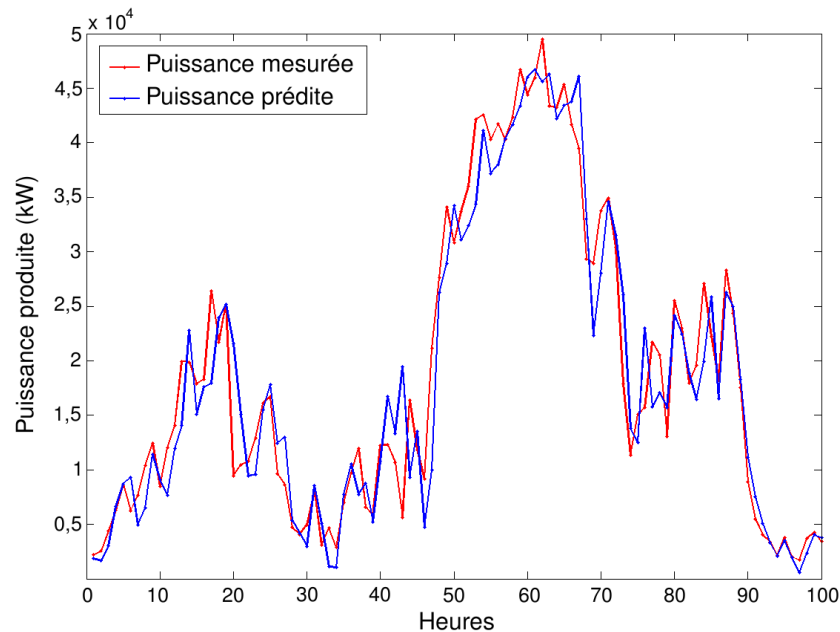


Figure 2.19 Résultats de la prévision de la puissance éolienne d'une heure à l'avance par la machine SVR et le modèle ANN, de 100 heures en janvier 2011

Critère	Prévision SVR + ANN	Persistance
MAE (kW)	2089,7	2454,6
MXE (kW)	21 378	27 114
RMSE (kW)	3457,8	4227,3
MAPE (%)	31,0025	32,0272
GMAPE (%)	3,9059	4,5880
SRL	0,9622	0,9622

Tableau 2.8 Critères d'évaluation de la prévision de la puissance éolienne, janvier 2011

Bien que l'apport de la méthode par rapport au modèle de persistance soit minime, cette amélioration peut se traduire par des bénéfices énormes en termes de coûts de production, surtout à grande échelle. L'avantage de l'utilisation de GMAPE par rapport à MAPE est bien évident, se traduisant par une erreur relative de 3,9 % au lieu de 31 %, ce qui est plus significatif en tant qu'évaluation. La réduction la plus importante réalisée par ce modèle est celle de l'erreur maximale MXE, qui a at-

teint 21 % par rapport au modèle de persistance (18 % pour RMSE et 14 % pour MAE et GMAPE). Les valeurs trouvées des critères d'erreur sont dans les limites reconnues [12].

2.7 Conclusion

La prévision de l'énergie éolienne est possible en utilisant les techniques d'intelligence artificielle, telles que ANN et SVR, qui surpassent les méthodes statistiques classiques et évitent les modèles NWP complexes. La technique SVR est utilisée dans ce chapitre, après une présentation détaillée du parc Sidi Daoud, afin de prédire l'énergie éolienne via une approche indirecte. Même si la vitesse du vent est le facteur qui influe le plus sur la puissance générée par un parc éolien, d'autres facteurs ont aussi leurs contributions, notamment la direction du vent et la disposition des aérogénérateurs. Cette influence est démontrée même pour les nacelles auto-orientables. De ce fait, un modèle est construit afin de prédire la direction du vent en parallèle avec la vitesse. La prévision de la vitesse par SVR a donné de bons résultats par rapport au modèle de persistance, notamment après l'introduction du concept de la moyenne spatiale des vitesses. Cette moyenne, s'inspirant de l'effet de lissage spatial, a permis la réduction de l'erreur maximale, et donc des pics d'erreurs. La prévision est quantifiée par plusieurs critères d'erreur qui ont été proposés. Cependant, l'application de la même configuration de la machine SVR à la prévision de la direction n'était pas appropriée. Donc, une optimisation du couple (C, γ) a été proposée afin de déterminer la configuration optimale, en minimisant une fonction objective bien choisie. Cette optimisation a abouti à des réductions importantes de tous les types d'erreur. La meilleure réduction est celle du critère MXE de la vitesse, réduit à 61 % par rapport au modèle de persistance, après le moyennage spatial et l'optimisation. Néanmoins, l'approche proposée a ses limitations ; il peut être nécessaire de répéter quotidiennement le processus d'optimisation, plusieurs choix restent empiriques, etc. Malgré ces inconvénients, il s'avère que la méthode SVR optimisée par l'algorithme PS est le choix le plus pertinent dans le cas de la prévision éolienne au sein du parc Sidi Daoud.

Une prédiction itérative a permis, à la fin de ce chapitre, d'étaler l'horizon de prévision à des heures à l'avance, allant jusqu'à 10 heures. Ce procédé conduit en général à des erreurs qui grandissent au fur et à mesure, et il est préférable qu'il soit accompagné par une méthode NWP. Finalement, la conversion des données de vent, à savoir la vitesse et la direction prédites, en puissance générée constitue la dernière partie du chapitre. Ne se contentant pas uniquement de la courbe de puissance, un réseau de neurones artificiels simple est utilisé afin de déduire la production du parc éolien. Une fois la prévision de la production renouvelable est effectuée avec précision, c'est au tour maintenant de la prévision de la charge. La relation entre les deux types de prévision est reprise en détails au prochain chapitre, qui s'attaque à la prévision de la demande d'électricité en général et qui propose un modèle de prédiction pour des horizons plus larges.

CHAPITRE 3 PRÉVISION DE LA DEMANDE

3.1 Introduction

Après avoir présenté en détails la procédure de prévision de l'énergie éolienne, c'est la demande d'électricité qui est mise en évidence dans ce chapitre. On rappelle que cette prévision est cruciale dans le processus d'équilibrage de l'offre et de la demande, et que son exactitude a des conséquences sur les décisions de dispatching. Les différents défis liés à l'intégration des ressources renouvelables ainsi que l'importance de couplage des deux types de prévision seront soulignés. Un modèle sera ensuite conçu pour la prévision d'une journée à l'avance par un pas temporel d'une heure, tout en respectant les spécificités du contexte tunisien ; petite surface du pays, climat modéré, marché non dérégulé et présence limitée des ressources intermittentes. La construction du modèle se base sur l'analyse du profil de charge et l'extraction des caractéristiques telles que la fonction d'autocorrélation. Le choix des entrées du modèle sera assuré par des mesures de corrélation et de validation croisée. L'apprentissage de la machine est en ligne, ce qui permet de prendre en compte les transitions entre les saisons. Les entrées seront raffinées par un filtrage spécial appelé sélection experte, afin de permettre à l'utilisateur du modèle d'ajouter ses propres règles concernant le climat ou le marché et généraliser davantage l'aptitude de prévision.

Le chapitre vise à répondre à ces questions en particulier : la forêt aléatoire, associée à la sélection experte des entrées, est-elle apte à manipuler n'importe quel profil de charge, et à s'adapter au comportement complexe des consommateurs ? L'approche proposée est-elle applicable sur les quatre saisons et sur les jours particuliers, comme les weekends et les jours fériés, qu'ils soient mobiles ou fixes ? Comment s'adapte-t-elle aux jours fériés religieux qui suivent le calendrier hégirien ?

3.2 Importance et défis de la prévision de la demande

La prévision de la demande d'électricité, qu'on appelle également prévision de la charge, devient de plus en plus difficile à cause de la dérégulation du marché d'électricité et l'émergence des ressources renouvelables. Pour faire face aux défis imposés par l'incertitude croissante et assurer une prévision précise pour différents horizons, plusieurs techniques intelligentes ont été développées au sein du smart grid. Plusieurs tâches des gestionnaires de réseaux, notamment le dispatching, dépendent de l'exactitude de la prévision. Après avoir prédit la production renouvelable dans le chapitre précédent, ce chapitre propose un modèle de prévision de la charge à court terme utilisant la technique des forêts aléatoires. Caractérisée par son immunité aux entrées non significatives et sa validation croisée interne, la forêt aléatoire est un choix pertinent pour les modèles de prévision

non paramétriques.

3.2.1 Importance et avantages de la prévision précise

La prévision de la demande d'électricité a été toujours un défi pour les gestionnaires de réseau. C'est un thème de recherche ancien qui a suivi l'évolution des installations de puissance et les techniques de calcul intelligentes. Il devient de plus en plus intéressant depuis le début du 21^e siècle à cause de l'émergence des ressources d'énergie renouvelable et des smart grids. La prévision de la demande est la prédiction de l'évolution future du signal de puissance consommée par un appartement, un micro-réseau, une région ou même un pays entier. La prévision précise d'une à quelques heures à l'avance est très importante pour le dispatching d'énergie, la gestion des unités de production et les échanges énergétiques. Prédire la charge pour des horizons plus larges sert à planifier la maintenance des centrales et à définir les politiques d'exploitation d'énergie. L'exactitude de la prédiction est tellement importante qu'une augmentation de 1 % dans l'erreur de prévision peut induire une hausse spectaculaire dans les prix opérationnels [76], ce qui fait que la plus petite amélioration peut faire la différence. L'erreur peut être une surestimation ou une sous-estimation de la charge, et les deux entraînent des difficultés dans l'équilibrage de l'offre et la demande d'électricité. Le concept de smart grid sous-entend l'utilisation des techniques de calcul intelligentes, y compris la prévision, afin de gérer l'offre d'électricité et l'égaliser à la demande en temps réel. La gestion de l'offre est liée à la réserve de puissance (spinning reserve), définie par la capacité de production synchronisée (tournante) moins la demande et les pertes. Prédire la charge revient donc à prédire la réserve de puissance qui est indispensable dans des cas de demande énorme et soudaine, coupures de lignes ou défaillance de quelques générateurs. Lorsque cette prévision est précise, la réserve de puissance peut compenser rapidement les déficiences potentielles. Sur le long terme, la détermination d'un profil de charge par prévision sert à estimer la puissance à ajouter (où à enlever) au réseau afin d'éviter les blackouts et le gaspillage d'énergie.

3.2.2 Défis de la prévision

La prévision de la charge devient plus difficile qu'auparavant à cause de deux raisons principales : la dérégulation du marché et l'intégration de l'énergie renouvelable.

Effet de la dérégulation

La dérégulation et la privatisation du marché d'électricité dans plusieurs pays du monde signifie que les consommateurs sont libres de choisir leur fournisseur parmi plusieurs. Cette liberté induit systématiquement un prix d'électricité variable, ce qui pousse les clients à préférer celui qui

propose le meilleur prix et à décaler leur consommation en dehors des heures de pointe. Dans des cas pareils, de nouvelles stratégies de prévision basées sur des analyses anthropologiques sont requises [129], ainsi que la prévision des prix en parallèle. La dérégulation induit aussi l'éparpillement des consommateurs, ce qui conduit à des courbes de puissance non lisses et faiblement corrélées avec les conditions météo. Cependant, ce n'est pas encore le cas du réseau tunisien, où la privatisation est encore limitée. En 2013, 3046 GWh ont été générés par des entreprises privées, alors que 13 947 GWh étaient fournis par la Société Tunisienne de l'Électricité et du Gaz (STEG), entreprise étatique. Le prix d'électricité est donc fixé par l'état et ne dépend pas du fournisseur ou de la période de consommation, et le problème de dérégulation ne se pose pas. Néanmoins, la privatisation du secteur ne cesse d'augmenter, et la dérégulation du marché serait probablement un choix inévitable.

Effet des ressources renouvelables

La haute pénétration des ressources intermittentes dans le réseau électrique, surtout les énergies éolienne et solaire, augmente le degré d'incertitude de la prévision à cause de leur comportement non régulier. Une ressource intermittente peut plus ou moins affecter le réseau, selon son degré de pénétration [130]. L'énergie éolienne étant intégrée au réseau tunisien, doit être prédite en parallèle avec la demande, et c'est ce qui a été expliqué en détails dans le chapitre 2. Lorsque ces deux types de prévision sont couplés, ils assurent une meilleure gestion de la charge résiduelle, c'est-à-dire la charge qui doit être produite par les centrales classiques. La charge résiduelle est la demande en électricité moins la production renouvelable, et donc par définition elle a besoin des deux prévisions pour être estimée. La petite surface de la Tunisie ainsi que la puissance installée limitée à 4425 MW permettent de prédire la demande du pays entier sans avoir à le diviser en zones. Il est également possible d'utiliser les données météo d'une seule station située à Tunis, la capitale et la ville la plus gourmande en énergie, au lieu de recourir à plusieurs stations éparpillées [70, 131].

3.3 Conception du modèle de prévision de la charge

3.3.1 Analyse du profil de charge

Le signal de charge est une série temporelle complexe pilotée par plusieurs facteurs. Il possède une tendance générale lente avec un comportement quotidien pseudo-périodique. La tendance générale est observée à travers la demande accrue chaque année. L'augmentation annuelle de la consommation est plus ou moins prévisible, par l'extrapolation de son évolution durant les années passées, comme le montre la figure 3.1. Les principales causes sont l'urbanisation, l'accroissement de la population, la commerce des appareils électroménagers notamment les climatiseurs, etc. Tous ces

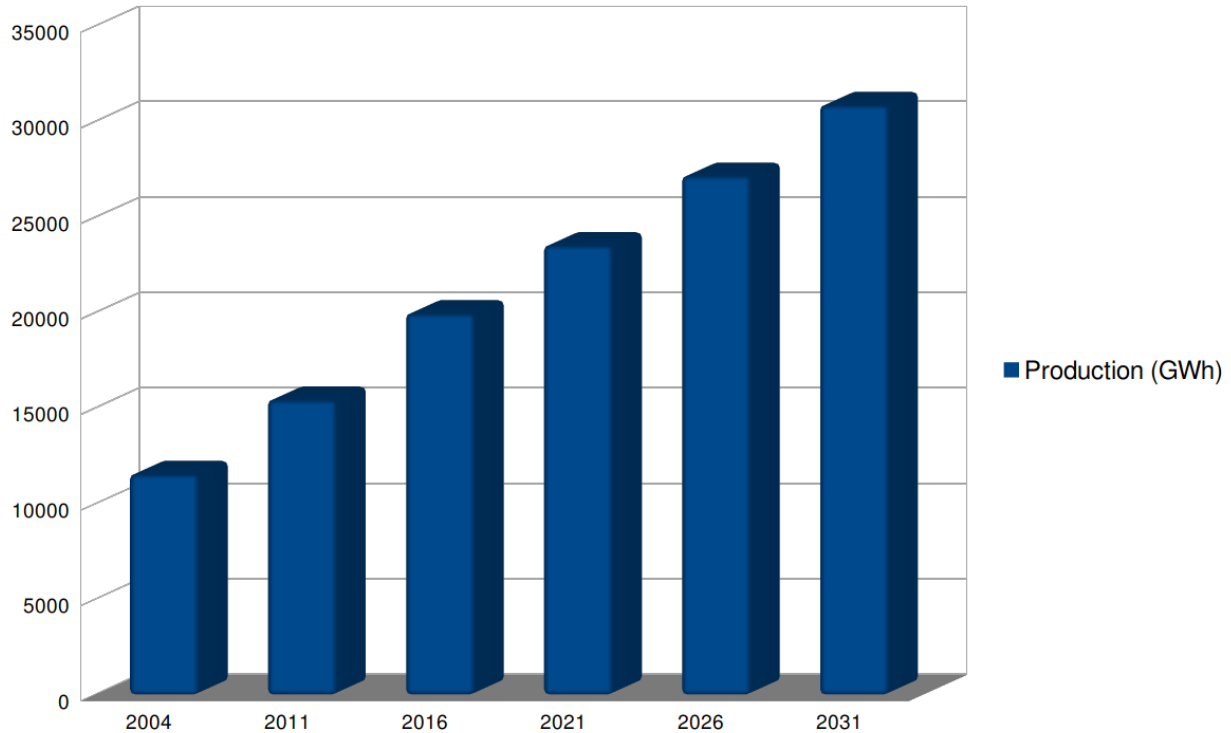


Figure 3.1 Évolution estimée de la production annuelle d'électricité en Tunisie

facteurs influent sur le long terme, et ne sont donc pas inclus dans le modèle proposé. La composante périodique, également appelée saisonnalité, apparaît à des niveaux différents. Au niveau saisonnier, il y a une augmentation spectaculaire chaque été à cause des canicules. Cette augmentation apparaît sur la figure 3.2(a) sous forme de crêtes cassant le rythme de croissance annuel lent. Au niveau hebdomadaire, les profils des weekends sont répétitifs. Sur la figure 3.2(b), le samedi a un comportement de demi-journée de travail, et le dimanche a le comportement d'une journée de repos. Au niveau quotidien, les profils du lundi au vendredi sont similaires, où le travail est à plein temps. Les jours fériés sont un peu particuliers ; ils sont semblables aux dimanches sauf pour certaines fêtes religieuses surtout en été. En effet, la courbe de charge est étroitement liée à la culture tunisienne.

Il est possible également de détecter les saisonnalités par une décomposition en ondelettes. En décomposant par exemple le signal de charge du 1^{er} au 30^e janvier 2013 (soit 720 heures) en ondelette de Daubechies jusqu'au niveau 4, on obtient les signaux qui apparaissent sur la figure 3.3. Le signal détail d4 présente une saisonnalité très claire. De plus, cette décomposition peut faciliter la prévision, surtout lorsqu'on tombe sur des signaux presque réguliers. L'inconvénient dans ce cas est qu'on ne connaît pas quel type d'ondelette choisir, et jusqu'à quel niveau il faut décomposer.

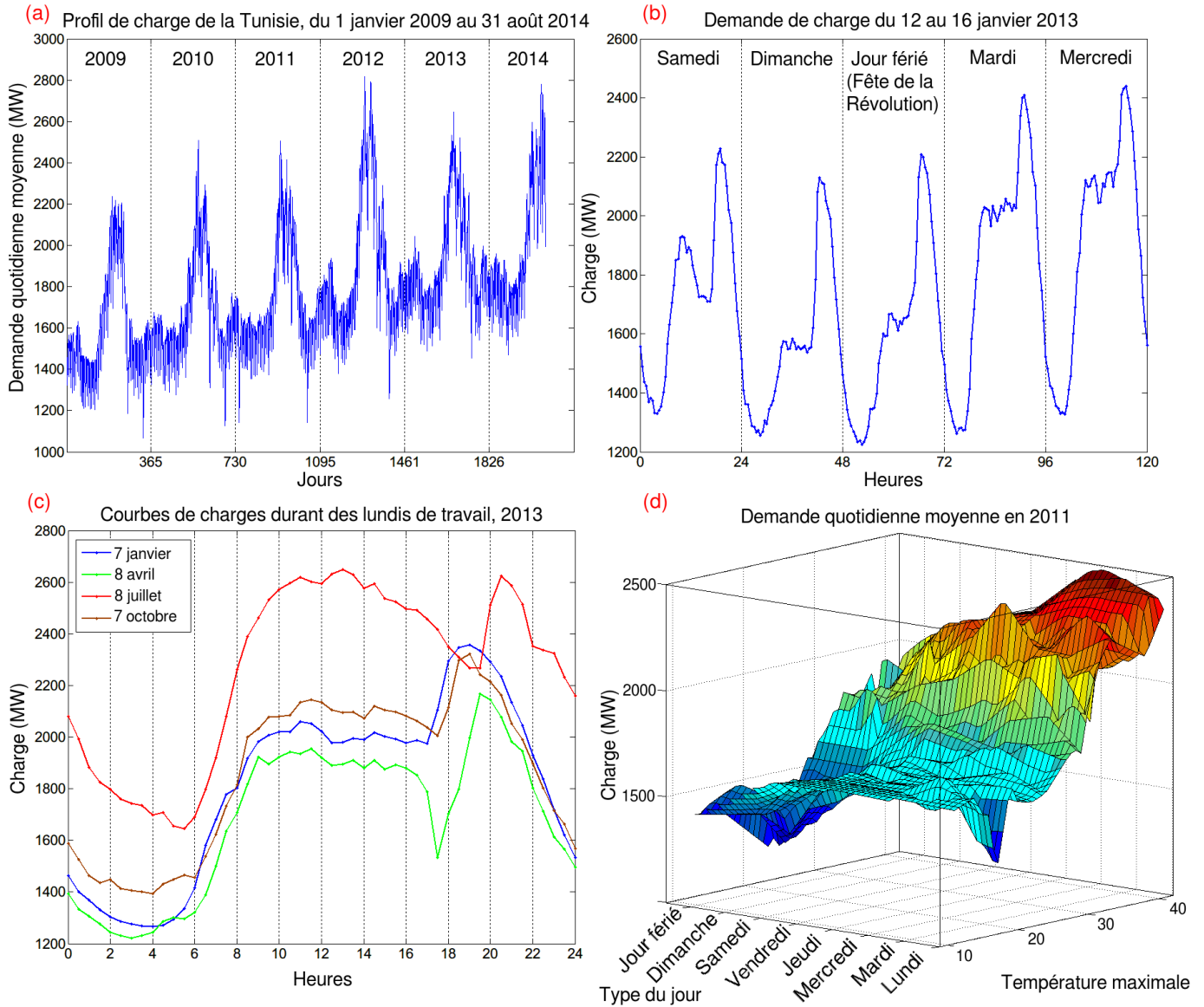


Figure 3.2 Profil de charge

En analysant la figure 3.2(c), il est possible de distinguer la hausse entre 6h et 8h du matin, qui correspond au début du travail de la journée. Ensuite, la demande se stabilise jusqu'à 17h environ, sauf en été où un pic plat apparaît entre 10h et 14h à cause de la climatisation excessive. Finalement, on voit le pic du soir qui est décalé selon les saisons, proportionnellement au nombre d'heures du jour. On remarque que, moyennement, le printemps est la saison qui consomme le moins. Ceci était le profil de consommation des jours de travail pendant plusieurs années.

La figure 3.2(d) montre l'évolution de la charge quotidienne en fonction de la température maxi-

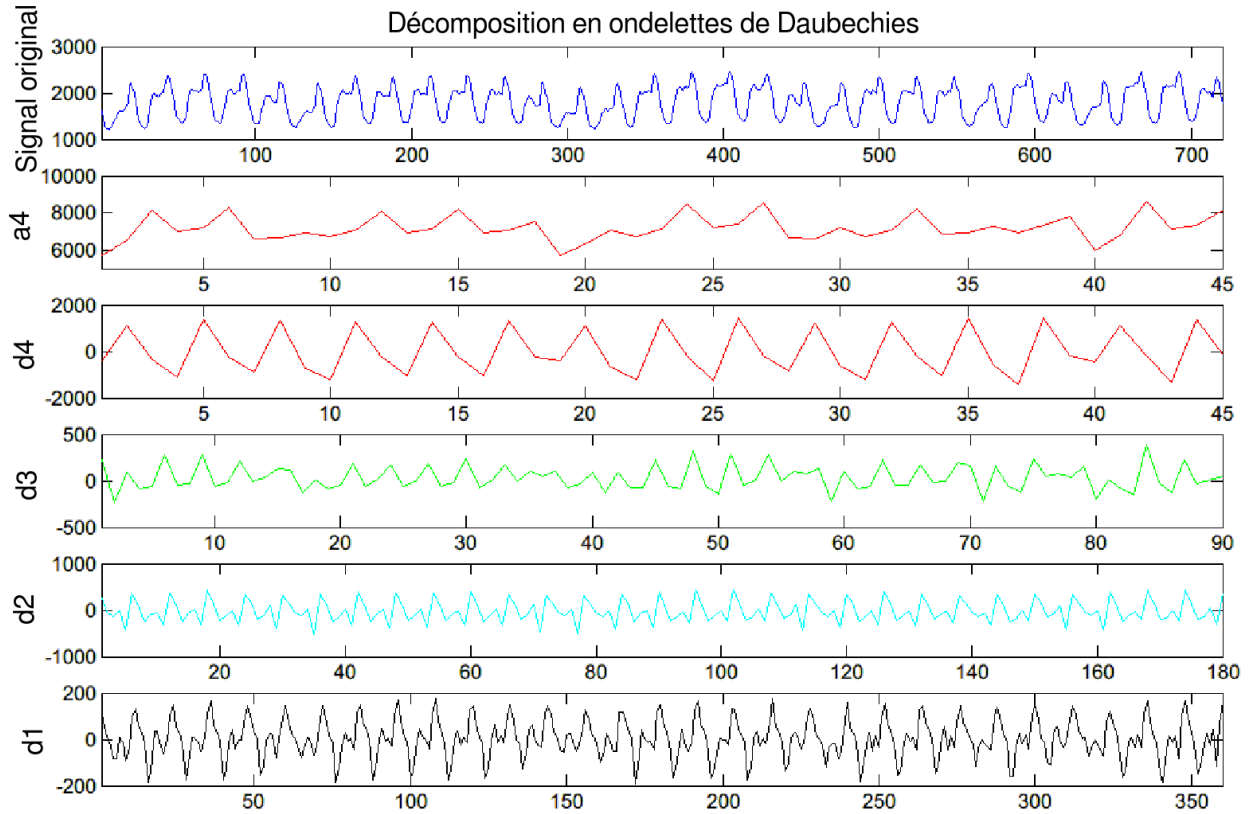


Figure 3.3 Décomposition en ondelettes du signal de charge

male et le type du jour, par une surface d'extrapolation. Évidemment, dans le cas de la Tunisie et des régions chaudes en général, la température élevée induit une consommation élevée. Le type du jour est aussi d'une importance capitale dans la détermination de la demande. À part les facteurs mentionnés dans cette section, il y a quelques événements qui se produisent rarement et qui influent à leur tour, ou plutôt perturbent la demande ; comme les troubles politiques, les blackouts et les canicules extrêmes. Dans de tels cas, il n'est pas possible de déterminer un profil de charge.

3.3.2 Choix de la méthode et des entrées

La méthode choisie pour la prévision de la charge est une méthode statistique à court terme, qui prédit la demande d'électricité sur un horizon d'une journée à l'avance ($s = 24$) par le modèle RF (on rappelle que s est la longueur de l'horizon de prévision en heures). L'horizon de 24 heures est choisi en analysant le profil de charge de la section 3.3.1, à travers lequel on voit une saisonnalité quotidienne très forte. Contrairement à la vitesse et la direction du vent qui sont très aléatoires, le signal de charge est beaucoup plus lisse et régulier, ce qui permet d'élargir la prévision à 24 heures

et plus. Dans de telles conditions, la forêt aléatoire est un choix pertinent pour la prévision. La RF est peu sensible à la variation de ses paramètres, et elle ne nécessite pas donc une optimisation. Bien évidemment, les autres méthodes (ANN et SVM notamment) donnent des résultats similaires lorsqu'elles sont optimisées par des métaheuristiques (voir le chapitre 1), mais cette optimisation n'est pas toujours évidente et facile à réaliser. La RF donne son meilleur résultat juste avec des paramètres génériques recommandés par le concepteur la méthode. Notons bien que la RF n'est pas adaptée à la prévision éolienne, puisque la nature du signal (très chaotique) exige une méthode optimisée.

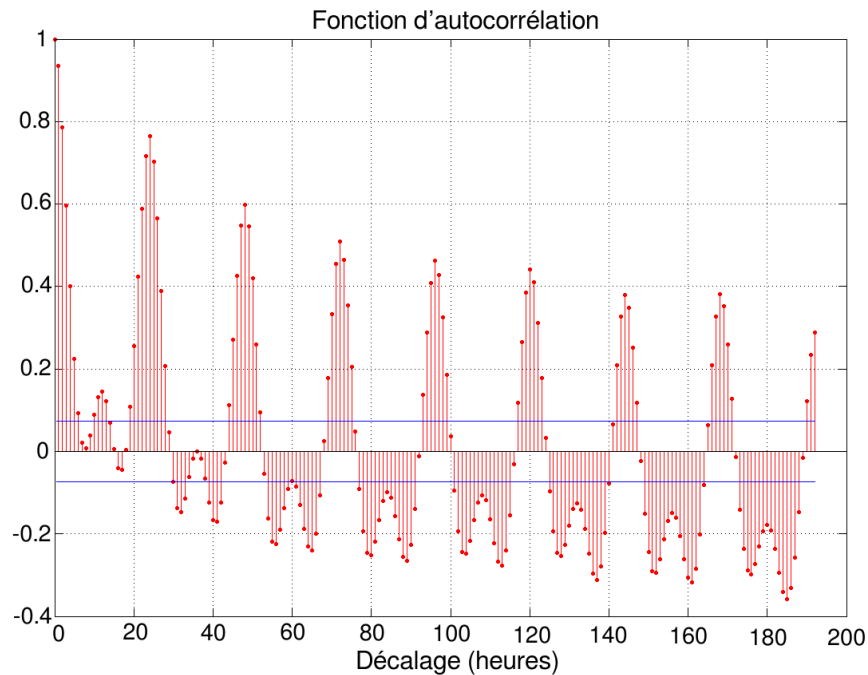


Figure 3.4 Courbe d'autocorrélation du signal de charge

La courbe d'autocorrélation de la figure 3.4 est utilisée pour détecter les similitudes entre le signal de charge et ses versions décalées dans le temps. Le but est de déterminer les échantillons passés les plus influents sur l'évolution future. Clairement, les valeurs de la charge à la même heure aux jours précédents ont la plus grande similarité, soit un décalage de 24 heures. Ceci amène à construire 24 modèles de prévision correspondant aux 24 heures du jour, au lieu d'utiliser un seul modèle appelé itérativement 24 fois. Le modèle itératif conduit généralement à une erreur croissante, et c'est ce qui a été prouvé au chapitre 2. Il n'est pas possible également d'utiliser un modèle MIMO (plusieurs entrées plusieurs sorties), puisque la forêt aléatoire n'admet qu'une seule sortie. La décomposition en ondelettes peut résoudre ce problème ; en utilisant la décomposition de la figure 3.3, il suffit de prédire un seul échantillon futur sur tous les signaux pour avoir une prédiction de 32 heures sur le signal original. Ceci est dû aux caractéristiques de la reconstruction, qui met des zéros entre les

échantillons puis effectue un filtrage. Mais cette démarche n'est pas adoptée, puisqu'elle minimise la taille de l'ensemble d'apprentissage par le sous-échantillonnage. Finalement, c'est le choix de 24 modèles qui a été adopté. Donc, afin de prédire la charge à une heure bien déterminée d'un jour quelconque, les valeurs de charge des deux jours précédents à la même heure sont utilisées comme entrées, aussi bien que les deux pics du matin et du soir du jour précédent. Ces quatre entrées sont endogènes, puisqu'elles sont liées aux échantillons passés (voir la section 1.2.4). De plus, elles sont bien justifiées par des mesures de corrélation et d'importance. Les facteurs externes choisis sont le mois, le type du jour, les températures minimale et maximale du jour à prédire. Ces facteurs sont supposés être les plus influents sur la demande, et ils sont classés comme des entrées exogènes. La figure 3.5 schématise le modèle ainsi proposé, où d représente le jour (à ne pas confondre avec la direction du vent du chapitre 2), et h représente l'heure à prédire avec $h = 0, \dots, 23$ (on n'utilise pas t car t représente l'heure absolue, alors que h ne dépasse pas 23). Le tableau 3.1 résume toutes les entrées choisies, leurs types, plages de variation, corrélations avec la sortie et mesures d'importance pour l'année 2013. La mesure d'importance, qui est une caractéristique intrinsèque des forêts aléatoires, est expliquée en détails dans la section 1.6.5 (voir aussi l'équation (1.42)).

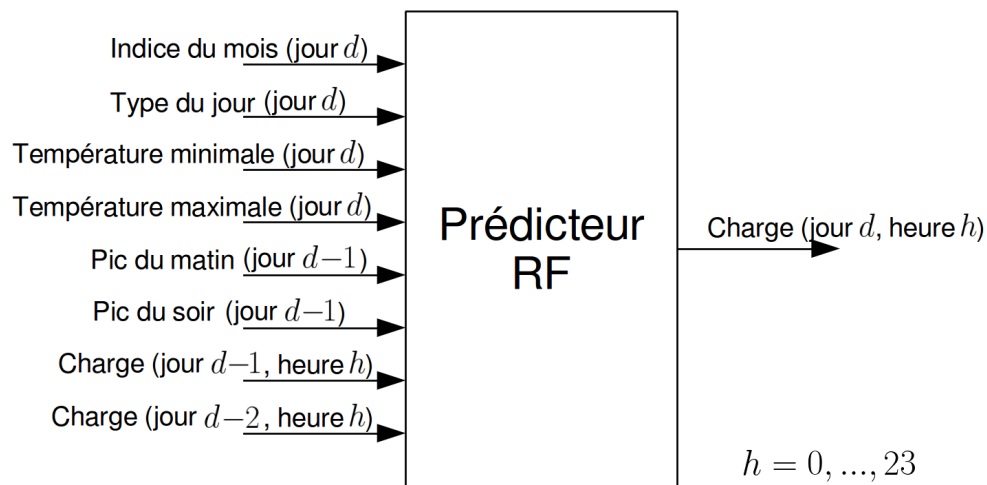


Figure 3.5 Modèle du prédicteur RF

Quelques entrées sont peu corrélées avec la sortie, ou ont une mesure d'importance faible, notamment l'entrée n°1. Elles peuvent donc être supprimées sans risque. Cependant, elles sont intentionnellement gardées afin de démontrer l'immunité de la machine face aux entrées non significatives, voire perturbantes. Il est aussi important d'indiquer que les températures du jour à prédire sont mesurées afin de valider le modèle, puisque les ensembles d'apprentissage et de test sont créés à partir de données historiques. Mais dans des cas réels, les températures sont aussi prédites, et une mauvaise prédiction de celles-ci entraîne une mauvaise prédiction de la charge. Quelques cher-

Entrée	Gamme	Corrélation avec la sortie	Importance (x 10 ⁶)
Entrées exogènes (liées au jour à prédire)			
1. Indice du mois	1 ... 12	0,1671	0,2186
2. Type du jour	1000 : lundi ... 7000 : dimanche >8000 : jour férié	-0,4663	6,7462
3. Température minimale	2 .. 29 °C	0,5289	0,5306
4. Température maximale	7 .. 43 °C	0,5453	1,2864
Entrées endogènes (liées aux jours précédents)			
5. Pic du matin du jour précédent	0 .. 4000 MW	0,7858	7,1279
6. Pic du soir du jour précédent	0 .. 4000 MW	0,5446	2,4659
7. La charge avant 24 heures	0 .. 4000 MW	0,7741 (à 10h)	7,8397 (à 10h)
8. La charge avant 48 heures	0 .. 4000 MW	0,6364 (à 10h)	1,8203 (à 10h)

Tableau 3.1 Les entrées choisies du modèle de prévision

cheurs effectuent la prévision de la température avant celle de la charge, en recourant par exemple à deux prédicteurs en cascade [76]. Mais en général, le service météo est largement suffisant pour cet horizon de prévision.

La mise en équation est alors la suivante (avec $m = 8$ et $s = 24$) :

$$\left\{ \begin{array}{l} X(t+i) = (X^1(t+i), \dots, X^m(t+i))^T, \quad i = 1, \dots, s \\ X^7(t+i) = Y(t+i-24), \quad i = 1, \dots, s \\ X^8(t+i) = Y(t+i-48), \quad i = 1, \dots, s \\ \hat{Y}(t+i) = F(X(t+i)), \quad i = 1, \dots, s \end{array} \right. \quad (3.1)$$

3.3.3 Stratégie adoptée et sélection experte

Le processus d'apprentissage de la machine est en ligne, tel que montré par la figure 3.6. Cela veut dire que pour chaque nouvelle journée de l'ensemble de test qui s'étend du jour a au jour b , un nouvel apprentissage est effectué de sorte que les valeurs réelles de charge de la journée précédente soient considérées. Dans ce cas, la période d'entraînement s'agrandit et chevauche la période de test. Cette démarche sert à surmonter les variations brusques dans le signal de charge surtout entre

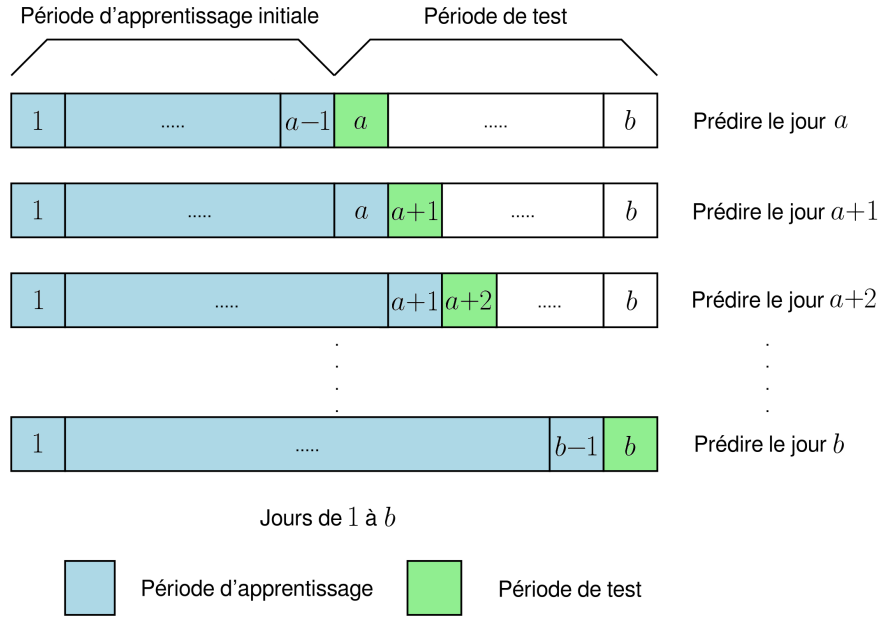


Figure 3.6 Processus d'apprentissage en ligne

les saisons. Au contraire, un apprentissage hors ligne signifie que l'apprentissage est effectué une seule fois avant le test. La période d'entraînement est alors fixe et séparée de la période de test. Ce type d'apprentissage est beaucoup plus rapide qu'un apprentissage en ligne. Afin de quantifier les différences entre les deux, un premier test de prévision a été effectué à titre de comparaison en termes de performance et de temps de calcul, avant de procéder à la sélection experte et la prévision sophistiquée. Les résultats, qui apparaissent au tableau 3.2 et sur la figure 3.7, montrent bien que l'apprentissage en ligne permet de surmonter quelques événements inattendus, tels que les journées chaudes. Cependant, il prend un certain temps de calcul même sur un ordinateur équipé d'un processeur i7.

Mathématiquement parlant, par un apprentissage hors ligne on effectue s apprentissages, puisqu'on dispose de s modèles de prédiction (on rappelle que $s = 24$). Par l'apprentissage en ligne, $\beta \cdot s$ apprentissages sont effectués, où β est le nombre de prédictions (dans ce cas, β est le nombre de jours à prédire dans la période de test). L'ensemble d'apprentissage S_n s'agrandit et ajoute une observation après chaque prédiction. À la fin, S_n doit contenir $n + \beta - 1$ observations. Évidemment, chaque modèle parmi les 24 modèles possède son propre ensemble S_n .

L'organigramme de la figure 3.8 détaille la stratégie de prévision adoptée. Deux boucles imbriquées représentent les routines de prévision horaire et quotidienne. Les phases d'apprentissage et de test sont exécutées en utilisant le code original de la forêt aléatoire [132]. La sélection des ensembles d'apprentissage et de test est d'une importance critique, puisqu'elle permet de modifier les entrées

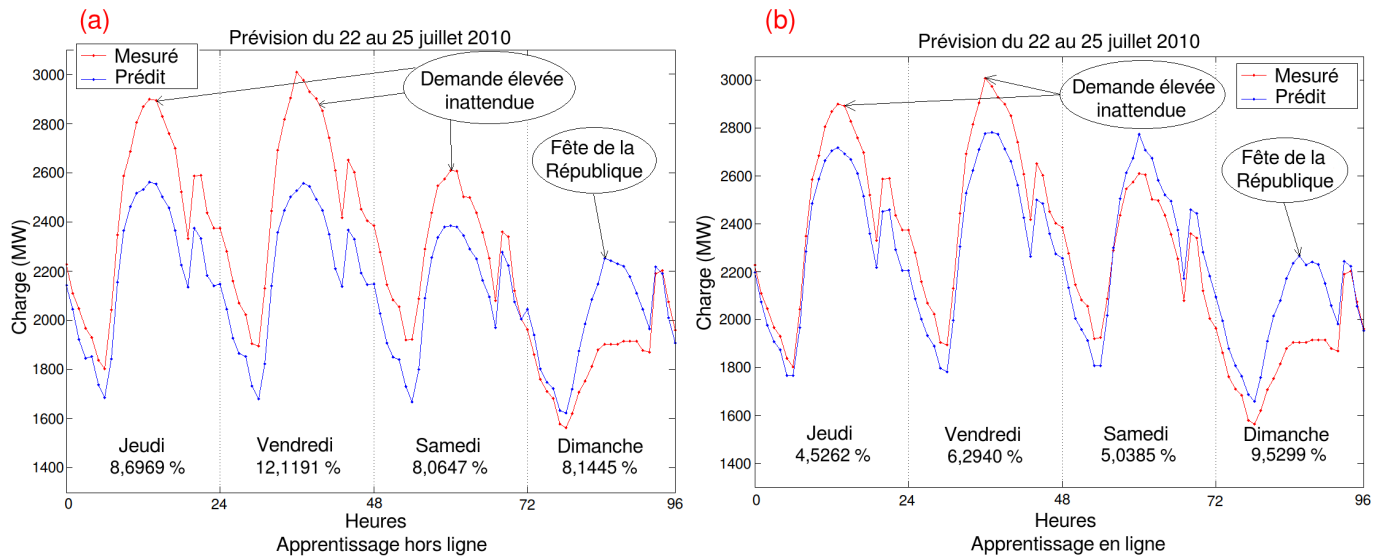


Figure 3.7 Illustration des différences entre les types d'apprentissage, avec la MAPE quotidienne

Période d'apprentissage	Période de test	MAPE mensuelle (%)	Temps de calcul (s)	MAPE mensuelle (%)	Temps de calcul (s)
		Apprentissage hors ligne		Apprentissage en ligne	
Janvier 2009 – Décembre 2009	Janvier 2010	2,5656	1,28	2,2524	40,4
Janvier 2009 – Mars 2010	Avril 2010	2,1903	1,7	2,1569	52,03
Janvier 2009 – Juin 2010	Juillet 2010	4,0202	2,16	3,1416	67,7
Janvier 2009 – Septembre 2010	Octobre 2010	2,2336	2,68	2,2092	83,25

Tableau 3.2 Quantification de la performance des types d'apprentissage

endogènes du tableau 3.1 selon le cas, à travers un ensemble de règles linguistiques si-alors. Par exemple, pour les lundis, les entrées 7 et 8 sont prises du vendredi et du jeudi précédents, que ce soit pour l'apprentissage ou le test. De même, les entrées 5 et 6 sont prises du vendredi précédent. Cela revient à lier les lundis avec les jours de la semaine les plus proches ayant des profils de charge similaires. Sans cet ajustement, la demande du lundi resterait toujours sous-estimée, puisqu'elle serait affectée par la faible charge du dimanche et du samedi. Le même ajustement est effectué pour le

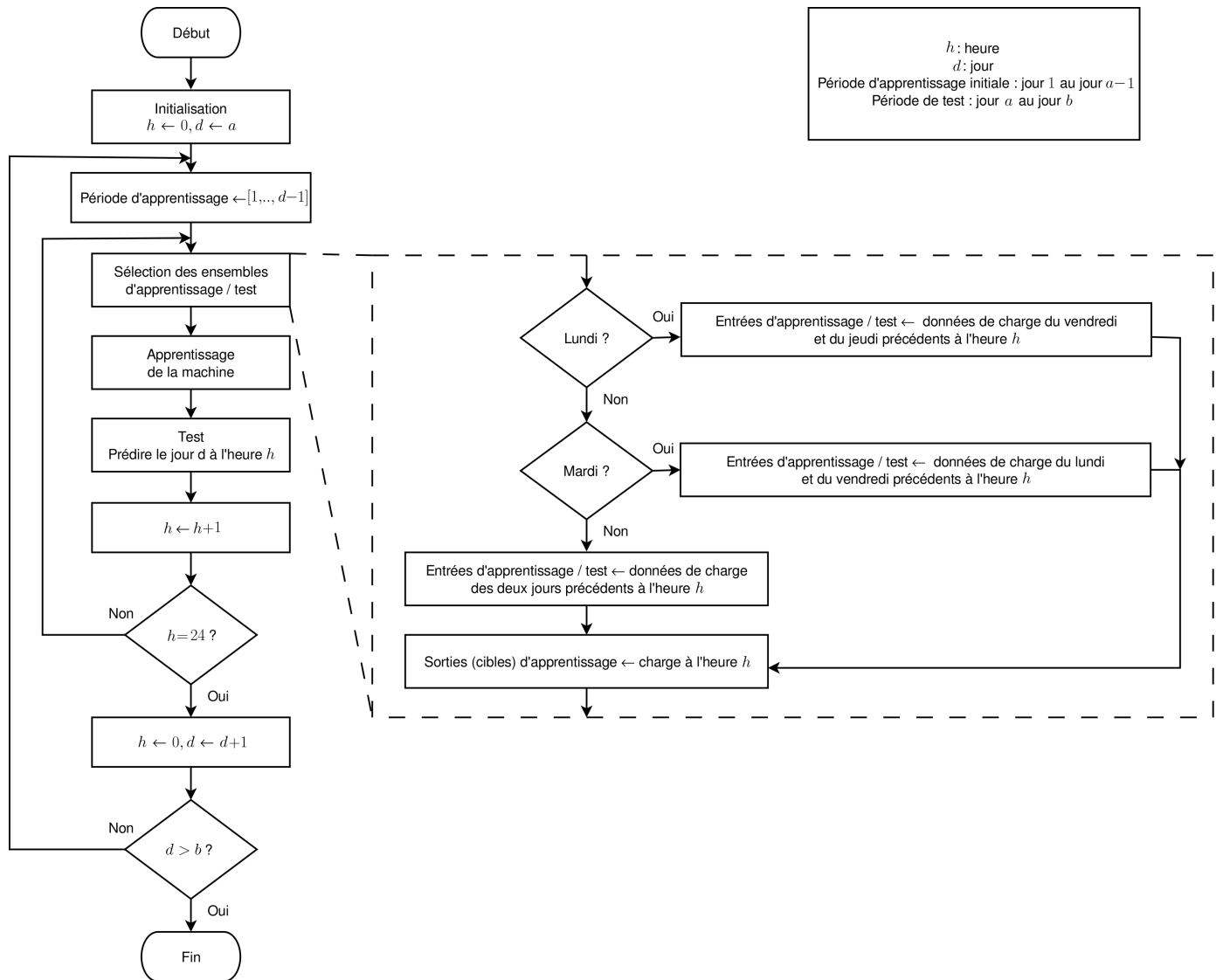


Figure 3.8 Organigramme décrivant la stratégie de prédiction

mardi, comme l'indique le zoom de la figure 3.8. Cependant, lier de la même manière les samedis et les dimanches avec des jours similaires, par exemple les samedis et les dimanches précédents, n'aboutit à aucune amélioration. Ceci s'explique par le fait que la transition de charge du vendredi au dimanche est lisse et lente, contrairement à la transition dimanche vers lundi, qui est brusque et aiguë. Ces transitions sont très claires sur la figure 3.9(d). De plus, l'énorme saut dans les indices du jour (entrée 2) du dimanche (7000) au lundi (1000) justifie le besoin de corriger les jours en aval du saut, donc les lundis et les mardis. Cet ajustement, qu'on appelle désormais sélection experte, est semblable à la recherche des jours similaires ou la sélection des caractéristiques élaborées par plusieurs chercheurs, mais il n'est pas automatisé et ne cherche pas à minimiser une certaine

distance. La sélection experte proposée se base sur des règles linguistiques si-alors, ce qui la rapproche des systèmes à inférence floue. Son avantage principal est qu'elle permet à l'utilisateur du modèle d'ajouter ses propres règles, relatives à son pays ou n'importe quelles autres règles qu'il juge être appropriées. Il est possible par exemple d'ajouter cette règle pour les régions chaudes : si la température est supérieure à 35°, alors les entrées 7 et 8 sont prises d'une journée historique très chaude. La sélection experte permet plus de flexibilité et améliore l'aptitude de généralisation du prédicteur proposé.

3.3.4 Données disponibles et prétraitement

Comme celles du chapitre précédent, les données disponibles dans ce chapitre sont fournies par la Société Tunisienne d'Électricité et de Gaz par l'intermédiaire du Centre National des Mouvements d'Énergie [117]. La durée considérée dans cette étude couvre le 1^{er} janvier 2009 jusqu'au 31^e août 2014, par des valeurs semi-horaires (échantillonnées à 30 minutes). Autrement dit, à chaque demi-heure on dispose de la puissance électrique instantanée demandée par la Tunisie entière (et non pas par zones). Les données incluent aussi le type de chaque jour (avec la notation du tableau 3.1) et les températures quotidiennes minimales et maximales mesurées à Tunis. Tous les chiffres sont ronds, sans virgules.

Le prétraitement des données inclut l'échantillonnage horaire (et non pas le moyennage), le remplacement des valeurs manquantes ou incorrectes par leurs antécédentes et l'élimination du redondant. Il n'y a pas de grosses quantités de données manquantes excepté l'année 2012, qui est généralement évitée. Aucune normalisation n'a été appliquée, donc les données sont restées au format brut et dans leurs gammes réelles. Les journées spéciales telles que les jours fériés sont intentionnellement gardés de sorte que la machine puisse apprendre leur comportement. Quelques contingences telles que les blackouts et les troubles politiques sont également gardées pour le test, afin de déterminer quand l'amélioration de la prévision est possible et quand elle ne l'est pas.

3.4 Résultats de la prévision

3.4.1 Résultats directs

La configuration de la machine RF est comme suit : le nombre d'arbres *ntree* est 500 (valeur par défaut) et le nombre de variables *mtry* à découper à chaque nœud est 4 (par défaut 2). Généralement, ces paramètres n'ont pas d'influence remarquable sur la performance de la prévision, et c'est ce qui constitue le point fort des forêts aléatoires. Il est cependant recommandé que *mtry* suive l'équation (1.38), et que *ntree* soit au moins égal à 500 pour avoir des résultats satisfaisants.

Afin de prédire la semaine du 21^e au 27^e octobre 2013, une période d'apprentissage initiale a été fixée du 1^{er} janvier au 20^e octobre 2013, et commence ensuite son expansion selon la stratégie expliquée dans la section 3.3.3. Entre autre, ça veut dire que $\beta = 7$ et $n = 291$ (nombre de jours dans la période d'apprentissage excepté les deux premiers jours). À la dernière prédiction, S_n contient 297 observations. Pour prédire 20 à 26 janvier 2014, la période d'apprentissage est prolongée au 19^e janvier et ainsi de suite. Donc l'ensemble d'apprentissage initial s'étend toujours du 1^{er} janvier 2013 jusqu'au dernier jour avant le début de la période de test, qui est une semaine de chaque saison. Ensuite, suivant le processus d'apprentissage en ligne, l'ensemble d'apprentissage s'agrandit à chaque jour de la période de test afin de prendre en compte la vraie charge du jour précédent. Il est important d'indiquer que le prédicteur ne voit pas les données de test durant son apprentissage même si ce dernier est en ligne. Autrement dit, on ne prédit pas des données qu'on a vues durant l'apprentissage, mais des données jamais vues auparavant. Les ensembles d'apprentissage et de test contiennent tous les deux les 8 entrées du tableau 3.1. Il n'y a pas de règle pour choisir la longueur de la période d'entraînement, quoiqu'elle doive être suffisamment longue pour couvrir toutes les saisons. De très longues périodes peuvent également dégrader la performance de la prévision, à part le temps de calcul excessif. En effet, les échantillons très anciens du signal de charge sont assez différents des échantillons actuels, à cause de l'évolution à long terme non considérée par le modèle.

Les résultats de simulation (figure 3.9) montrent une prédiction très précise du lundi au dimanche pour les quatre saisons. Néanmoins, la prévision de l'été trouve quelques difficultés par rapport à celles des autres saisons. Ceci revient au fait que le profil de charge de l'été soit particulier à cause de la climatisation (figure 3.2(a), expliquée en détails dans la section 3.3.1). Par suite, la machine a du mal à apprendre le profil de l'été. Afin d'évaluer davantage la qualité de la prévision, un test d'une demi-année est effectué, du 1^{er} janvier au 31^e juin 2014. L'erreur MAPE obtenue (dont le choix est justifié au paragraphe suivant) est de 2,24 %, excepté les jours fériés qui sont étudiés à part. Une prévision spéciale est effectuée pour certains jours fériés des 4 saisons, dont les résultats apparaissent sur la figure 3.10. Dans de tels cas, la prévision est moins performante, surtout lorsqu'il y a une succession de jours fériés comme le cas du 13 et 14 janvier 2014. Les jours fériés religieux comme El-Moulid (en 2014, c'était le 13 janvier) sont mobiles tout au long de l'année, suivant le calendrier lunaire (hégirien), et c'est ce qui rend la prédiction plus difficile.

Afin de quantifier les résultats numériquement, l'erreur MAPE est choisie comme critère d'évaluation. Les autres critères ne sont pas calculés vu qu'ils dépendent énormément de la quantité de puissance installée. Ce n'était pas le cas de la vitesse du vent par exemple, où une erreur de 1 m/s a toujours le même sens, puisque les valeurs qui maximisent la production sont toujours de 15 à 25 m/s. Par contre, une erreur de 100 MW n'a pas la même signification pour une installation de 4000 MW que pour une autre de 100 000 MW. De plus, le critère MAPE ne diverge pas ici, vu que

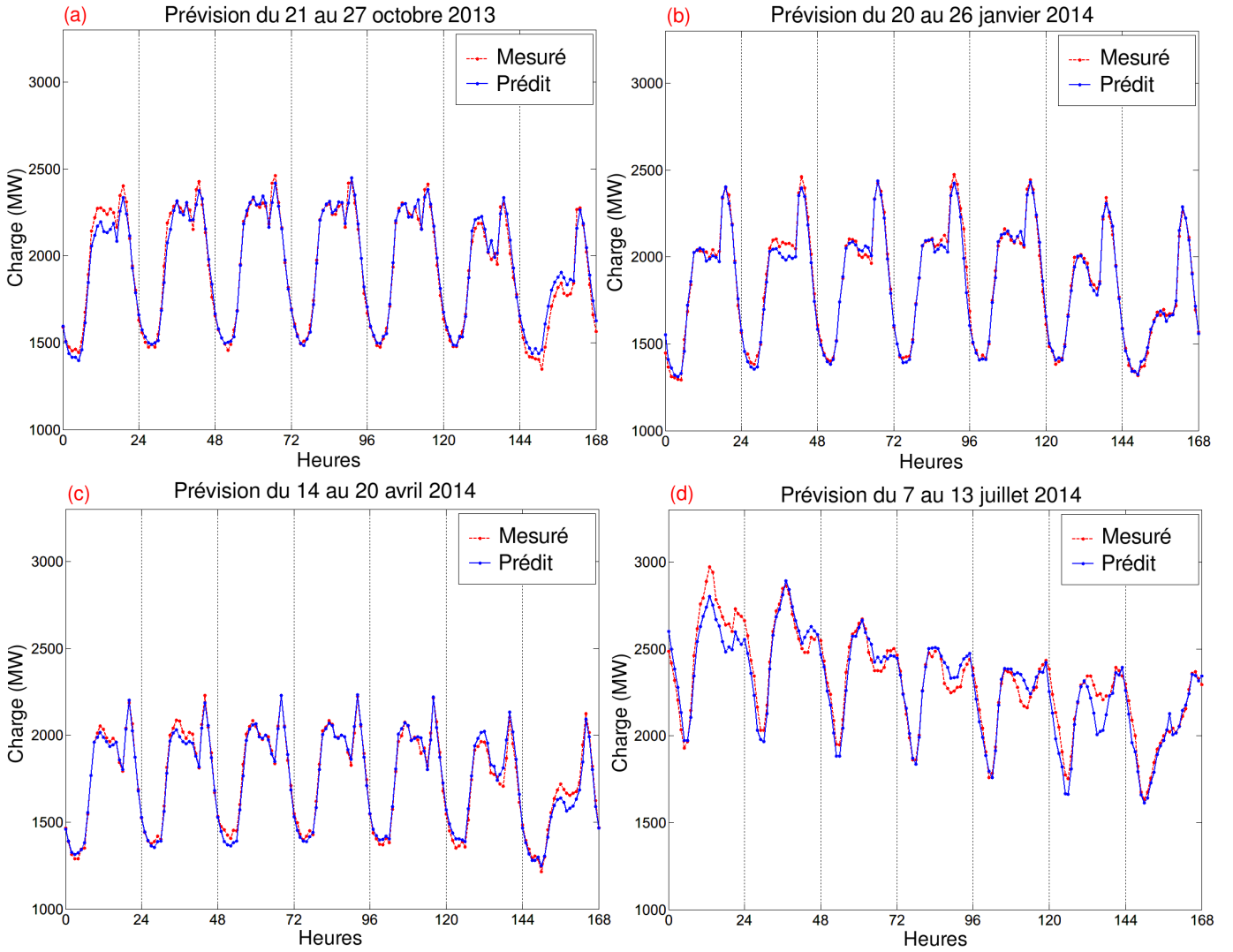


Figure 3.9 Résultats de prévision des jours réguliers par la machine RF

la valeur de charge ne s'approche jamais de zéro (sauf pour un cas très particulier qui sera étudié en détails à la section 3.4.2). Il assure donc une quantification précise en tant qu'erreur relative (pourcentage) et dispense la prévision des autres critères. On rappelle que MAPE s'écrit sous la forme :

$$MAPE = \frac{1}{k} \sum_{t=1}^k \frac{|\hat{L}(t) - L(t)|}{L(t)} \quad (3.2)$$

$L(t)$ désigne la valeur mesurée de la charge à l'heure t , $\hat{L}(t)$ représente la valeur prédite et $k = \beta.s$ est le nombre total d'erreurs. Ce critère définit l'erreur de l'ensemble de test en totalité (dans ce

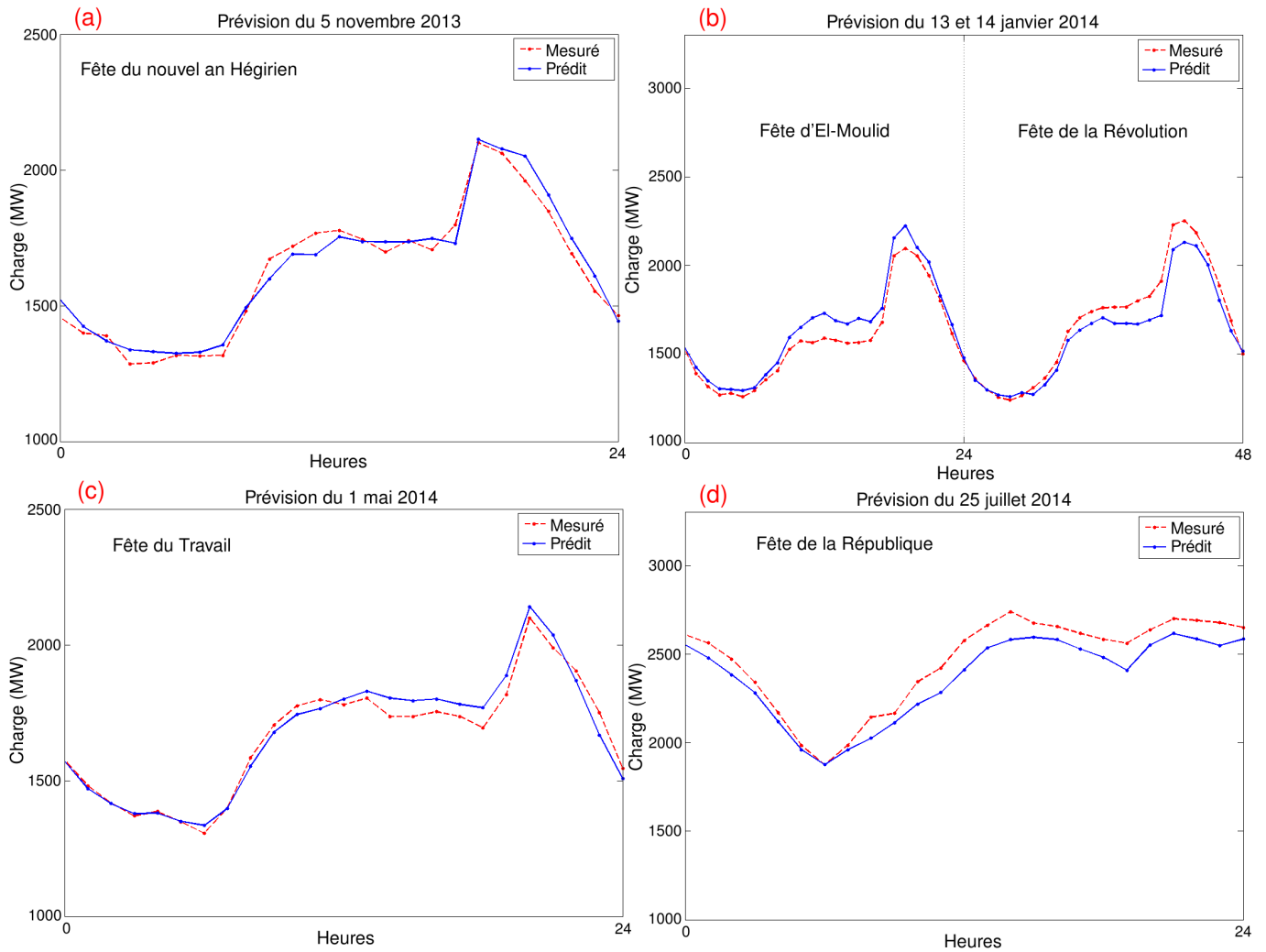


Figure 3.10 Résultats de prévision des jours fériés par la machine RF

cas, l'erreur hebdomadaire). La MAPE quotidienne est naturellement définie par :

$$MAPE = \frac{1}{s} \sum_{t=s.(i-1)+1}^{s.i} \frac{|\hat{L}(t) - L(t)|}{L(t)}, \quad i = 1, \dots, \beta \quad (3.3)$$

Les erreurs MAPE des jours prédits ainsi que leurs moyennes hebdomadaires sont affichées au tableau 3.3.

	Automne 2013 21 au 27 octobre	Hiver 2014 20 au 26 janvier	Printemps 2014 14 au 20 avril	Été 2014 7 au 13 juillet
Lundi	2,9573	1,7651	1,0019	4,2302
Mardi	2,0358	2,3948	1,5630	2,3781
Mercredi	1,0339	1,1314	1,7281	2,2637
Jeudi	1,0388	1,8062	1,1029	1,8342
Vendredi	1,1931	1,3294	1,3353	2,1097
Samedi	1,9190	1,2111	2,7520	4,2180
Dimanche	3,7984	1,2074	2,5791	1,7955
Moyenne	1,9966	1,5493	1,7232	2,6899
Jour férié	5 novembre 2,3962	13 et 14 janvier 4,2686 / 3,7361	1 mai 1,9130	25 juillet 3,6064

Tableau 3.3 Erreurs de prévision

3.4.2 Améliorations possibles

Même si les résultats semblent être précis, quelques événements inattendus peuvent se produire et induire un comportement inhabituel de la charge, comme ceux de la figure 3.11. Par exemple, la fête d'Aïd Al-Fitr (fête religieuse de deux jours) en 2014 était en été, et les gens (en tant que consommateurs d'énergie électrique) ne se comporteraient pas de la même manière si cette fête était en hiver, sans oublier l'augmentation spectaculaire de la température au deuxième jour, qui a entraîné une climatisation excessive. Autre complication, le jour juste avant la fête était un dimanche, et sa propre demande était plus faible que celle attendue, à cause des préparations des gens à fêter ; déplacements, achats de vêtements et de cadeaux, etc. Ces préparations, qui ne demandent pas d'énergie électrique, ne seraient pas si intenses s'il n'était pas dimanche. Au 31^e août 2014, un blackout complet a envahi tout le pays à cause d'une panne empirée par l'effet papillon, ce qui est un événement très rare. Dans des cas comme celui-ci, l'erreur MAPE s'approche de l'infini. En janvier 2011, la révolution tunisienne survenait, et des troubles politiques énormes ont conduit à une demande beaucoup plus faible qu'on prévoyait. Le problème majeur rencontré dans tous ces cas est la rareté, et la machine de prévision est incapable d'apprendre les cas rares. Un autre exemple de cas rares est celui donné par l'histogramme de la figure 3.12. La distribution des températures maximales quotidiennes montre que les valeurs supérieures à 41 °C ou inférieures à 10 °C sont très rares. Donc lorsque la température atteint ces valeurs extrêmes, le prédicteur ne sait plus comment réagir.

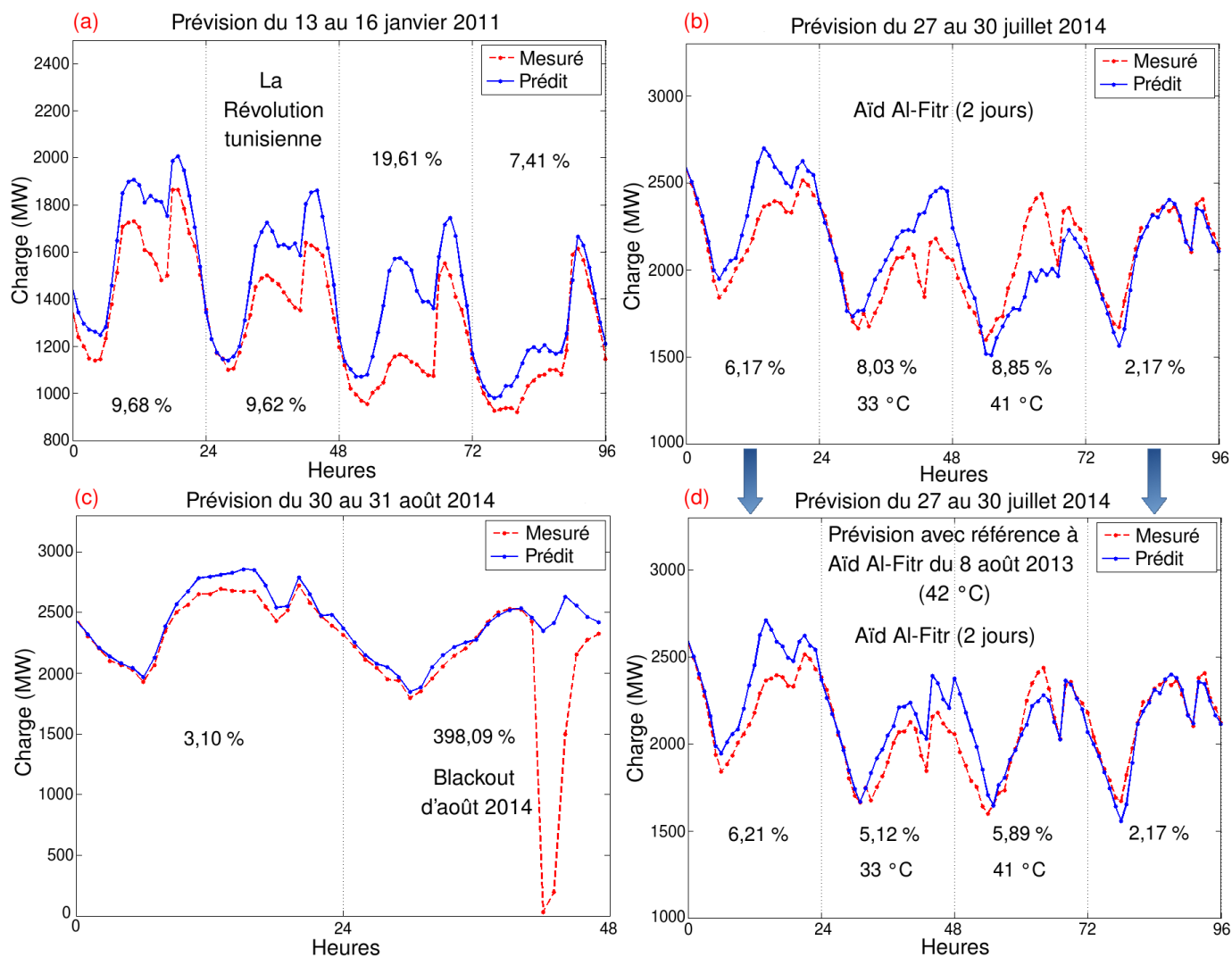


Figure 3.11 Mauvaises prévisions et améliorations possibles, avec la MAPE quotidienne

Une solution possible consiste à modifier la sélection des ensembles d'apprentissage et de test dans l'organigramme en y ajoutant d'autres règles. Dans le cas d'Aïd Al-Fitr de 2014, cette règle est ajoutée : si c'est Aïd Al-Fitr et la température est supérieure à 35 °C, alors l'entrée 7 est prise d'Aïd Al-Fitr de l'année précédente, c'est-à-dire le 8 août 2013. Les conditions sont similaires ; même jour férié et c'était une journée très chaude (42 °C). La figure 3.13, extension du zoom de la figure 3.8, montre la nouvelle règle ajoutée. Cette modification conduit à l'amélioration visible sur la figure 3.11(d), où l'erreur MAPE a passé de 8 % à 5 %. Cependant, la prévision du dimanche (juste avant la fête) reste la même si on la relie à l'année précédente, puisque le jour avant la fête de 2013 n'était pas dimanche. En général, il est possible d'ajouter ces règles génériques : s'il fait très chaud, se référer à une journée historique très chaude (sur l'entrée 7), et si c'est un jour férié, se référer

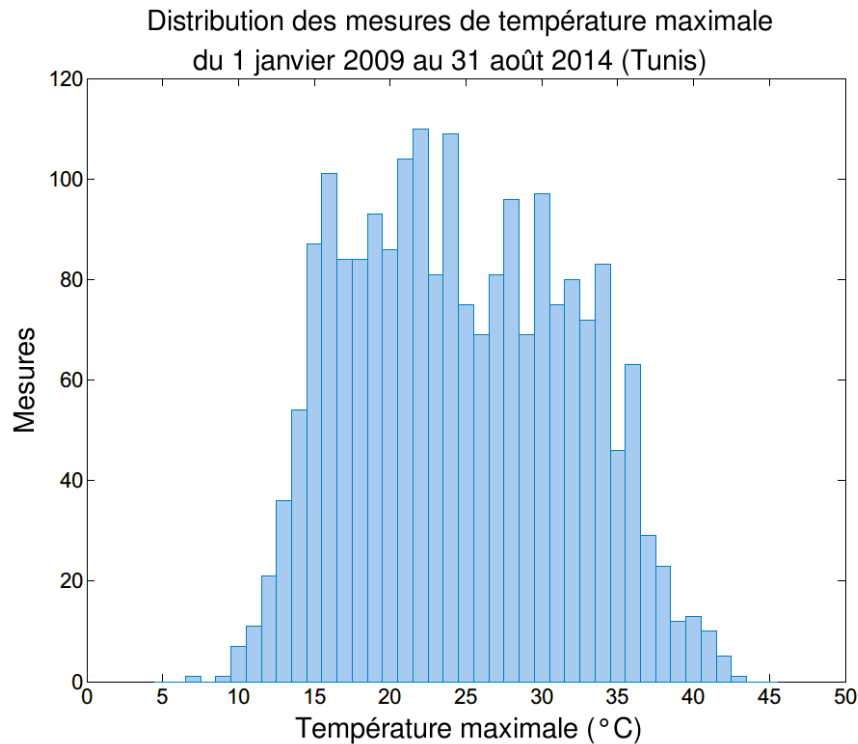


Figure 3.12 Distribution des températures quotidiennes maximales

au même jour férié de l'année précédente (sur l'entrée 7 également). Contrairement au cas d'Aïd Al-Fitr et aux autres fêtes annuelles dont la prévision peut être améliorée, les cas de blackout et des troubles politiques ne peuvent pas être corrigés, puisqu'ils se produisent une seule fois pendant des dizaines d'années, et même lorsqu'ils arrivent, ils ne suivent aucun motif particulier qui puisse être appris par la machine.

3.4.3 Comparaisons

Le but de cette section est de comparer le modèle RF proposé aux machines d'apprentissage concurrentes, à savoir ANN et SVM, et d'évaluer en même temps la capacité de généralisation à travers des tests sur d'autres marchés. Le modèle de persistance (PER), dans lequel la prédiction est littéralement égale à la mesure précédente, est utilisé comme référence. Afin d'assurer une comparaison équitable, aucun processus d'optimisation n'a été appliqué pour régler les paramètres des trois machines. Mais puisque ANN et SVM sont très sensibles à leurs paramètres, leurs configurations ont été dérivées des configurations par défaut recommandées par les constructeurs de leurs codes, avec de légères modifications manuelles pour avoir le meilleur rendement possible. Le but final est de pouvoir comparer ces trois méthodes en termes de performance sans recourir aux algorithmes d'optimisation. La stratégie de la figure 3.8 est appliquée pour chacune d'elles ; où elles doivent prédire

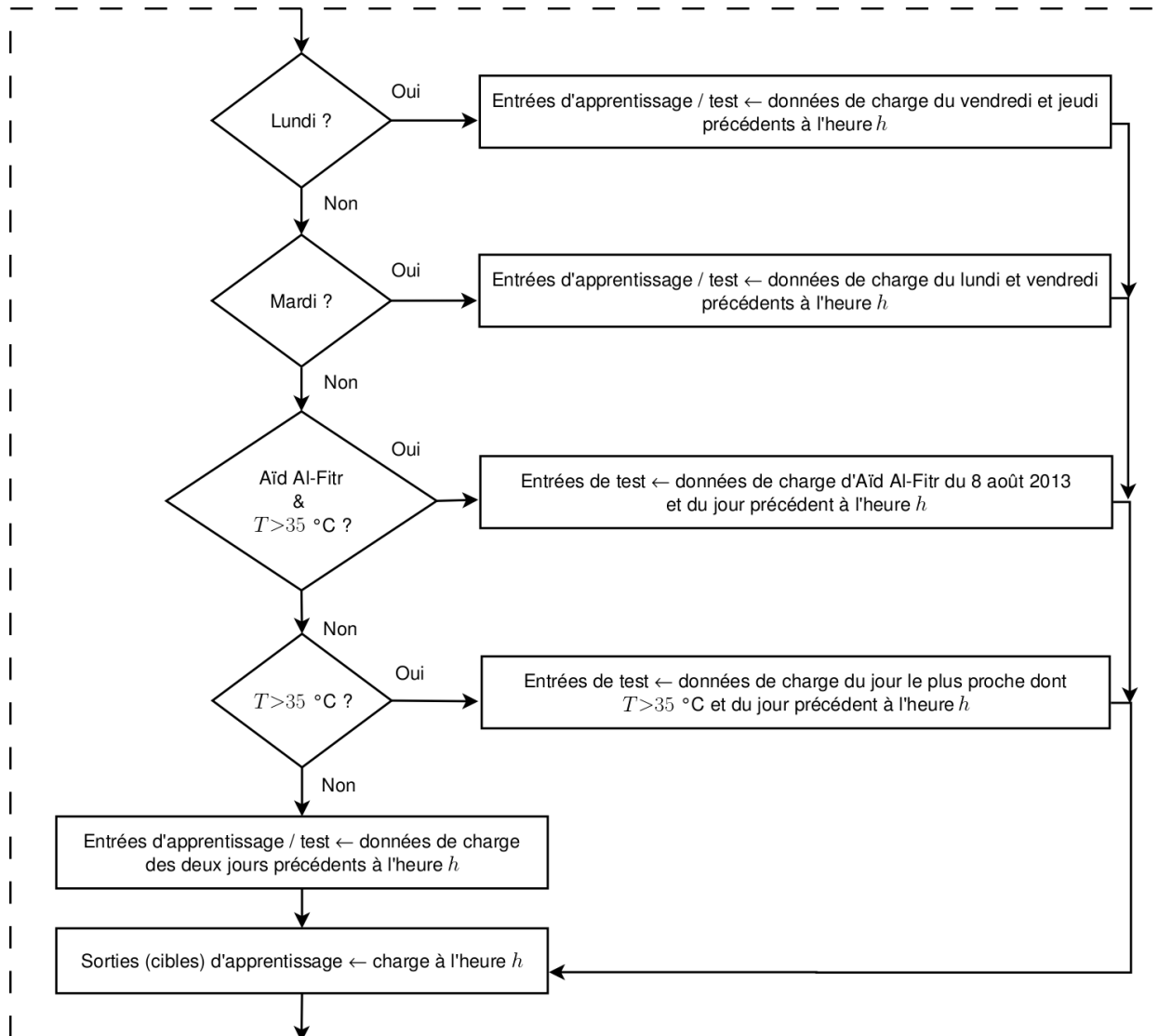


Figure 3.13 Sélection raffinée des ensembles d'apprentissage / test

la charge à l'heure t , les 8 entrées du tableau 3.1 étant données. Les configurations proposées sont comme suit :

- ANN : réseau 3 couches de type feedforward ; couche d'entrée à 8 neurones (pour les 8 entrées), couche cachée à 20 neurones et une couche de sortie à 1 neurone. L'algorithme d'apprentissage est la rétropropagation Levenberg-Marquardt, avec l'erreur quadratique moyenne comme fonction de performance. La fonction d'activation de chaque neurone est la sigmoïde tangente hyperbolique donnée par l'équation (3.4) (revoir aussi la section 1.6.3).

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (3.4)$$

– SVM : c’est en réalité SVR (régression à vecteurs de support), type epsilon. Le noyau est polynomial de degré 1 (équation (3.5)). Le paramètre γ est égal à 0,0003, r égal à 0 et le paramètre de régularisation C est fixé à 10 en utilisant la recherche par grille [127]. Le critère de terminaison est 0,001 (revoir la section 1.6.4).

$$K(u, v) = \gamma u^T v + r \quad (3.5)$$

– RF : la même configuration, $ntree = 500$ et $mtry = 4$.

Les trois machines sont testées en utilisant les données du réseau tunisien et du marché PJM américain (Pennsylvania-New Jersey-Maryland Interconnection) en vue de comparaison. Les résultats de prédiction de 4 semaines sont affichés au tableau 3.4. Les périodes de test sont choisies de sorte que les jours fériés soient évités. Les résultats pour le réseau tunisien, dans la plupart des cas, sont plus précis en utilisant le prédicteur RF qu’en utilisant les autres machines, comme le montre les barres de la figure 3.14. Bien sûr, la performance des machines ANN et SVM peut s’améliorer par optimisation, mais ceci conduit à d’autres problèmes de choix de l’algorithme d’optimisation et de ses paramètres.

Les résultats relatifs au marché PJM sont moins précis à cause de trois raisons principales. Premièrement, le marché PJM est dérégulé et contient une volatilité très élevée. Deuxièmement, il n’y a pas de raffinement appliqué aux jours particuliers, donc les règles demeurent génériques (uniquement la correction des lundis et mardis). Finalement, la température est mesurée uniquement à New York, qui est supposée la ville la plus gourmande en énergie, bien que le marché couvre une surface assez importante. Pour améliorer les résultats, l’utilisateur doit ajuster le modèle selon les caractéristiques du marché PJM ; en prenant les informations de température de plusieurs régions de la zone de couverture, et en ajoutant ses propres règles afin de prendre en compte les jours particuliers. En effet, le modèle proposé montre son efficacité lorsqu’il est raffiné selon la culture et le comportement des consommateurs.

En général, le critère MAPE d’une journée à l’avance est entre 1 % et 3 % dans la plupart des articles de recherche récents [76, 91, 92, 133–136]. Les résultats trouvés dans cette étude sont similaires, mais en utilisant une machine de prévision RF dont les paramètres sont génériques, sans aucun algorithme d’optimisation. La méthodologie proposée est à son tour adaptée à la prédiction RF, puisque l’accent est mis sur l’organisation des ensembles d’apprentissage et de test avec la sélection experte, plutôt que sur le réglage des paramètres de la machine. L’atout principal de la forêt aléatoire est donc le peu de paramètres à régler. En général, garder les valeurs par défaut des paramètres $ntree$

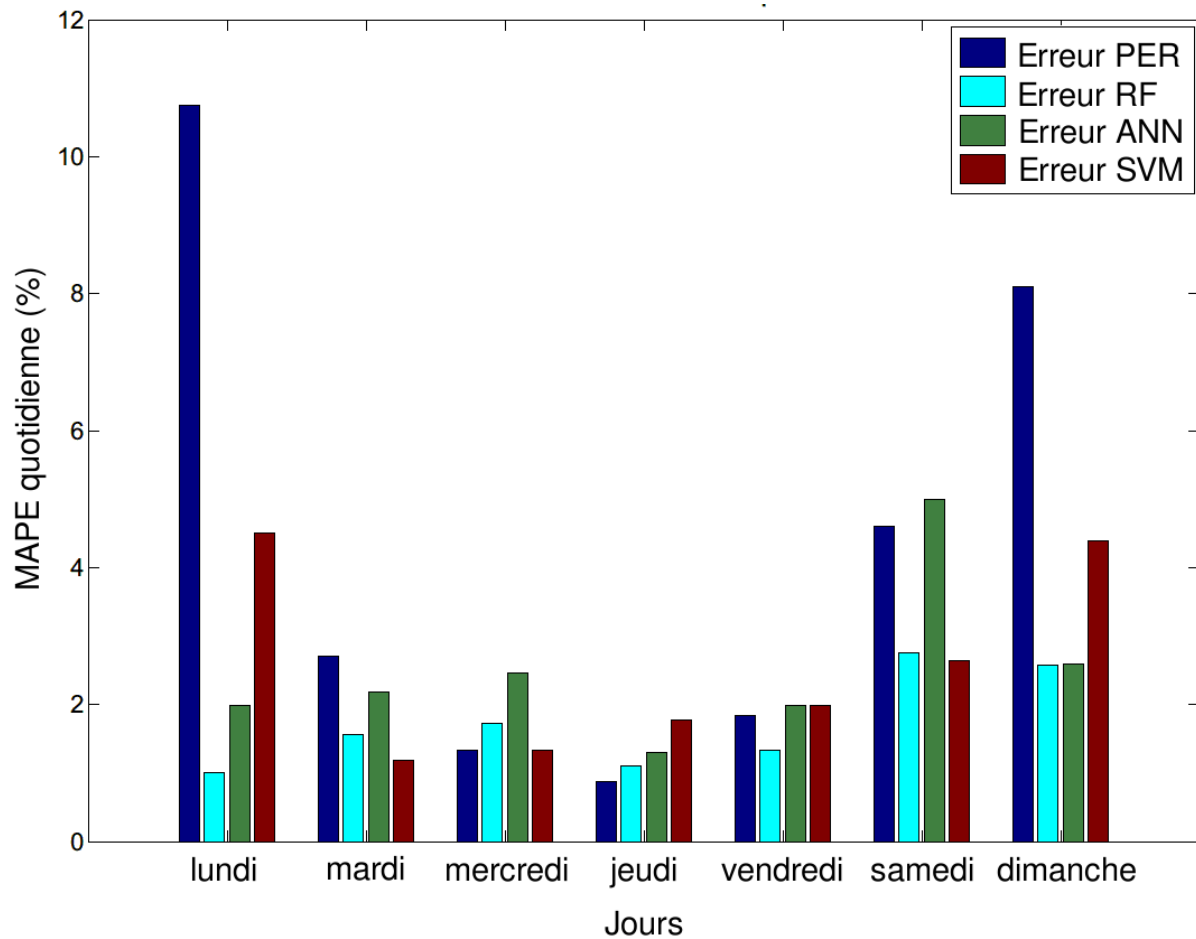


Figure 3.14 Illustration des différences entre les méthodes proposées en termes d'erreur de prévision, du 20 au 26 janvier 2014

et *mtry* suffit pour obtenir le meilleur possible de cette méthode. Par contre, la performance des machines ANN et SVM dépend énormément de leurs paramètres. De plus, la validation croisée interne permet d'obtenir une bonne capacité de généralisation. Comparée à ANFIS et les systèmes experts, la machine RF possède l'avantage de pouvoir manipuler des ensembles d'apprentissage énormes, ce qui est compliqué et fastidieux en utilisant les règles floues. La sélection experte des entrées est la contribution majeure qui distingue le modèle proposé des autres prédictors RF, où les ensembles d'apprentissage et de test raffinés permettent une prévision plus flexible. D'autres nouveautés sont également ajoutées et contribuent à cette distinction, telles que le choix des entrées par l'analyse de la fonction d'autocorrélation et des mesures de corrélation croisée, la mesure de l'importance des variables et le processus d'apprentissage en ligne.

	Réseau tunisien				Marché PJM			
	PER	RF	ANN	SVM	PER	RF	ANN	SVM
Automne	21 au 27 octobre 2013				21 au 27 octobre 2013			
Lundi	10,2473	2,9573	5,9970	1,1842	11,6758	2,8623	3,9795	1,8143
Mardi	1,4937	2,0358	3,4987	1,0261	1,2088	0,9653	2,9234	2,7147
Mercredi	1,1342	1,0339	2,2240	1,0531	3,8949	4,2642	5,0863	2,4308
Jeudi	1,3563	1,0388	2,5128	0,9776	2,4319	2,9714	3,5950	3,0299
Vendredi	1,0737	1,1931	1,8585	1,7933	2,0794	1,9389	4,6577	2,6091
Samedi	4,5218	1,9190	3,5893	2,6264	7,7393	2,4416	4,4413	2,6219
Dimanche	10,8898	3,7984	3,8425	6,8002	5,5892	3,0055	5,5879	2,1995
Moyenne	4,3881	1,9966	3,3604	2,2087	4,9456	2,6356	4,3244	2,4886
Hiver	20 au 26 janvier 2014				13 au 19 janvier 2014			
Lundi	12,7193	1,7651	1,8818	5,1333	8,9871	1,1432	2,2401	3,6559
Mardi	3,6586	2,3948	2,4101	1,9979	2,5049	2,6942	5,2133	1,8779
Mercredi	1,7609	1,1314	1,4194	1,5722	5,3825	5,1880	3,0915	5,7661
Jeudi	2,1357	1,8062	1,8745	3,0358	2,9150	5,8280	5,3839	4,3063
Vendredi	2,0410	1,3294	1,9987	2,8130	1,4735	3,4029	4,6590	4,8541
Samedi	5,4608	1,2111	1,8029	2,7589	2,8463	3,4065	5,0160	3,7134
Dimanche	9,9884	1,2074	2,4129	5,2611	5,8810	1,5673	2,4824	2,0096
Moyenne	5,3949	1,5493	1,9715	3,2246	4,2843	3,3186	4,0123	3,7405
Printemps	14 au 20 avril 2014				7 au 13 avril 2014			
Lundi	10,7557	1,0019	1,9814	4,5069	10,5069	1,9821	2,8370	1,4882
Mardi	2,7122	1,5630	2,1754	1,1834	2,6453	2,5960	4,3617	3,0996
Mercredi	1,3405	1,7281	2,4537	1,3258	2,1431	3,2794	3,7305	2,7160
Jeudi	0,8667	1,1029	1,3012	1,7798	1,9861	2,8144	2,3001	1,6258
Vendredi	1,8373	1,3353	1,9889	1,9818	2,8596	2,1364	5,1039	3,1789
Samedi	4,6041	2,7520	4,9910	2,6412	10,6370	1,4290	2,3294	6,2736
Dimanche	8,0959	2,5791	2,5934	4,3904	4,1369	4,3520	4,6895	3,9764
Moyenne	4,3161	1,7232	2,4979	2,5442	4,9879	2,6556	3,6217	3,1941
Été	7 au 13 juillet 2014				21 au 27 juillet 2014			
Lundi	13,6542	4,2302	3,3521	2,2994	15,5319	6,7983	4,1693	5,1931
Mardi	4,1347	2,3781	2,4733	3,6409	7,3031	5,8869	3,8159	4,5301
Mercredi	5,1276	2,2637	2,6060	2,8871	3,3315	1,7904	1,5886	2,1203
Jeudi	4,9159	1,8342	3,0465	2,1154	12,2722	7,5953	5,4861	9,3693
Vendredi	3,5551	2,1097	2,5774	1,4218	4,8534	1,3925	4,4891	4,6822
Samedi	1,7505	4,2180	3,8909	3,9019	3,6804	2,6274	4,0350	4,0649
Dimanche	8,386	1,7955	2,4515	3,6495	1,8542	2,0841	2,3800	2,6695
Moyenne	5,9320	2,6899	2,9140	2,8452	6,9752	4,0250	3,7091	4,6614

Tableau 3.4 Comparaison entre les différentes méthodes en termes d'erreur de prévision

3.5 Conclusion

Ce chapitre propose un modèle de prédiction de la charge électrique demandée en Tunisie d'un jour à l'avance, par un pas temporel d'une heure. La forêt aléatoire, qui fait partie des machines d'apprentissage intelligentes, est choisie pour construire le modèle. Caractérisé par la validation croisée interne, la mesure d'importance intrinsèque et l'immunité aux changements des paramètres et aux entrées non significatives, un prédicteur RF est construit tout en ayant 24 versions représentant les 24 heures du jour, en vue d'éviter les modèles itératifs. Un procédé de sélection experte est ajouté afin de raffiner les ensembles d'apprentissage et de test. Un processus d'apprentissage en ligne pilote la prévision sur toute la période de test, dans le but de surmonter les changements brusques de la charge.

Après l'analyse des conditions du marché et du profil de la demande, des tests ont été effectués sur des jours de travail réguliers, des weekends et des jours fériés choisis parmi les 4 saisons. L'erreur de prévision obtenue, excédant rarement une moyenne de 2,3 % en termes de MAPE, reflète l'exactitude de l'approche. La procédure de sélection des ensembles d'apprentissage et de test à travers des règles linguistiques si-alors prouve aussi son efficacité et son aptitude à s'ajuster selon la culture du pays et les spécificités du marché. La machine RF couplée à la sélection experte est capable de capturer le comportement complexe de la demande et de résoudre quelques cas spéciaux qui sont liés à la culture, les canicules extrêmes, les événements religieux et les jours fériés mobiles grâce à un choix approprié des entrées. Bien que la prévision de la demande soit bien précise, elle n'est pas suffisante dans un marché dérégulé. Dans un tel contexte, il faut absolument prédire en parallèle les prix de l'électricité et bien choisir la méthode de prédiction qui n'est pas nécessairement identique à celle de la demande, et c'est ce qui constitue l'objectif du prochain chapitre.

CHAPITRE 4 PERSPECTIVES POUR LE RÉSEAU TUNISIEN : TARIFICATION ET PRÉVISION DES PRIX DE L'ÉLECTRICITÉ

4.1 Introduction

La prévision de la charge, qui était le sujet du chapitre précédent, n'est pleinement exploitée que si on l'associe à la prévision des prix, car le profil du prix a son impact sur le comportement des consommateurs. Cependant, la prévision des prix en particulier n'apparaît que dans un contexte de dérégulation et de concurrence libre entre les différentes entreprises génératrices d'électricité, et en exploitant les compteurs intelligents. Ce chapitre propose donc des tâches de prévision et des mesures d'efficacité énergétique qui ne sont encore pas possibles sur le réseau tunisien en cette deuxième décennie du 21^e siècle. Cette solution technologique est étudiée en détails dans ce chapitre et proposée comme perspective pour le futur proche. Mais avant d'attaquer la prévision des prix qui est surtout utile du côté client, il est intéressant de proposer une nouvelle procédure de tarification, qui peut se voir comme une procédure de prévision des prix côté fournisseur. La tarification souple permet de prendre en compte les contraintes de production et sert à guider les consommateurs et à les éloigner des périodes de pointes.

Le présent chapitre se présente alors comme suit : quelques éléments de littérature sont d'abord présentés afin de clarifier la différence entre la tarification et la prévision des prix, et analyser leurs différentes approches. Par la suite, deux approches seront proposées pour la tarification de l'électricité dans le marché tunisien. La première consiste à estimer les coûts de production de chaque type de générateurs, et d'en déduire un profil de prix qui encourage ou décourage les clients à consommer selon les conditions. La deuxième qui est l'inverse de la première, consiste à analyser la tarification actuelle de l'électricité en Tunisie afin de pouvoir extraire un profil de prix qui reflète les coûts de production. Et dans les deux cas, le prix serait corrélé avec la demande.

La tarification sera ensuite étudiée dans un marché dérégulé, on prendra un cas réel qui correspond à celui de l'Australie. Le marché sera présenté, les profils de la charge et du prix seront analysés. En dernière partie, la prévision du prix sera introduite, avant de passer à la prévision d'une demi-heure et d'une journée à l'avance. Les machines d'apprentissage, ANN, SVM et RF seront choisies comme prédicteurs, et leurs performances seront évaluées en termes d'erreur de prédiction. Cette étude serait utile et applicable lorsque le marché tunisien d'électricité devient dérégulé. La question à laquelle ce chapitre essaye de répondre est la suivante : quelle méthode de tarification serait la plus adaptée au contexte tunisien, et qui contribuerait à l'équilibre de l'offre et la demande ? Et quel modèle de prévision des prix permet d'avoir les meilleurs résultats, que ce soit pour une demi-heure

ou une journée à l'avance, dans un contexte de dérégulation ?

4.2 Tarification et prévision des prix : différences et approches utilisées

La tarification de l'électricité dans n'importe quel marché, qu'il soit dérégulé ou non, dépend d'un certain nombre de paramètres, à savoir le coût de production, le coût de transport, l'entretien, les subventions, etc. Cependant, l'objectif de cette étude n'est pas d'estimer les différents coûts, mais de manipuler le prix afin de contrôler la consommation d'électricité pour éviter les pics et les pertes. Ceci permettra d'améliorer l'efficacité énergétique. La tarification sera alors étudiée dans ce sens. La prévision des prix est l'estimation future de cette tarification, et qui est nécessaire lorsque le pourcentage des ressources intermittentes dans le parc de production est élevé, ou lorsqu'il y a de la concurrence entre les générateurs d'électricité. Ces paramètres, l'intermittence des ressources et la concurrence, sont deux grandes sources d'incertitude dans la détermination des prix.

4.2.1 Tarification

On introduit la réponse à la demande (DR : demand response) comme un mécanisme qui incite les utilisateurs finaux (consommateurs) à modifier leurs usages de l'énergie par rapport à leurs habitudes de consommation normales selon les changements dans le prix de l'électricité au fil du temps. Les stratégies de la DR peuvent être regroupées en deux catégories générales : la DR incitative et la DR basée sur le temps. Dans la DR incitative, on donne aux consommateurs des incitations de paiement pour abaisser leur consommation selon la fiabilité du système. Et pour la DR basée sur les temps, les consommateurs vont ajuster dynamiquement leur consommation en fonction de la tarification variante au cours du temps tout en maximisant leurs profits. Il existe aussi des programmes DR qui combinent les deux ; l'incitation et le temps [137]. La DR est aussi appelée la gestion côté demande (DSM : demand side management), définie comme une approche proactive pour la planification, la mise en œuvre et le suivi de la consommation d'énergie pendant les heures de pointe. Elle repose principalement sur la gestion de la demande en contrôlant les consommations du côté de l'utilisateur. Les principaux objectifs de la DSM sont : l'écêtement des pics, le remplissage des creux, le déplacement de la charge, la croissance stratégique de la charge, la conservation stratégique, et l'assouplissement du profil de charge [138].

La DSM incitative

Le système de tarification basé sur l'incitation le plus couramment utilisé est le contrôle direct de la charge (DLC : direct load control). Dans le DLC, l'utilitaire prend le contrôle de l'équipement électrique du consommateur et a une autorité d'arrêter son matériel électrique. Les programmes

DSM basés sur l'incitation comme le DLC paient l'argent d'incitation aux consommateurs pour le temps où ils sont invités à réduire ou arrêter leur charge en période de pointe. La plupart du temps, les charges importantes qui contribuent aux pics les plus élevés sont entraînées de façon thermostatique, à savoir les climatiseurs, les chauffages et les chauffe-eaux. Dans le DLC, l'utilitaire installe dans le local du consommateur l'équipement qui peut contrôler ses appareils électriques [138].

La Société Tunisienne de l'Électricité et du Gaz (STEG) utilise un outil DLC d'optimisation de la demande d'électricité centralisée à distance appelé Télécommande Centralisée (TLC). Cet outil permet d'assurer l'équilibre suivant : $\text{production} = \text{consommation} + \text{pertes}$, en lissant la courbe de charge. Le lissage est obtenu en réduisant les différences entre les pics et les creux de la demande. La technique TLC est basée sur la transmission d'un signal à une fréquence audible (216,66 Hz) superposée à la fréquence du réseau de 50 Hz, pour commander des relais de puissance installés chez les clients. Ce dispositif permet de supprimer certaines charges de pointe, tels que le pompage agricole, les chauffe-eaux électriques, l'allumage et l'extinction de l'éclairage public, etc. La TLC est assurée par une baie de commande qui contrôle un convertisseur de fréquence triphasé. La baie de commande contient un programme géré par un moniteur en temps réel qui fait une scrutation périodique afin de sélectionner la tâche qui doit être exécutée en tenant compte de sa priorité et de sa périodicité. La TLC permet aussi de commuter les compteurs entre les tarifs, ce qui permet de remplacer partiellement les compteurs intelligents [139].

La DSM basée sur le temps

Cette technique de DSM est celle qui évoque la tarification intelligente. Le programme DSM basé sur le temps permet aux utilisateurs de choisir la période d'utilisation de l'électricité en tenant compte de ses prix. Les différentes approches de ce programme sont les suivantes :

- Temps d'utilisation (TOU : time of use)

Ce type de système de tarification intelligente offre des prix pendant les heures de pointe différents de ceux en dehors. Les prix de pointe sont naturellement plus élevés. Certains utilitaires divisent le système de tarification TOU en trois parties, à savoir les heures de pointe, le temps de la mi-pointe et le temps hors pointe. Aux heures de pointe, les prix sont maintenus élevés en raison de la demande élevée d'énergie qui déclenche l'utilisation de centrales de pointe. En effet, les centrales sont classées en base, intermédiaire, et pointe. Les centrales en charge de base ont généralement un coût de capital élevé, mais un faible coût d'exploitation, et donc fonctionnent tout le temps (par exemple, hydraulique, nucléaire, géothermique). Les centrales de charge intermédiaire (par exemple, charbon) ont un coût de fonctionnement plus élevé, et les centrales de charge de pointe (par exemple, gaz naturel, essence, diesel) ont le coût d'exploitation le plus élevé. Dans une période donnée, si la demande dépasse la capacité de charge de base, le générateur recourt aux centrales intermédiaires

de charge et puis enfin, aux centrales de pointe pour produire de l'électricité supplémentaire [137].

En outre, pour répondre à la demande de pointe, l'expansion des installations existantes comme l'installation de nouvelles centrales électriques, l'extension des réseaux de transport et de distribution sera nécessaire. Les consommateurs peuvent bénéficier de leur propre production distribuée comme un système solaire ou éolien installé dans leurs maisons. Ils peuvent vendre l'électricité au réseau à un prix élevé et utiliser l'électricité pendant les heures creuses à des tarifs moins chers avec leur propre génération. Le réseau tunisien STEG utilise ce genre de tarification actuellement juste pour l'irrigation.

– La tarification de pointe critique (CPP : critical peak pricing)

La CPP est une forme modifiée du TOU. Cela implique une certaine période de l'année durant laquelle la demande d'énergie est très élevée par rapport au reste des heures de pointe de l'année. Le TOU est incapable de distinguer le temps de l'année pendant lequel le coût de l'énergie produite dépasse la valeur de crête fixée dans le schéma de tarification horaire. Parfois, il coûte à l'utilitaire un prix de gros de l'énergie très élevé qui est hors de la portée du TOU. Pour garder un système de tarification équitable, pendant les jours normaux, le prix de pointe facturé au consommateur est le même que celui du TOU. Et lorsque les prévisions indiquent une charge très élevée, ce qui fait de la journée une journée critique, la CPP est déclarée afin de remédier à ce problème économique et écrêter les pointes. Habituellement, le jour du CPP est annoncé 24 heures à l'avance. La CPP peut être déclarée pour un certain nombre de jours pendant un an, et le prix peut être pendant le pic 15 fois plus élevé que celui des heures hors pointe.

– La tarification en temps réel (RTP : real time pricing)

La RTP est le système qui reflète le mieux les coûts supportés par l'utilitaire qui fournit l'électricité utilisée par le consommateur. Dans le régime RTP, les consommateurs sont facturés à un prix très proche du prix réel de génération (production) à l'intervalle de temps donné. Le régime RTP peut être de deux types, la tarification horaire et la tarification d'une journée à l'avance. Pour les prix horaires, le prix de l'électricité pour une heure est annoncé chaque heure pour l'heure qui suit. Alors que pour la RTP d'une journée à l'avance, le prix de l'électricité pour une livraison de 24 h est annoncé au préalable, selon la prévision de la demande et le coût de production. La RTP nécessite l'implication des consommateurs, qui profitent de l'électricité bon marché lorsqu'elle est produite à faible coût. La tarification d'une journée à l'avance peut être plus efficace puisque les consommateurs obtiennent suffisamment de temps pour planifier leur consommation d'électricité, alors que les prix horaires peuvent être fastidieux. La RTP peut être efficace seulement avec la participation active des consommateurs et lorsque le niveau technologique le permet. En outre, les systèmes de contrôle de charge peuvent être conçus pour faire bénéficier à la fois l'utilitaire (fournisseur) et le consommateur. Les signaux RTP combinés avec l'automatisation au niveau du client font bénéficier

non seulement le consommateur en réduisant la charge, mais il aide aussi l'utilitaire par l'écrêtage des pics et la réduction de la charge grâce à la DSM dans le cas de production ou de capacité de distribution limitées. Ainsi, un système RTP bien conçu augmente la fiabilité du système, réduit le coût de production et réduit la facture d'électricité des consommateurs [138]. Les approches de tarification proposées dans ce chapitre sont toutes les deux de type RTP.

4.2.2 Prévision des prix

On ne parle de prévision des prix que dans un contexte de tarification en temps réel (RTP), car pour les méthodes TOU et CPP, les prix sont connus à l'avance, ou à la limite déterminés à partir de la prévision de la charge. Dans un contexte de RTP, on dispose généralement d'un marché qui met en relations un / plusieurs générateurs avec un / plusieurs fournisseurs qui sont eux-mêmes en relations avec plusieurs consommateurs. La prévision des prix n'est pas utile uniquement pour les agents de consommation (clients), mais aussi pour les gestionnaires de réseaux (fournisseurs) qui ont besoin d'un préavis sur les prix afin de vérifier que leur planification est faisable et qu'elle respecte les contraintes de transport.

La prévision des prix de l'électricité est devenue une information nécessaire pour les mécanismes de prise de décision des entreprises du secteur d'énergie. Les prévisions de prix allant de quelques heures à quelques mois à venir ont un intérêt particulier pour les gestionnaires de puissance. Un générateur, une entreprise de service public ou un grand consommateur industriel est en mesure de prévoir les prix volatiles avec un niveau de précision raisonnable. Cette prévision lui permet d'ajuster sa stratégie d'appel d'offres et son propre calendrier de production ou de consommation, afin de réduire le risque et maximiser ses profits.

Dans un marché day-ahead (journée à l'avance), les agents (générateurs) soumettent leurs offres de livraison d'électricité pour chaque heure de la prochaine journée (ou une période plus courte) avant une certaine heure de clôture du marché. Le rôle du marché organisé est de faire correspondre l'offre à la demande d'électricité de manière à déterminer le prix d'équilibre du marché (MCP : Market Clearing Price). Typiquement, le MCP est établi dans une vente aux enchères menée une fois par jour, comme l'intersection entre la courbe d'offre et la courbe de demande. Bien évidemment, ce prix qui n'est pas connu à l'avance, dépend de la concurrence entre ces différents générateurs et de la nature des ressources de chacun, et c'est ce qui donne à la prévision des prix toute son utilité. Les offres avec des prix négatifs sont autorisées dans de nombreux marchés, généralement lorsque la demande est très faible (les pertes dues à l'arrêt d'une unité de production peuvent dépasser les pertes dues à accepter des prix négatifs) ou lorsque la production des sources renouvelables est très élevée. Certains marchés suivent une structure de règlement en temps réel unique. Dans un tel système, les prix sont fixés par l'opérateur du marché chaque 5 min, et les prix comptants sont alors

déterminés dans les demi-heures. Ce qui fait qu'ils sont beaucoup plus volatiles que la plupart des autres marchés.

Le prix comptant de l'électricité dépend d'un grand nombre de facteurs fondamentaux, y compris les charges du système (la demande, les chiffres de consommation), les variables météorologiques (températures, vitesse du vent, précipitations, rayonnement solaire), les coûts de carburant (pétrole et gaz naturel, et dans une moindre mesure charbon), la marge de réserve (surplus de génération : la production disponible moins la demande prévue), et la maintenance programmée ou arrêts forcés des composants importants du réseau électrique. Leurs dernières valeurs historiques et leurs estimations futures pour l'horizon de prévision considéré sont précieuses pour la construction et l'étalonnage correct des modèles.

Les différentes modélisations des prix proposées par la littérature sont les suivantes [140] :

- Les approches multi-agents : ce sont des modèles qui simulent le fonctionnement d'un système d'agents hétérogènes (unités génératrices, entreprises) qui interagissent les uns avec les autres, afin de construire le processus de prix en faisant correspondre l'offre et la demande dans le marché.

Ces modèles basés sur les coûts de production ont la capacité de prévoir les prix heure par heure. Ils sont appropriés pour les marchés réglementés avec peu d'incertitude sur les prix et ayant une structure stable, mais ne sont pas adaptés aux marchés concurrentiels de l'électricité.

- Les méthodes fondamentales (structurelles), qui décrivent la dynamique des prix en modélisant les impacts des facteurs physiques et économiques importants sur le prix de l'électricité.

- Les méthodes de forme réduite (quantitative, stochastique), modèles qui caractérisent les propriétés statistiques des prix de l'électricité au fil du temps, avec l'objectif ultime de la gestion des risques.

- Les méthodes statistiques (économétrie, analyse technique) : ils sont des applications directes des techniques statistiques de prévision de charge ou de marché d'énergie, donc des implémentations de modèles économétriques

- L'intelligence informatique (artificielle fondée sur le renseignement, statistiques non-linéaires) : ce sont des techniques qui combinent des éléments de l'apprentissage, l'évolution et la logique floue pour créer des approches adaptables aux systèmes dynamiques complexes.

Les deux dernières familles de méthodes (statistiques et intelligence artificielle) sont généralement utilisées par les chercheurs en génie électrique. Les autres sont élaborées par les chercheurs en économie, surtout celles basées sur l'estimation des coûts de production. Et quoi qu'il en soit, il est nécessaire d'avoir des études comparatives de prévision des prix impliquant (i) les mêmes ensembles de données, (ii) les mêmes procédures robustes d'évaluation d'erreur, et (iii) des tests statistiques évaluant la performance d'un modèle par rapport à l'autre [140]. Les approches de prévision des

prix dans ce chapitre appartiennent à la dernière famille, qui est l'intelligence informatique.

4.3 Tarification de l'électricité dans le marché tunisien : approches proposées

Deux approches sont proposées, et qui nécessitent toutes les deux l'utilisation des compteurs intelligents. La première se base sur les coûts de production, et la deuxième sur l'analyse de la tarification actuelle. Les deux méthodes utilisent un système de tarification en temps réel (RTP) d'une journée à l'avance.

4.3.1 Première approche

Cette approche se base sur l'analyse la stratégie adoptée par le gestionnaire du réseau tunisien STEG pour répondre à la demande, et sur les coûts de production de chaque type de générateur. Elle utilise les données de 2010. La méthode adoptée est la tarification en temps réel RTP d'une journée à l'avance avec des valeurs semi-horaires, qui dépendent de la prévision de la demande L et d'un coefficient α représentant la prévision de la production renouvelable. Néanmoins, La méthode CPP est aussi présente en quelque mesure, car la tarification estivale diffère de la tarification hivernale.

En Tunisie, le parc de génération est presque entièrement alimenté par le gaz, et le pourcentage de cette ressource s'élève même à 97 %. La gamme des centrales à gaz en Tunisie comprend trois technologies principales : les centrales à Cycle Combiné (CC), les turbines à Cycle Ouvert (CO) (également appelées Turbines à Gaz TG) et les Turbines à Vapeur (TV). Il est à noter que, certaines centrales existantes en Tunisie permettent également d'être opérées avec des carburants liquides (notamment le fuel). Néanmoins, vu le prix élevé de ces carburants, ce type de fonctionnement est considéré seulement comme une option d'appoint pour les centrales à gaz.

La structure de fonctionnement des centrales électriques pour répondre à la courbe de charge est basée sur l'ordre successif suivant (données 2010) :

1. Les IPP (Independent Power Producer), étant donné les exigences du contrat de cession entre la STEG et les producteurs indépendants. Ces producteurs détiennent 471 MW CC (avec des TV de faible capacité). Le coût typique de génération de cette ressource est 61 €/MWh, soit environ 122 millimes/kWh.
2. Les centrales de base, du type cycle combiné. Cette technologie, dont la capacité installée s'élève à 360 MW, est en train de devenir la technologie dominante dans le parc de génération tunisien. Ce type de centrales sert à satisfaire la charge de base et intermédiaire. Le coût typique de génération est 61 €/MWh, soit environ 122 millimes/kWh.
3. Les centrales de base, du type turbine à vapeur, dont la capacité installée est de 990 MW. Le coût

typique de génération (estimé) est aussi de 61 €/MWh, soit environ 122 millimes/kWh.

4. Et en dernier lieu, les turbines à gaz cycle ouvert (les grandes puissances puis les petites), dont la capacité est de 1348 MW. La mise en marche peut être effectuée très rapidement : seulement quelques minutes après le démarrage, les turbines atteignent leur capacité nominale. C'est pourquoi ces centrales sont considérées comme une technologie de choix pour sécuriser la production électrique pendant les pointes de charge. Disponibles en plusieurs tailles sur le marché, les turbines à gaz CO permettent la construction de petites unités génératrices (10 MW) jusqu'à de grandes centrales de plusieurs mégawatts (300 MW). Le coût typique de génération est 130 €/MWh, soit environ 260 millimes/kWh.

Les turbines gas-oil ont une capacité de 100 MW. Leur coût de génération typique n'est pas spécifié, mais on l'estime par 520 millimes/kWh (double des turbines à gaz), vu le coût élevé du carburant.

Le placement des groupes hydrauliques (capacité 66 MW) est tributaire du programme journalier de l'irrigation, sachant que toute la production hydraulique est acceptée systématiquement par le réseau STEG quel que soit le programme d'irrigation. Pareillement pour la centrale éolienne, toute la production est acceptée par le réseau STEG quelle que soit l'heure et quelle que soit la journée. La capacité installée de cette ressource est de 54 MW, avec un coût typique de génération égal à 60 €/MWh, soit environ 120 millimes/kWh (données 2010) [141].

La température peut dégrader énormément le rendement des turbines. Par exemple, pour le cycle combiné de Ghannouch (sous essai), la puissance maximale à 20 °C est 420 MW, mais elle s'abaisse à 390 MW à 35 °C, soit une diminution de 7 %. Pour la turbine à gaz de Bir Mcherga, la puissance maximale à 20 °C est 120 MW, et à 35 °C est 105 MW, soit une diminution de 12,5 %.

Pour simplifier la procédure d'estimation des prix, on considère que le rendement de tous les types de générateurs à gaz est diminué de 10 % à 35 °C et plus. On suppose aussi que le prix serait égal au coût de production sans tenir en compte les coûts de transport, subventions, taxes, bénéfices et ainsi de suite. De plus, on considère que le coût du kWh produit par les centrales privées est égal à celui du kWh produit par la STEG. Étant donné toutes ces suppositions, il est possible d'identifier les tranches de production du tableau 4.1, qui représentent la charge de base, intermédiaire et de pointe, respectivement. Les centrales hydrauliques et éoliennes sont placées à la fin puisque leur production est intermittente.

En 2010, la demande minimale était de 928 MW et la demande maximale était 3010 MW. Donc pour toutes les saisons sauf l'été, le réseau est capable de satisfaire la demande et même assurer une réserve de puissance allant jusqu'à 379 MW (en présence des ressources intermittentes). En été par contre, et avec la dégradation du rendement des centrales à cause de la chaleur, la production peut être incapable de satisfaire la demande. La solution consiste à élever le prix au maximum sous ces

Tranches de production à 35°C (Été)	Tranches de production à 20°C (Autres saisons)	Coût de production
0-1638 MW (IPP+CC+TV)-10 %	0-1821 MW (IPP+CC+TV)	122 millimes/kWh
1638-2852 MW (TG)-10 %	1821-3169 MW (TG)	260 millimes/kWh
2852-2942 MW (gas-oil)-10 %	3169- 3269 MW (gas-oil)	520 millimes/kWh
2942-3062 MW (hydraulique+éolien)	3269-3389 MW (hydraulique+éolien)	120 millimes/kWh (intermittent)

Tableau 4.1 Tranches de production

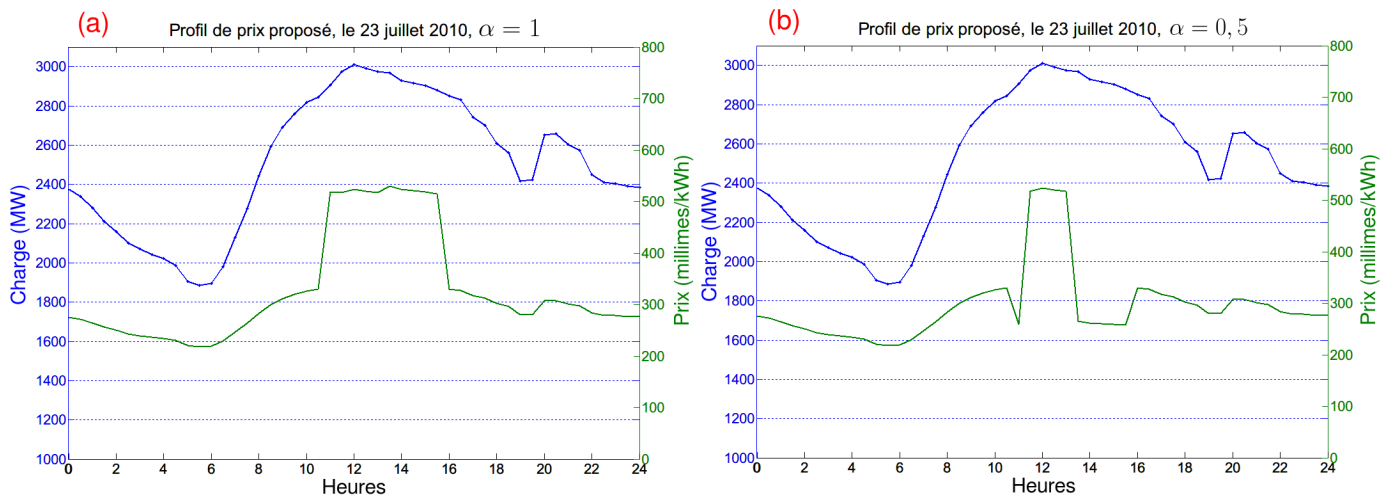


Figure 4.1 Profil du prix proposé, première approche

conditions. Par ailleurs, le prix est multiplié par un coefficient α ($0 < \alpha \leq 1$) dans des cas de faible demande avec présence de la production intermittente, afin d'encourager les consommateurs à en profiter. La réduction du prix dans ce cas serait proportionnelle à la quantité d'énergie intermittente disponible. Le tableau 4.2 récapitule la procédure de calcul du prix moyen de chaque tranche en été. Le prix réel P est ensuite calculé à partir du prix moyen et de la charge L . La prévision des prix est ainsi directement déduite à partir de la prévision de la charge par une transformation linéaire par intervalles.

En appliquant la démarche proposée, on obtient les courbes de prix visibles sur la figure 4.1, avec $\alpha = 1$ et $\alpha = 0,5$, respectivement. On remarque que la tarification est plus souple, qu'elle suit l'évolution de la charge et qu'elle reflète mieux les coûts de production. On remarque aussi qu'avec

Tranche	Charge moyenne (MW)	Prix moyen (millimes/kWh)	Calcul du prix
928-1638 MW	1283	122 α	$P = L/(1283/122\alpha)$
1638-2852 MW	2245	260	$P = L/(2245/260)$
2852-2972 MW	2912	520 α	$P = L/(2912/520\alpha)$
2972-3010 MW	2991	520	$P = L/(2991/520)$

Tableau 4.2 Tarification proposée en été de type RTP, première approche

la contribution de l'énergie renouvelable, la période caractérisée par un prix élevé rétrécit.

4.3.2 Deuxième approche

La deuxième approche s'appuie sur l'analyse de la tarification actuelle, et utilise les données de 2013. On a recours à cette méthode car la première peut être non précise, vu que les coûts de production sont estimés, et puisqu'elle ne tient pas en compte les autres facteurs (coûts de transport, entretien, subvention,...). La deuxième méthode consiste donc à analyser la tarification actuelle, qui est de type TOU, afin de pouvoir extraire les trois tranches de charge : base, intermédiaire, et pointe. Elle élabore par la suite une procédure RTP d'une journée à l'avance avec des valeurs semi-horaires, qui sont corrélées avec la prévision de la demande L .

La puissance installée du parc national tunisien de production en 2013 est divisée comme suit : 40 % turbines à gaz, 28 % cycle combiné, 25 % turbines à vapeur et 7 % énergie renouvelable. Il est difficile d'estimer les coûts de production de chaque ressource, puisqu'il faut avoir un aperçu total de toute la stratégie de production adoptée par le gestionnaire du réseau. Cependant, la tarification actuelle de l'électricité (figure 4.2) en donne une idée.

Le prix est généralement fixe, et ne dépend que de la quantité d'énergie consommée et de l'installation en kVA. Ceci revient essentiellement à l'utilisation des compteurs classiques avec lesquels il n'est pas possible de connaître la période de consommation. La seule exception est l'irrigation, dans laquelle le prix varie moyennant trois compteurs différents selon les postes horaires définis par la figure 4.3, adoptant ainsi l'approche TOU. Le prix est clairement maximal pendant les heures de pointe. L'idée consiste donc à en déduire les coûts de production.

La figure 4.4(a) donne le profil de la charge et du prix de l'irrigation pendant le lundi 5 mai 2014, divisé sur 3 postes horaires, selon la tarification de la figure 4.2. En s'appuyant sur la courbe de charge, on peut clairement classer les prix comme suit : 93 millimes/kWh si la demande est inférieure à 1800 MW, 103 millimes/kWh si elle est entre 1800 et 2000 MW, et 295 millimes/kWh si elle est supérieure à 2000 MW. Cette tarification est visiblement proportionnelle aux coûts de

Basse Tension Générale

TARIF	SECTEUR	REDEVANCE DE PUISSANCE (mill/kVA/mois)	PRIX D'ENERGIE POUR CHAQUE TRANCHE DE CONSOMMATION MENSUELLE (mill/kWh)					
			1-50	51-100	101-200	201-300	301-500	501 et +
<u>Tranche économique</u> (1 et 2 kVA & C° ≤ à 200 kWh/mois)	Résidentiel	500	75					
	Résidentiel & Non Résidentiel		108					
	Résidentiel & Non Résidentiel		140					
<u>Tranche économique</u> (1 et 2 kVA & C° > à 200 kWh/mois) <u>Tranche Normale</u> (> à 2 kVA)	Résidentiel	500	151			184	280	350
	Non Résidentiel						250	295

Basse Tension Spéciaux

TARIF	REDEVANCE		PRIX D'ENERGIE (mill/kWh)			
	D'ABONNEMENT (mill/Ab/mois)	DE PUISSANCE (mill/kVA/mois)	Jour	Pointe matin été	Pointe soir	Nuit
Eclairage Public	-	700	218			
Chauffe-eau	500	-	280			280
Chauffage & Climatisation	-	500	350			
Irrigation	Uniforme	300	135			
	Trois postes horaires	1 000	103		295	93

Figure 4.2 Tarification de l'électricité en Tunisie, à partir du 1^{er} mai 2014

production. Jusqu'ici, tout va bien. Le problème se pose en été, avec le profil de la figure 4.4(b), dans laquelle le prix est clairement sous-estimé. L'idée consiste à étendre les règles déjà établies à partir de la figure 4.4(a) à toutes les saisons en s'appuyant directement sur la courbe de charge et non pas sur les postes horaires, dans le but d'établir une tarification qui reflète mieux les coûts de production.

Se basant sur la figure 4.4, il est possible de diviser la plage de variation de la demande en trois parties en définissant la charge moyenne et le prix moyen de chacune d'elles (tableau 4.3). Le prix à chaque demi-heure est calculé en divisant la charge actuelle par le rapport charge moyenne/prix moyen, et ceci pour chaque partie. Par exemple, si la charge actuelle est 1376 MW, le prix estimé est $1376/(1500/93)$, soit 85 millimes/kWh. Par ailleurs, pour une charge de 2560 MW, le prix serait $2560/(2300/295)$, soit 328 millimes/kWh. La figure 4.5 propose cette nouvelle tarification pour les mêmes jours, 5 mai et 15 juillet 2014. On remarque que la courbe de prix est bien plus souple, et qu'elle reflète mieux les coûts de production. Cette démarche peut évidemment être étendue à toutes les tranches de consommation, dans lesquelles les prix moyens doivent être calculés de façon similaire. Comme pour la première approche, la prévision des prix est déduite à partir de la

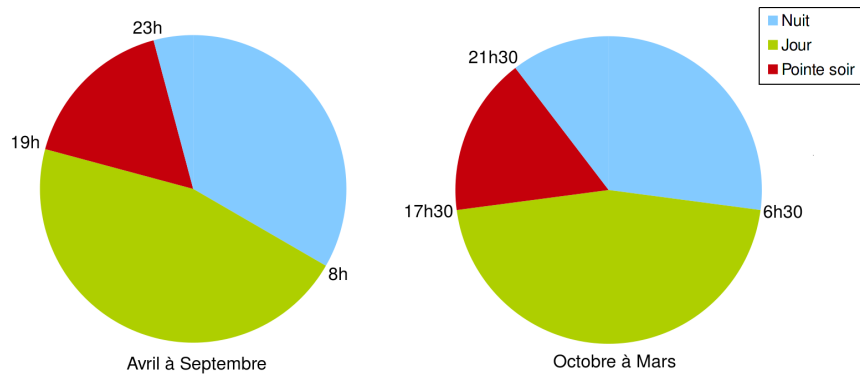


Figure 4.3 Postes horaires

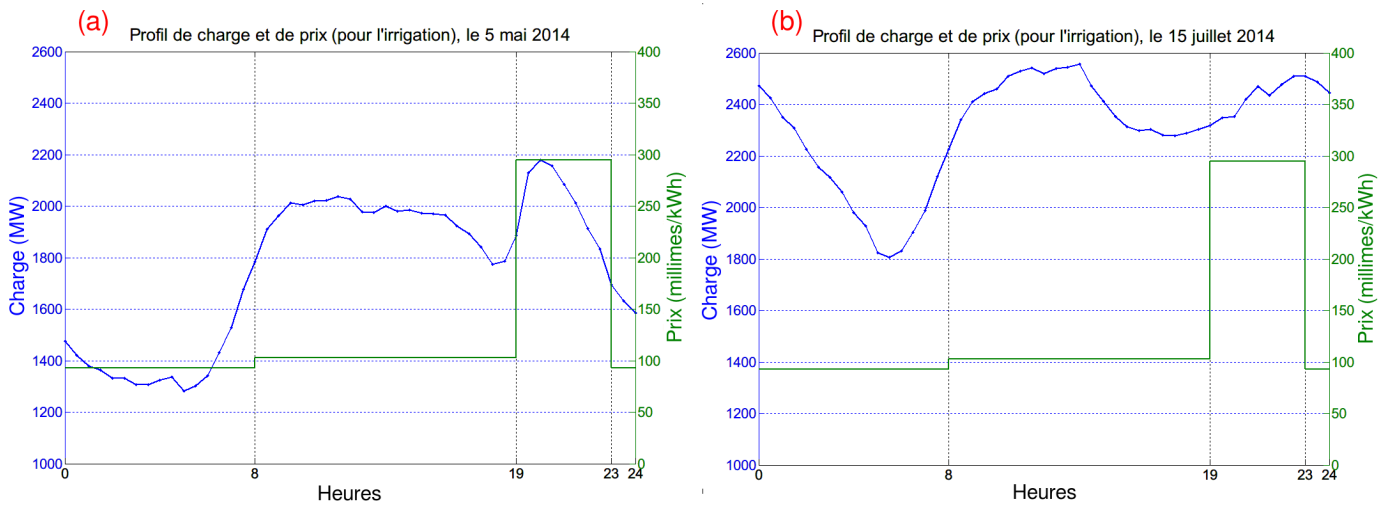


Figure 4.4 Profil actuel du prix en Tunisie (pour l'irrigation)

prévision de la charge par une transformation linéaire par intervalles.

Plage de variation de la charge	Charge moyenne (MW)	Prix moyen (millimes/kWh)	Calcul du prix
1200 – 1800 MW	1500	93	$P = L/(1500/93)$
1800 – 2000 MW	1900	103	$P = L/(1900/103)$
2000 – 2600 MW	2300	295	$P = L/(2300/295)$

Tableau 4.3 Tarification proposée de type RTP, deuxième approche

L'adoption d'une de ces deux approches nécessite des compteurs intelligents qui communiquent avec le gestionnaire du réseau, et qui affichent en temps quasi-réel la tarification de l'électricité

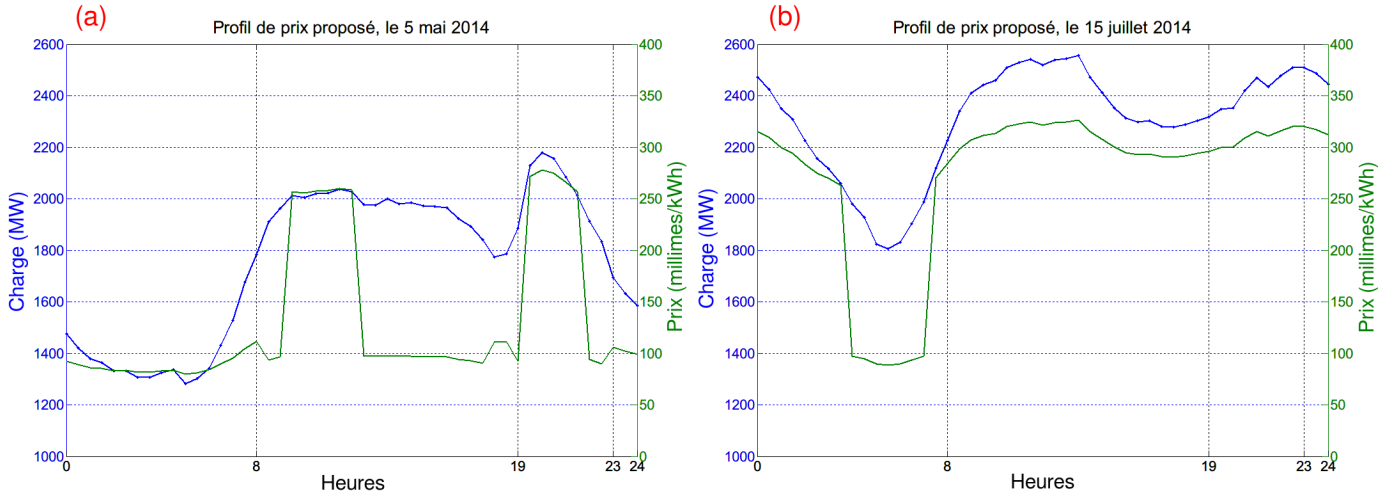


Figure 4.5 Profil du prix proposé, deuxième approche

ainsi que la consommation actuelle. Cette démarche est applicable en Tunisie tant que le marché d'électricité n'est pas encore dérégulé, car si c'est le cas, c'est la concurrence entre les différents générateurs qui définit les prix plutôt que les coûts de production. Et dans ces conditions, la transformation linéaire de la courbe de charge ne suffit pas, et il faut aborder la prévision des prix.

4.4 Tarification dans le marché australien dérégulé

Une fois la procédure de tarification est établie, on peut passer à la prévision des prix, qui n'est pas encore possible en Tunisie. Un marché dérégulé a donc été choisi pour étudier les relations entre la demande et le prix, et analyser la tarification dans des cas extrêmes. L'étude de cas réel est effectuée sur le marché australien, pour lequel certaines données sont accessibles et qui a déjà été considéré comme référence dans plusieurs travaux de la littérature. On rappelle que le marché est dérégulé s'il y a de la concurrence libre entre les différentes entreprises génératrices d'électricité. Le marché australien est caractérisé par sa très grande volatilité, vu qu'il utilise une structure de règlement en temps réel RTP, selon laquelle un nouveau prix est établi chaque demi-heure. Cette analyse servira par la suite dans la conception du prédicteur dans la section 4.5, dédiée à la prévision des prix.

4.4.1 Présentation du marché d'énergie australien

L'opérateur de marché d'énergie australien (AEMO) a été créé pour gérer les marchés d'électricité et de gaz à partir du 1^{er} juillet 2009. Depuis le début et le milieu des années 1990, la déréglementation et la privatisation de certains actifs étatiques a conduit à la création de marchés de négociation

de plus en plus ouverts et transparents, facilitant ainsi le commerce et la concurrence.

Depuis 1998, la production, la distribution et l'approvisionnement en électricité dans l'est et le sud de l'Australie a été menée dans le cadre du marché national de l'électricité (NEM). Le NEM relie cinq marchés régionaux (Queensland, New South Wales, Victoria, l'Australie du Sud et la Tasmanie). Il assure également la production de gros qui est transportée par des lignes de transmission à haute tension aux distributeurs d'électricité, qui la délivrent aux maisons et aux entreprises.

Le transport de l'électricité des producteurs vers les consommateurs est facilité par un « pool » (réservoir), dans lequel la production de tous les générateurs est agrégée et programmée à des intervalles de cinq minutes pour répondre à la demande. Le « pool » n'est pas une chose physique, mais un ensemble de procédures que l'AEMO gère en conformité avec la loi et le règlement national de l'électricité. Le marché utilise des systèmes sophistiqués pour envoyer des signaux aux producteurs leur demandant combien d'énergie ils doivent produire chaque cinq minutes pour que la production soit adaptée aux exigences des consommateurs, les réserves soient maintenues prêtes pour les situations d'urgence, et le prix actuel d'énergie puisse être calculé. L'infrastructure NEM comprend à la fois des actifs étatiques et privés gérés par de nombreux participants.

4.4.2 Analyse des profils

La restructuration et la déréglementation des marchés de l'électricité en Australie a entraîné des changements fondamentaux dans le comportement des prix de gros. Ces prix sont toujours caractérisés par une forte volatilité (la variance des prix est très grande), un fort retour à la moyenne (les prix ont tendance à fluctuer autour d'un point d'équilibre à long terme), et des hausses brusques et imprévues, des sauts ou des pointes qui se désintègrent rapidement. En effet, ils reflètent les caractéristiques inhérentes des marchés concurrentiels de l'électricité : la saisonnalité, les faibles coûts marginaux de production, l'impact des défaillances ou des pannes système, les contraintes sur l'interconnexion entre les marchés, le stockage limité, et même la manipulation du marché. En conséquence, les principaux participants à ces marchés, y compris les générateurs, les détaillants et les grands utilisateurs industriels, sont exposés à des risques de marché significatifs et sont obligés de prendre des mesures coûteuses de gestion des risques [142].

Le marché de New South Wales est un des cinq marchés gérés par l'AEMO. Le profil de charge et de prix de ce marché pendant trois jours du mois de mars 2013 est donné dans la figure 4.6.

Le profil de charge est plus ou moins régulier ; on peut distinguer vaguement les deux pointes de consommation du jour et du soir, et la baisse aux premières heures du matin. Les variations sont lisses et progressives. Cependant, le profil du prix est beaucoup plus chaotique ; les variations sont aiguës et quasi-aléatoires. Néanmoins, sa corrélation avec la charge reste visible, avec la baisse

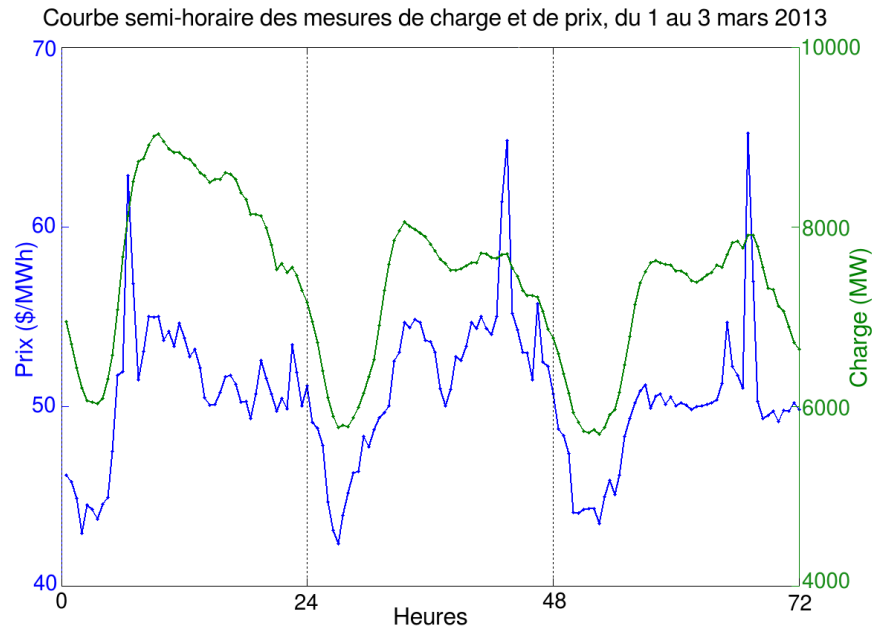


Figure 4.6 Profil de charge et de prix, mois de mars

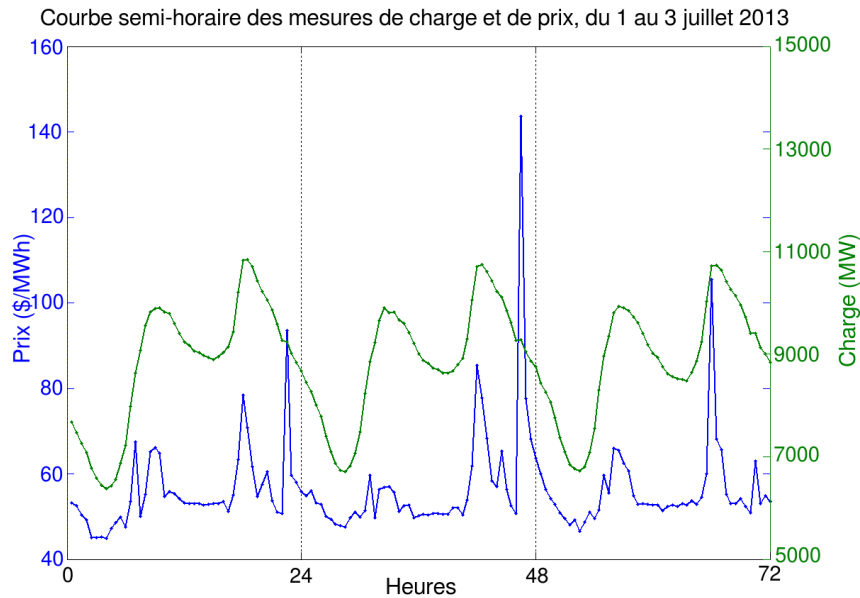


Figure 4.7 Profil de charge et de prix, mois de juillet

remarquable du prix aux premières heures du matin. Il est difficile d'expliquer toutes les variations de la courbe de prix et surtout les pics, vu qu'on ne dispose pas de beaucoup d'informations sur le marché et la procédure de tarification. En fait et comme on l'a déjà dit au chapitre 3, la prévision devient plus précise lorsque les entrées sont raffinées selon les caractéristiques du marché et des

consommateurs.

La figure 4.7 montre les deux courbes au mois de juillet. De même, la courbe de charge présente deux pics le matin et le soir, nettement plus visibles. Par contre, le prix est davantage plus dur à prédire ; dans le sens où il garde une valeur constante pendant de longues durées, mais présente des pics brusques parfois spectaculaires et très élevés. Les pics élevés se produisent généralement durant les pentes raides ou le changement de tendance de la courbe de charge. Ceci s'explique par le fait que les prix sont intentionnellement élevés durant les transitions afin de décourager les consommateurs, pour que les centrales aient le temps nécessaire pour démarrer ou pour s'arrêter.

4.5 Prédiction du prix de l'électricité dans le marché australien

Il s'agit dans cette section de développer un modèle de suivi des prix au marché australien décrit dans la section 4.4. Cette prédiction n'est pas étudiée en détails à cause du manque d'informations sur le réseau électrique australien, son parc de production, ses procédures de tarification, le climat dominant, la distribution de la population et le comportement des clients. La prédiction dans cette partie sert juste à évaluer et à comparer la performance des machines apprenantes par une prédiction « à l'aveugle », et à démontrer l'applicabilité des prédicteurs proposés dans d'autres contextes autres que le contexte tunisien.

4.5.1 Prédiction d'une demi-heure à l'avance

Dans cette section, la prédiction des prix d'électricité est effectuée sur un horizon d'une demi-heure (très court terme), par un pas d'une demi-heure également, soit $s = 1/2$. Des entrées endogènes sont considérées ; les valeurs passées et actuelles du signal de prix, ainsi que des entrées exogènes ; les valeurs passées, actuelles et futures du signal de charge. Les autres facteurs qu'on suppose influents ne sont pas considérés, telles que les conditions météo et les prix de carburant, à cause du manque d'informations à propos de la nature des générateurs utilisés. On suppose que la valeur future de charge (d'une demi-heure à l'avance) est déjà donnée et exacte, par recours à une méthode de prédiction. Le prix d'une demi-heure à l'avance serait donc en fonction de celle-ci, et de 3 échantillons passés de charge et de prix, comme indiqué dans la figure 4.8. Le nombre d'entrée m est alors égal à 7. Le modèle de prédiction s'écrit sous la forme :

$$\left\{ \begin{array}{l} X\left(t + \frac{1}{2}\right) = \left[L\left(t + \frac{1}{2}\right) \ L(t) \ L\left(t - \frac{1}{2}\right) \ L(t - 1) \ P(t) \ P\left(t - \frac{1}{2}\right) \ P(t - 1) \right]^T \\ Y(t) = P(t) = X^5(t) \\ \hat{Y}\left(t + \frac{1}{2}\right) = F\left(X\left(t + \frac{1}{2}\right)\right) \end{array} \right\} \quad (4.1)$$

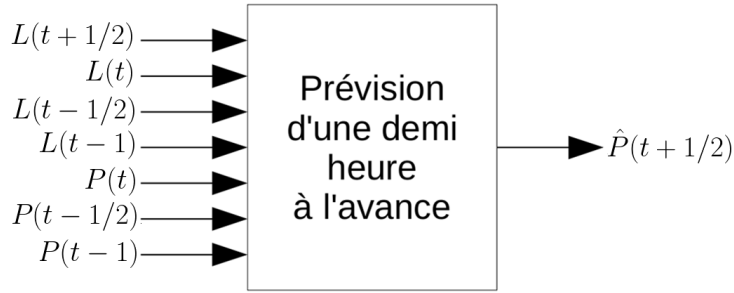


Figure 4.8 Modèle de prévision, demi-heure à l'avance

$L(t)$ représente la charge et $P(t)$ le prix à l'heure t . Donc $P(t + 1/2)$ est le prix futur (dans une demi-heure), $P(t)$ est le prix actuel, $P(t - 1/2)$ est le prix passé (il y a une demi-heure) et $P(t - 1)$ est le prix passé (il y a une heure). Il est de même pour la charge L .

Il est intéressant de tracer la réponse du modèle de persistance avant de procéder aux autres modèles. Ce modèle, dans lequel la prévision est simplement égale à la valeur actuelle, est une référence pour les méthodes de prévision à court terme. La figure 4.9(a) donne l'allure de cette réponse et l'erreur MAPE calculée de la même manière que celle des chapitres précédents.

Le critère MAPE pour les trois jours de test est déjà assez faible, d'environ 2,6 %. Le défi donc pour les autres modèles est de fournir une erreur bien plus petite.

Les mois de janvier et février 2013 sont choisis pour la phase d'apprentissage pour les trois machines issues de l'intelligence artificielle. L'apprentissage hors ligne est largement suffisant, puisque la période de test est courte ; les trois premiers jours du mois de mars 2013. À titre d'information, le nombre d'observations dans l'ensemble d'apprentissage $n = 2829$, le nombre de prédictions $\beta = 144$ et le nombre total d'erreurs $k = \beta$ (et non pas $k = \beta.s$, afin de prendre en compte toutes les erreurs puisque $s = 1/2$). Aucune méthode d'optimisation n'est proposée, afin d'assurer une comparaison équitable. Le réglage des machines est empirique et manuel. Les configurations choisies sont les suivantes :

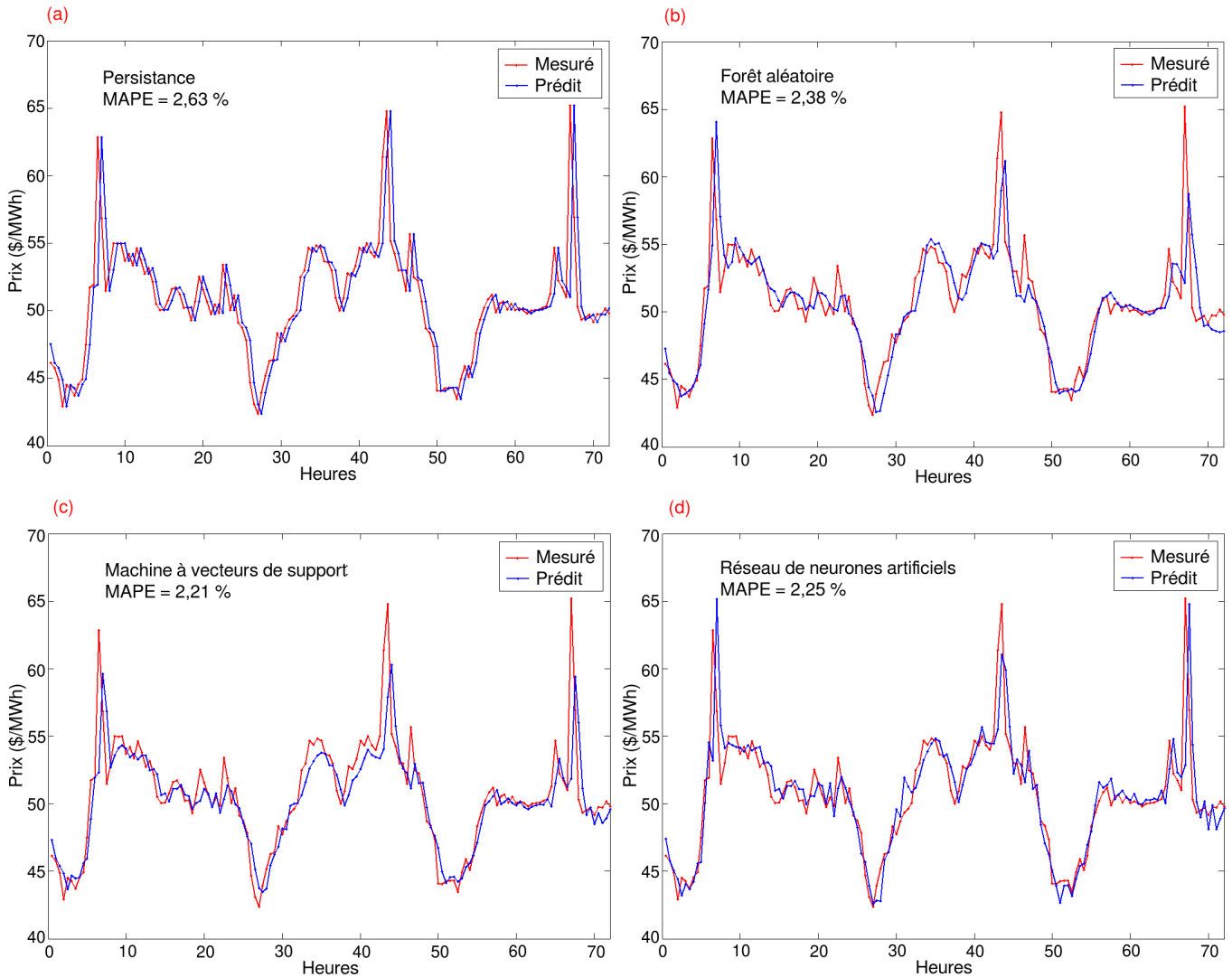


Figure 4.9 Prédiction d'une demi-heure à l'avance, 1 à 3 mars 2013

- La forêt aléatoire a comme paramètres : $ntree = 500$ et $mtry = 2$.
- La régression à vecteurs de support utilise un noyau polynômial de degré 1, $C = 10$ et $\gamma = 3.10^{-4}$.
- Le réseau de neurones artificiels est de type feedforward, et a une seule couche cachée de 20 neurones, la fonction d'activation sigmoïde et l'algorithme d'apprentissage Levenberg-Marquardt. La réponse de ce modèle, contrairement à ses deux prédécesseurs, peut varier énormément après chaque apprentissage.

On remarque que l'amélioration dans tous les cas est minime, l'erreur ne passe même pas sous le seuil de 2 % (figure 4.9). Afin d'avoir une meilleure prédiction, on essaye de changer certains

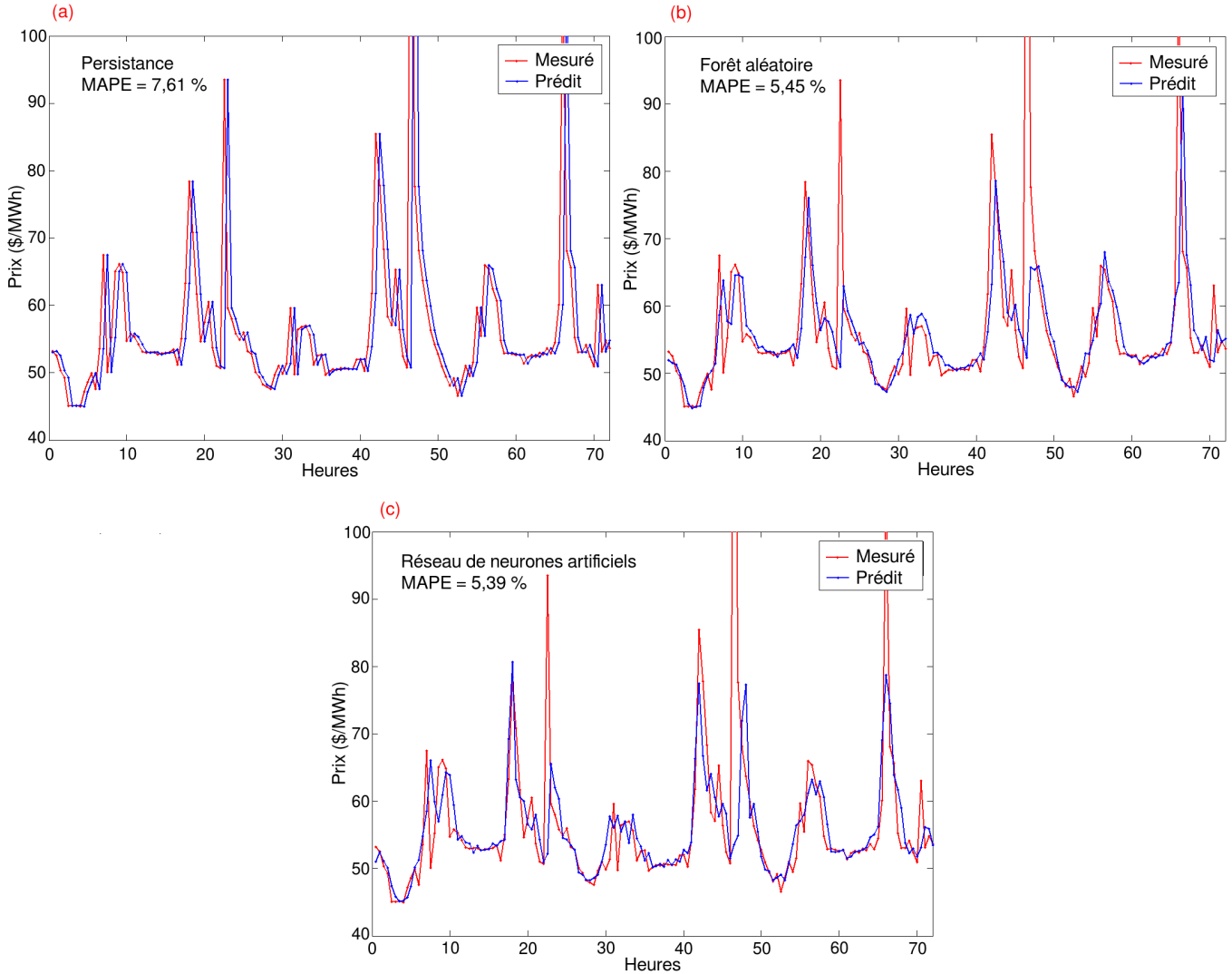


Figure 4.10 Prévision d'une demi-heure à l'avance, 1 à 3 juillet 2013

paramètres. C'est alors la taille de l'ensemble d'apprentissage qui est modifiée.

Un deuxième test est effectué sur les données de juillet, dont les résultats apparaissent sur la figure 4.10. Dans ce cas, l'ensemble d'apprentissage est bien plus grand, comportant 6 mois, de janvier à juin. La période de test couvre les trois premiers jours de juillet, soit donc $n = 8685$, $\beta = 144$ et $k = \beta$. Le profil est de sa part légèrement différent, comme expliqué dans la section 4.4.2. Le modèle de persistance conduit à une erreur assez élevée, à cause des pics spectaculaires de prix. Les autres prédicteurs gardent leurs configurations.

La forêt aléatoire abaisse l'erreur MAPE de plus de 2 %, ainsi que les réseaux de neurones artificiels. Par contre, la machine à vecteurs de support ne converge plus avec ses paramètres actuels, vu

que l'ensemble d'apprentissage S_n est devenu très grand.

La courbe de prévision a un aspect plus « intelligent » avec les machines à apprentissage automatique, et ne suit plus « machinalement » la courbe mesurée comme c'était le cas avec le modèle de persistance. Le réseau de neurones arrive à prédire efficacement les pics dans plusieurs reprises, et il donne le meilleur rendement dans ce contexte de prévision. Il n'y a surtout pas une méthode toujours performante quelles que soient les conditions, c'est toujours le contexte qui définit l'exactitude des approches utilisées.

4.5.2 Prévision d'une journée à l'avance

Le modèle proposé de prévision d'une journée à l'avance est donné par la figure 4.11.

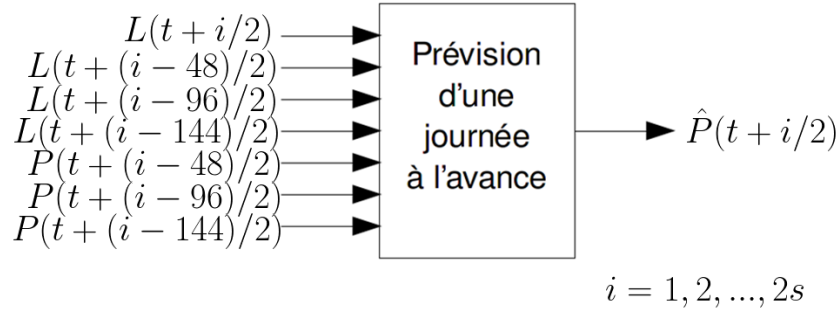


Figure 4.11 Modèle de prévision d'une journée à l'avance

Il s'agit dans cette section de construire $2s$ instances du modèle précédent, dédié à la prévision d'une demi-heure à l'avance. L'horizon de prévision est $s = 24$ puisqu'il s'agit d'une journée à l'avance, et il faut avoir $2s$ instances (soit 48) parce que le pas temporel est une demi-heure. Les entrées sont les mêmes, donc $m = 7$. Ce procédé est choisi puisqu'il a abouti à une prévision très performante au chapitre 3, contrairement au modèle dynamique.

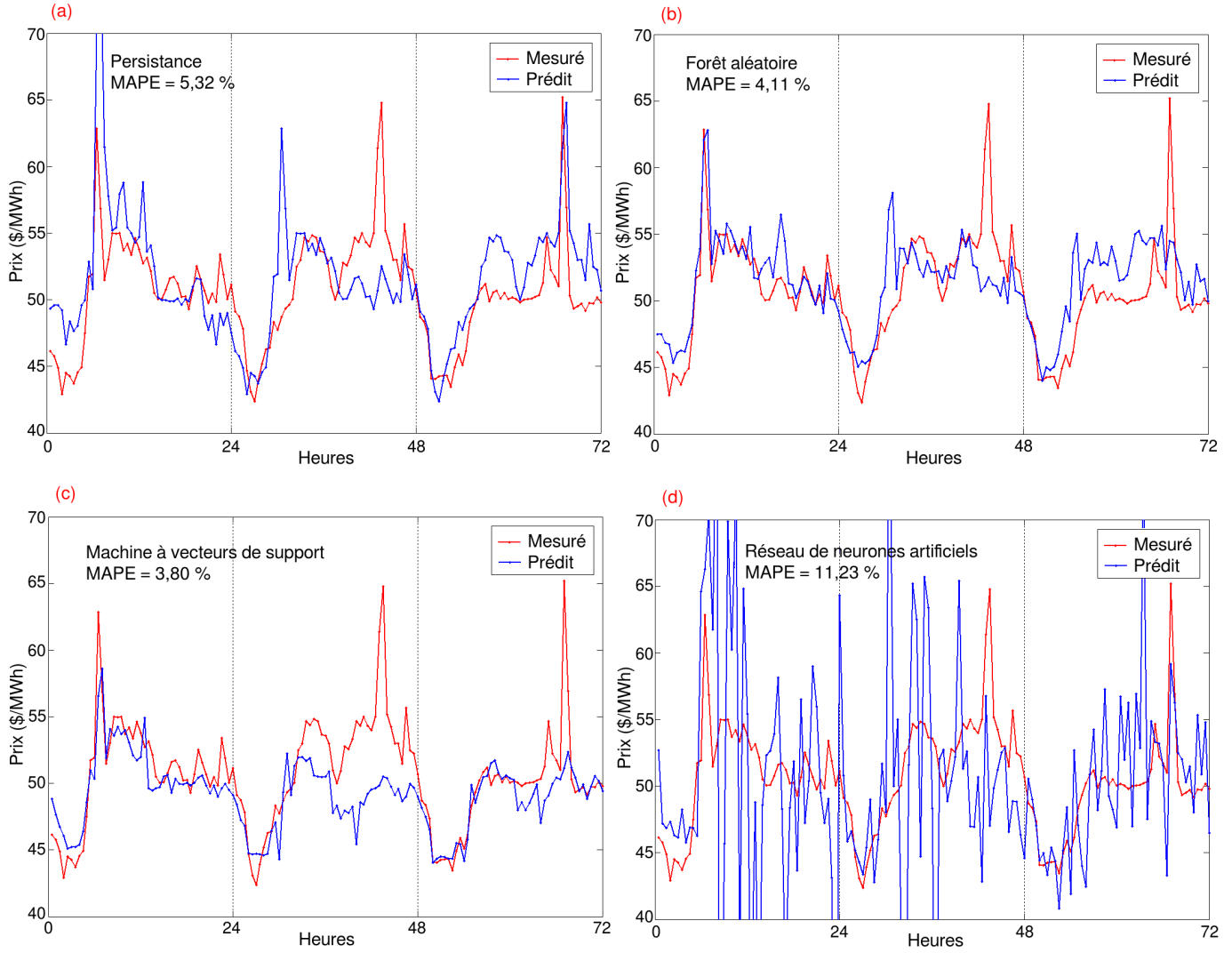


Figure 4.12 Prédiction d'une journée à l'avance, 1 à 3 mars 2013

Le modèle est le suivant :

$$\left\{ \begin{array}{l} X_L = \left[L\left(t + \frac{i}{2}\right) \quad L\left(t + \frac{i-48}{2}\right) \quad L\left(t + \frac{i-96}{2}\right) \quad L\left(t + \frac{i-144}{2}\right) \right] \\ X_P = \left[P\left(t + \frac{i-48}{2}\right) \quad P\left(t + \frac{i-96}{2}\right) \quad P\left(t + \frac{i-144}{2}\right) \right] \\ X\left(t + \frac{i}{2}\right) = [X_L \quad X_P]^T \\ Y(t) = P(t) = X^5(t) \text{ pour } i = 2s \\ \hat{Y}\left(t + \frac{i}{2}\right) = F\left(X\left(t + \frac{i}{2}\right)\right) \end{array} \right. \quad (4.2)$$

Avec $i = 1, 2, 3, \dots, 2s$.

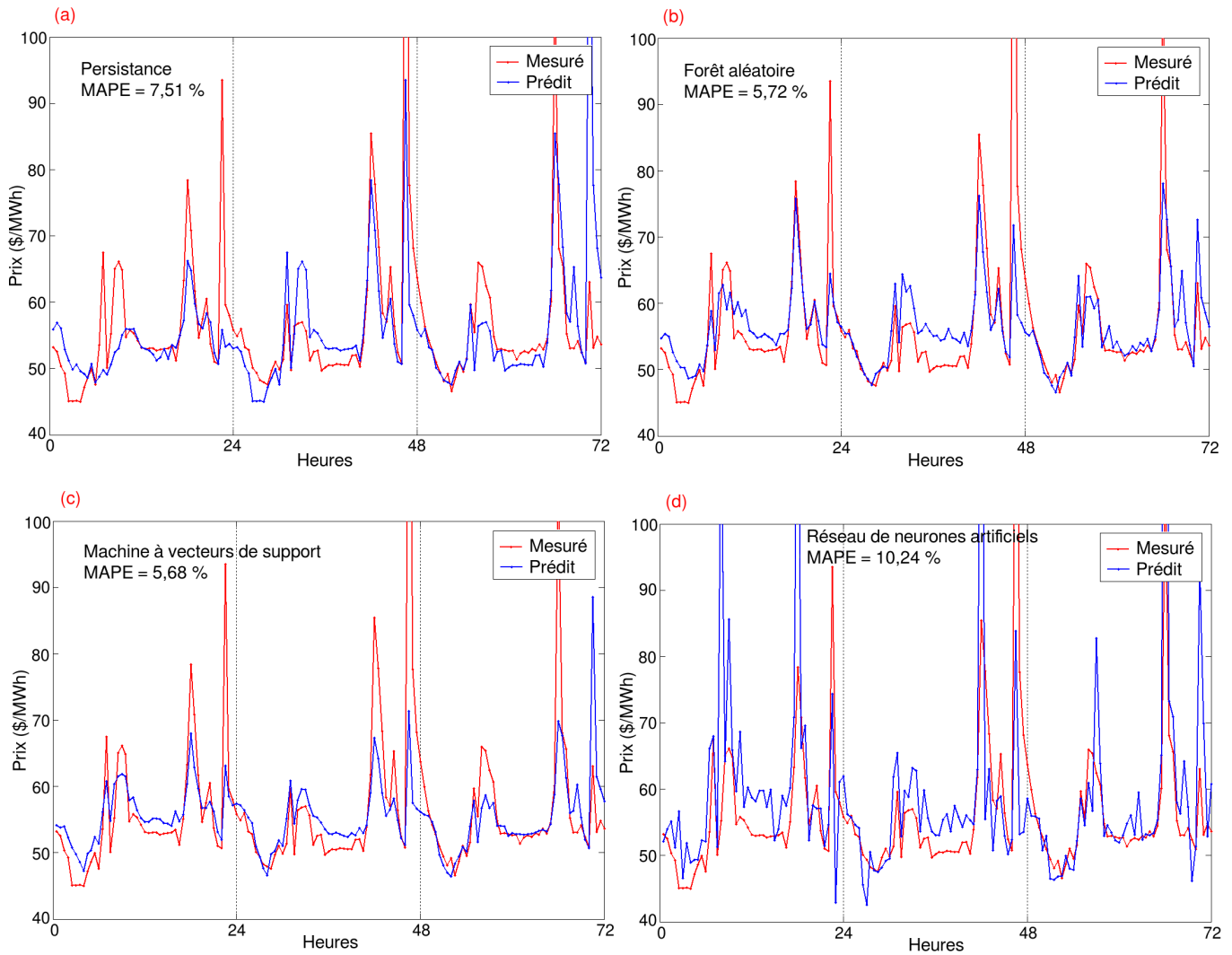


Figure 4.13 Prédiction d'une journée à l'avance, 1 à 3 juillet 2013

Cette démarche a été appliquée en utilisant les mêmes méthodes de la section 4.5.1, soit la persistance, la forêt aléatoire, la régression à vecteurs de support et le réseau de neurones (figure 4.12). Les ensembles d'apprentissage et de test sont les mêmes ; janvier et février pour l'apprentissage (hors ligne), et les trois premiers jours de mars pour le test. Évidemment, chacun des 48 modèles possède son propre ensemble S_n , avec $n = 56$. Le nombre de prédictions pour chaque modèle est $\beta = 3$ est le nombre total d'erreurs est $k = 2.s.\beta = 144$. La persistance dans ce cas est une recopie de toute la journée précédente, et non pas de la demi-heure précédente (voir la section 1.6.2).

La première remarque qui saute aux yeux est que le réseau de neurones n'est pas du tout adapté à ce

contexte de prévision. Le modèle de la figure 4.11 ne lui convient pas, surtout parce que l'ANN est de nature plus stochastique que celle des autres machines. Les deux autres méthodes, SVM et RF améliorent visiblement le modèle de persistance. Cependant, l'erreur reste relativement importante.

Pour le mois de juillet, il n'y a pas de grandes différences. Seulement, l'ensemble d'apprentissage est plus grand (janvier à juin), ce qui fait que $n = 178$ pour chacun des 48 modèles. La période de test de 3 jours fait que $\beta = 3$ et $k = 2.s.\beta = 144$. Cet élargissement de la période d'apprentissage a conduit à des résultats légèrement meilleurs (figure 4.13). L'algorithme de la régression à vecteurs de support donne la plus faible erreur MAPE. En contrepartie, le modèle RF permet de mieux détecter les pics. L'erreur de prévision dans la littérature varie énormément à cause de la grande volatilité et des spécificités de chaque marché, mais en général, elle varie entre 3 et 14 % [7, 103]. Les erreurs trouvées sont donc dans la marge de l'acceptable, bien qu'elles ne reflètent pas une prévision optimale.

4.6 Conclusion

Ce chapitre a permis d'aborder un aspect d'optimisation de l'énergie non disponible actuellement sur le réseau électrique tunisien, mais qui le serait dans le futur proche. Il consiste à introduire la variation des prix d'électricité en fonction de la demande, en commençant d'abord par proposer deux approches de tarification différentes, qui reflètent les coûts de production et qui sont mieux corrélées avec la charge demandée. Les approches proposées nécessitent l'utilisation des compteurs intelligents, et utilisent une procédure de tarification en temps réel RTP, qui permet de déterminer les prix d'une journée à l'avance. Elles sont introduites dans le but de gérer la demande côté client, afin d'éloigner les consommateurs des périodes de pointe, et améliorer par conséquence le rendement énergétique.

Ensuite, la prévision des prix proprement dite est étudiée pour un horizon d'une demi-heure et d'une journée à l'avance dans un marché dérégulé étranger. En utilisant les différentes techniques d'intelligence artificielle, les modèles sont évalués dans le but d'étendre ce type de prévision au réseau tunisien quand la variation du prix suivrait une tarification dynamique. Il est démontré que l'ANN donne de meilleurs résultats à très court terme, alors que la RF est plus performante pour des horizons plus longs. La SVM rencontre cependant des problèmes purement techniques. Ceci démontre que l'horizon de prévision a bel et bien son influence sur le choix de la méthode optimale. La conclusion générale permettra d'évaluer toutes les solutions proposées, de cerner leurs limitations et d'ouvrir les perspectives pour le travail futur.

CONCLUSION GÉNÉRALE

Ce mémoire de thèse a présenté une variété de méthodes de prévision à court terme de la production, de la charge et des prix de l'électricité dans un smart grid. À titre de rappel, dans un réseau électrique, il faut en permanence équilibrer l'offre et la demande d'électricité, ce qui entraîne nécessairement des procédures de prévision de la production d'une part, et de la consommation d'autre part. Et ce qui caractérise le smart grid en particulier, c'est le taux élevé des ressources intermittentes, la dérégulation des marchés et la tarification en temps réel, qui font que la prévision soit de plus en plus difficile. Dans ce sens, ce mémoire propose trois chapitres dédiés aux méthodes de prévision de la production éolienne, la prévision de la demande et la tarification dans le contexte tunisien, ainsi que des procédures de prévision des prix dans des marchés dérégulés étrangers.

La prévision de la production s'est limitée à l'énergie éolienne produite par le parc Sidi Daoud, pour un horizon d'une heure à l'avance, en raison de la nature très chaotique de ce signal. Le modèle conçu adopte une approche statistique et une méthode de prédiction indirecte, qui prédit d'abord les données du vent (vitesse et direction) et qui les convertit ensuite en puissance générée. Les techniques utilisées sont issues de l'intelligence artificielle et dont le choix a été bien justifié. La stratégie générale de prévision a exploité l'effet de lissage spatial pour limiter la variabilité de la ressource, en faisant l'agrégation de l'ensemble des aérogénérateurs. Des processus d'optimisation ont été ensuite proposés afin d'améliorer les résultats trouvés. La prévision de la charge a été développée pour le pays entier, pour un horizon d'une journée à l'avance par pas horaire. La problématique et les défis rencontrés ont été présentés en premier lieu. Le signal de charge, qui est beaucoup plus lisse que celui de la production renouvelable, a été analysé pour en extraire des profils et des saisonnalités. Ensuite, les entrées du modèle de prédiction ont été choisies avec un processus de raffinement et de sélection intelligente, et une stratégie d'apprentissage en ligne a été adoptée. Des tests dans les quatre saisons ainsi que dans des jours fériés sont par la suite effectués, et dont les résultats ont été évalués. Des améliorations sont ensuite ajoutées, en exploitant le processus de sélection intelligente et en concordance avec le contexte et la culture tunisienne locale. Des comparaisons sont aussi réalisées avec d'autres méthodes et marchés pour pouvoir se situer par rapport à l'existant dans la littérature. Une procédure de tarification dynamique est par la suite proposée pour le réseau tunisien, dans lequel la prévision des prix ne se pose encore pas, afin de mieux refléter le coût de production de l'électricité et d'être en corrélation avec la charge demandée. La prévision des prix a été appliquée dans un marché dérégulé pour un horizon d'une demi-heure et d'une journée à l'avance respectivement. Plusieurs méthodes ont été utilisées tout en évaluant leurs performances.

La nouveauté dans les solutions proposées se manifeste dans plusieurs aspects : l'exploitation de l'effet de lissage spatial pour prédire la vitesse du vent ainsi que la démonstration de l'effet de la direction et de l'agencement des aérogénérateurs sur la production, le développement des méthodes rarement utilisées telles que les forêts aléatoires pour la prévision de la charge, la construction d'un modèle de prédiction spécifique à la culture tunisienne et aux fêtes religieuses, l'élaboration d'une nouvelle procédure de tarification et d'estimation des prix, etc. Néanmoins, plusieurs limitations existent encore. La validité restreinte uniquement au court terme est parmi les limitations les plus évidentes. Et même au sein du court terme, certains modèles ne sont valides que pour une heure à l'avance, surtout dans le cas de la prévision éolienne. L'étude de la prévision à moyen et long terme nécessite évidemment d'autres procédures et implique d'autres facteurs d'influence. La performance de la prévision éolienne par l'approche statistique se dégrade énormément avec l'horizon de prévision, et c'est la raison pour laquelle la plupart des chercheurs utilisent en parallèle la prévision numérique de la météo ; un outil sophistiqué qui manque encore en Tunisie. Des limitations purement techniques existent aussi, c'est que les machines de prévision intelligentes ne peuvent pas apprendre les cas très rares et très particuliers, et ceci représente le défi majeur pour les gestionnaires du smart grid ; comment éviter les contingences et les pics. Un autre problème auquel fait face la procédure de tarification dynamique proposée, est que les compteurs intelligents ne soient pas encore disponibles sur le marché tunisien, ce qui limite la solution à l'aspect théorique.

Le travail futur peut être présenté à deux niveaux différents. Au niveau smart grid, il est possible d'abord d'étudier la prévision à moyen et à long terme, en analysant les facteurs qui pilotent les signaux à prédire sur ces horizons. Plusieurs combinaisons de méthodes intelligentes et de procédures d'optimisation sont possibles et pas encore exploitées, et ceci pour tous les types de prévisions. L'introduction des compteurs intelligents peut également ouvrir la perspective de la tarification en temps réel et de la prévision des prix. L'exploitation de l'énergie photovoltaïque à grande échelle amène à appliquer de nouvelles méthodes de prévision pour cette ressource renouvelable. Au niveau smart home maintenant, et qui est une issue naturelle du smart grid, des projets d'optimisation de la consommation à domicile sont en cours d'exploitation. Des systèmes embarqués dotés de capteurs et communicants à travers un réseau local prennent en charge la gestion énergétique en analysant les différentes informations envoyées par un compteur intelligent. De tels systèmes peuvent intégrer la prévision de la charge ou des prix afin de permettre aux habitants d'une résidence de bien gérer leur consommation. Que ce soit pour le smart grid ou le smart home, la prévision est un outil indispensable pour prendre des mesures d'efficacité énergétique et d'équilibrer en temps réel l'offre et la demande d'électricité, en évitant les coûts supplémentaires, le gaspillage d'énergie et les coupures inattendues.

RÉFÉRENCES

- [1] Y. Matsuo, *Prediction, Forecasting, and Chance Discovery*, in : Y. Ohsawa, P. McBurney (Eds.), *Chance Discovery*, Advanced Information Processing, Springer Berlin Heidelberg, 2003, pp. 30–43.
- [2] K. Moslehi, R. Kumar, *A Reliability Perspective of the Smart Grid*, *Smart Grid*, IEEE Transactions on 1 (1) (2010) 57–64.
- [3] C. Bergmeir, J. M. Benítez, *On the use of cross-validation for time series predictor evaluation*, *Information Sciences* 191 (2012) 192 – 213, data Mining for Software Trustworthiness.
- [4] K. Cheung, R. Rios-Zalapa, *Smart dispatch for large grid operations with integrated renewable resources*, in : *Innovative Smart Grid Technologies (ISGT)*, 2011 IEEE PES, 2011, pp. 1–7.
- [5] M. Q. Raza, A. Khosravi, *A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings*, *Renewable and Sustainable Energy Reviews* 50 (2015) 1352 – 1372.
- [6] K. He, L. Yu, L. Tang, *Electricity price forecasting with a {BED} (Bivariate {EMD} Denoising) methodology*, *Energy* 91 (2015) 601 – 609.
- [7] H. Shayeghi, A. Ghasemi, M. Moradzadeh, M. Nooshyar, *Simultaneous day-ahead forecasting of electricity price and load in smart grids*, *Energy Conversion and Management* 95 (0) (2015) 371 – 384.
- [8] Z. Zheng, Y. Chen, M. Huo, B. Zhao, *An Overview: the Development of Prediction Technology of Wind and Photovoltaic Power Generation*, *Energy Procedia* 12 (0) (2011) 601 – 608, the Proceedings of International Conference on Smart Grid and Clean Energy Technologies (ICSGCE 2011).
- [9] M. Zamo, O. Mestre, P. Arbogast, O. Pannekoucke, *A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production*, *Solar Energy* 105 (0) (2014) 792 – 803.
- [10] J. W. Taylor, *An evaluation of methods for very short-term load forecasting using minute-by-minute British data*, *International Journal of Forecasting* 24 (4) (2008) 645 – 658, energy Forecasting.
- [11] S. Soman, H. Zareipour, O. Malik, P. Mandal, *A review of wind power and wind speed forecasting methods with different time horizons*, in : *North American Power Symposium (NAPS)*, 2010, 2010, pp. 1–8.

- [12] J. Jung, R. P. Broadwater, *Current status and future advances for wind speed and power forecasting*, Renewable and Sustainable Energy Reviews 31 (2014) 762 – 777.
- [13] Y. Zhang, J. Wang, X. Wang, *Review on probabilistic forecasting of wind power generation*, Renewable and Sustainable Energy Reviews 32 (2014) 255 – 270.
- [14] L. J. Ricalde, G. Catzin, A. Y. Alanis, E. N. Sanchez, *Higher Order Wavelet Neural Networks with Kalman learning for wind speed forecasting*, in : Computational Intelligence Applications In Smart Grid (CIASG), 2011 IEEE Symposium on, 2011, pp. 1–6.
- [15] A. Anvari Moghaddam, A. Seifi, *Study of forecasting renewable energies in smart grids using linear predictive filters and neural networks*, Renewable Power Generation, IET 5 (6) (2011) 470–480.
- [16] F. Grimaccia, M. Mussetta, R. Zich, *Advanced predictive models towards PV energy integration in smart grid*, in : Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on, 2012, pp. 1–6.
- [17] A. M. Foley, P. G. Leahy, A. Marvuglia, E. J. McKeogh, *Current methods and advances in forecasting of wind power generation*, Renewable Energy 37 (1) (2012) 1 – 8.
- [18] A. Lahouar, J. B. H. Slama, *Day-ahead load forecast using random forest and expert input selection*, Energy Conversion and Management 103 (2015) 1040 – 1051.
- [19] J. G. D. Gooijer, R. J. Hyndman, *25 years of time series forecasting*, International Journal of Forecasting 22 (3) (2006) 443 – 473, twenty five years of forecasting.
- [20] H. Daumé, *A Course in Machine Learning*, 2012.
- [21] A. Tascikaraoglu, M. Uzunoglu, *A review of combined approaches for prediction of short-term wind speed and power*, Renewable and Sustainable Energy Reviews 34 (0) (2014) 243 – 254.
- [22] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences 454 (1971) (1998) 903–995.
- [23] S. Luke, *Essentials of Metaheuristics*, 2nd Edition, Lulu, 2013.
- [24] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, Z. Yan, *A review on the forecasting of wind speed and generated power*, Renewable and Sustainable Energy Reviews 13 (4) (2009) 915 – 920.
- [25] X. Zhao, S. Wang, T. Li, *Review of Evaluation Criteria and Main Methods of Wind Power Forecasting*, Energy Procedia 12 (0) (2011) 761 – 769.

- [26] I. Colak, S. Sagiroglu, M. Yesilbudak, *Data mining and wind power prediction: A literature review*, Renewable Energy 46 (0) (2012) 241 – 247.
- [27] L. Lazic, G. Pejanovic, M. Zivkovic, *Wind forecasts for wind power generation using the Eta model*, Renewable Energy 35 (6) (2010) 1236 – 1243.
- [28] W. Mahoney, K. Parks, G. Wiener, Y. Liu, W. Myers, J. Sun, L. Delle Monache, T. Hopson, D. Johnson, S. Haupt, *A Wind Power Forecasting System to Optimize Grid Integration*, Sustainable Energy, IEEE Transactions on 3 (4) (2012) 670–682.
- [29] F. Cassola, M. Burlando, *Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output*, Applied Energy 99 (0) (2012) 154 – 166.
- [30] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, G. Kallos, *Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering*, Journal of Wind Engineering and Industrial Aerodynamics 96 (12) (2008) 2348 – 2362.
- [31] C. Stathopoulos, A. Kaperoni, G. Galanis, G. Kallos, *Wind power prediction based on numerical and statistical models*, Journal of Wind Engineering and Industrial Aerodynamics 112 (0) (2013) 25 – 38.
- [32] S. Al-Yahyai, A. Gastli, Y. Charabi, *Probabilistic wind speed forecast for wind power prediction using pseudo ensemble approach*, in : Power and Energy (PECon), 2012 IEEE International Conference on, 2012, pp. 127–132.
- [33] J. Torres, A. Garcia, M. D. Blas, A. D. Francisco, *Forecast of hourly average wind speed with {ARMA} models in Navarre (Spain)*, Solar Energy 79 (1) (2005) 65 – 77.
- [34] A. Sfetsos, *A novel approach for the forecasting of mean hourly wind speed time series*, Renewable Energy 27 (2) (2002) 163 – 174.
- [35] L. Xie, Y. Gu, X. Zhu, M. Genton, *Short-Term Spatio-Temporal Wind Power Forecast in Robust Look-ahead Power System Dispatch*, Smart Grid, IEEE Transactions on 5 (1) (2014) 511–520.
- [36] M. A. Mohandes, S. Rehman, T. O. Halawani, *A neural networks approach for wind speed prediction*, Renewable Energy 13 (3) (1998) 345 – 354.
- [37] W. Wei, G. Wu, M. Yang, Y. Zhang, S. Qiu, A. Sun, *Short-term forecasting for wind speed based on wavelet decomposition and LMBP neural network*, in : Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011 4th International Conference on, 2011, pp. 1126–1131.
- [38] M. Bilgili, B. Sahin, A. Yasar, *Application of artificial neural networks for the wind speed prediction of target station using reference stations data*, Renewable Energy 32 (14) (2007) 2350 – 2360.

- [39] Z. Men, E. Yee, F.-S. Lien, D. Wen, Y. Chen, *Short-term wind speed and power forecasting using an ensemble of mixture density neural networks*, Renewable Energy 87, Part 1 (2016) 203 – 211.
- [40] T. Barbounis, J. Theocharis, *Locally recurrent neural networks for wind speed prediction using spatial correlation*, Information Sciences 177 (24) (2007) 5775 – 5797.
- [41] T. Barbounis, J. Theocharis, *A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation*, Neurocomputing 70 (7-9) (2007) 1525 – 1542, advances in Computational Intelligence and Learning, 14th European Symposium on Artificial Neural Networks 2006.
- [42] M. Mohandes, T. Halawani, S. Rehman, A. A. Hussain, *Support vector machines for wind speed prediction*, Renewable Energy 29 (6) (2004) 939 – 947.
- [43] S. Salcedo-Sanz, E. G. Ortiz-Garcia, A. M. Perez-Bellido, A. Portilla-Figueras, L. Prieto, *Short term wind speed prediction based on evolutionary support vector regression algorithms*, Expert Systems with Applications 38 (4) (2011) 4052 – 4057.
- [44] J. Hu, J. Wang, G. Zeng, *A hybrid forecasting approach applied to wind speed time series*, Renewable Energy 60 (0) (2013) 185 – 194.
- [45] Y. Ren, P. Suganthan, N. Srikanth, *A Comparative Study of Empirical Mode Decomposition-Based Short-Term Wind Speed Forecasting Methods*, Sustainable Energy, IEEE Transactions on 6 (1) (2015) 236–244.
- [46] S. Alessandrini, S. Sperati, P. Pinson, *A comparison between the {ECMWF} and {COSMO} Ensemble Prediction Systems applied to short-term wind power forecasting on real data*, Applied Energy 107 (0) (2013) 271 – 280.
- [47] L. LI, Y. qian LIU, Y. ping YANG, S. HAN, Y. mei WANG, *A physical approach of the short-term wind power prediction based on {CFD} pre-calculated flow fields*, Journal of Hydrodynamics, Ser. B 25 (1) (2013) 56 – 61.
- [48] H. Pousinho, V. Mendes, J. Catalao, *A hybrid PSO-ANFIS approach for short-term wind power prediction in Portugal*, Energy Conversion and Management 52 (1) (2011) 397 – 402.
- [49] J. Wu, B. Zhang, H. Li, Z. Li, Y. Chen, X. Miao, *Statistical distribution for wind power forecast error and its application to determine optimal size of energy storage system*, International Journal of Electrical Power & Energy Systems 55 (0) (2014) 100 – 107.
- [50] X. An, D. Jiang, C. Liu, M. Zhao, *Wind farm power prediction based on wavelet decomposition and chaotic time series*, Expert Systems with Applications 38 (9) (2011) 11280 – 11285.

- [51] M. Ozkan, P. Karagoz, *A Novel Wind Power Forecast Model: Statistical Hybrid Wind Power Forecast Technique (SHWIP)*, Industrial Informatics, IEEE Transactions on PP (99) (2015) 1–1.
- [52] J. Yan, Y. Liu, S. Han, M. Qiu, *Wind power grouping forecasts and its uncertainty analysis using optimized relevance vector machine*, Renewable and Sustainable Energy Reviews 27 (0) (2013) 613 – 621.
- [53] T. H. M. El-Fouly, E. El-Saadany, M. Salama, *One Day Ahead Prediction of Wind Speed and Direction*, Energy Conversion, IEEE Transactions on 23 (1) (2008) 191–201.
- [54] E. Erdem, J. Shi, *{ARMA} based approaches for forecasting the tuple of wind speed and direction*, Applied Energy 88 (4) (2011) 1405 – 1414.
- [55] S. Velázquez, J. A. Carta, J. Matías, *Influence of the input layer signals of {ANNs} on wind power estimation for a target site: A case study*, Renewable and Sustainable Energy Reviews 15 (3) (2011) 1556 – 1566.
- [56] A. Goia, C. May, G. Fusai, *Functional clustering and linear regression for peak load forecasting*, International Journal of Forecasting 26 (4) (2010) 700 – 711.
- [57] R. Ramanathan, R. Engle, C. W. Granger, F. Vahid-Araghi, C. Brace, *Short-run forecasts of electricity loads and peaks*, International Journal of Forecasting 13 (2) (1997) 161 – 174.
- [58] N. Amral, C. Ozveren, D. King, *Short term load forecasting using Multiple Linear Regression*, in : Universities Power Engineering Conference, 2007. UPEC 2007. 42nd International, 2007, pp. 1192–1198.
- [59] S. Pappas, L. Ekonomou, D. Karamousantas, G. Chatzarakis, S. Katsikas, P. Liatsis, *Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models*, Energy 33 (9) (2008) 1353 – 1360.
- [60] M. Cho, J. Hwang, C.-S. Chen, *Customer short term load forecasting by using ARIMA transfer function model*, in : Energy Management and Power Delivery, 1995. Proceedings of EMPD '95., 1995 International Conference on, Vol. 1, 1995, pp. 317–322 vol.1.
- [61] C.-M. Lee, C.-N. Ko, *Short-term load forecasting using lifting scheme and {ARIMA} models*, Expert Systems with Applications 38 (5) (2011) 5902 – 5911.
- [62] H. Al-Hamadi, S. Soliman, *Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model*, Electric Power Systems Research 68 (1) (2004) 47 – 59.
- [63] A. Bakirtzis, V. Petridis, S. Kiartzis, M. Alexiadis, A. Maissis, *A neural network short term load forecasting model for the Greek power system*, Power Systems, IEEE Transactions on 11 (2) (1996) 858–863.

- [64] S. Kiartzis, C. Zoumas, J. Theocharis, A. Bakirtzis, V. Petridis, *Short-term load forecasting in an autonomous power system using artificial neural networks*, Power Systems, IEEE Transactions on 12 (4) (1997) 1591–1596.
- [65] T. Senjyu, H. Takara, K. Uezato, T. Funabashi, *One-hour-ahead load forecasting using neural network*, Power Systems, IEEE Transactions on 17 (1) (2002) 113–118.
- [66] P. Mandal, T. Senjyu, T. Funabashi, *Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market*, Energy Conversion and Management 47 (15–16) (2006) 2128 – 2142.
- [67] P. Mastorocostas, J. Theocharis, A. Bakirtzis, *Fuzzy modeling for short term load forecasting using the orthogonal least squares method*, Power Systems, IEEE Transactions on 14 (1) (1999) 29–36.
- [68] S. Kiartzis, A. Bakirtzis, J. Theocharis, G. Tsagas, *A fuzzy expert system for peak load forecasting. Application to the Greek power system*, in : Electrotechnical Conference, 2000. MELECON 2000. 10th Mediterranean, Vol. 3, 2000, pp. 1097–1100 vol.3.
- [69] C.-T. Lin, L.-D. Chou, *A novel economy reflecting short-term load forecasting approach*, Energy Conversion and Management 65 (0) (2013) 331 – 342, global Conference on Renewable energy and Energy Efficiency for Desert Regions 2011.
- [70] S. Fan, K. Methaprayoon, W. Lee, *Multi-area load forecasting for system with large geographical area*, in : Industrial and Commercial Power Systems Technical Conference, 2008. ICPS 2008. IEEE/IAS, 2008, pp. 1–8.
- [71] G. Dudek, *Short-Term Load Forecasting Using Random Forests*, in : D. Filev, J. Jabłkowski, J. Kacprzyk, M. Krawczak, I. Popchev, L. Rutkowski, V. Sgurev, E. Sotirova, P. Szynekarczyk, S. Zadrozny (Eds.), Intelligent Systems’2014, Vol. 323 of *Advances in Intelligent Systems and Computing*, Springer International Publishing, 2015, pp. 821–828.
- [72] Y.-Y. Cheng, P. Chan, Z.-W. Qiu, *Random forest based ensemble system for short term load forecasting*, in : Machine Learning and Cybernetics (ICMLC), 2012 International Conference on, Vol. 1, 2012, pp. 52–56.
- [73] J. Wang, L. Li, D. Niu, Z. Tan, *An annual load forecasting model based on support vector regression with differential evolution algorithm*, Applied Energy 94 (0) (2012) 65 – 70.
- [74] J. Nagi, K. S. Yap, F. Nagi, S. K. Tiong, S. K. Ahmed, *A computational intelligence scheme for the prediction of the daily peak load*, Applied Soft Computing 11 (8) (2011) 4773 – 4788.
- [75] J. Che, J. Wang, G. Wang, *An adaptive fuzzy combination model based on self-organizing map and support vector regression for electric load forecasting*, Energy 37 (1) (2012) 657 – 664, 7th Biennial International Workshop “Advances in Energy Studies”.

- [76] S. Kulkarni, S. P. Simon, K. Sundareswaran, *A spiking neural network (SNN) forecast engine for short-term electrical load forecasting*, Applied Soft Computing 13 (8) (2013) 3628 – 3635.
- [77] R. Abdel-Aal, *Improving electric load forecasts using network committees*, Electric Power Systems Research 74 (1) (2005) 83 – 94.
- [78] Y. Lu, *Short-term load forecasting method based on structural neural network*, in : Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on, 2008, pp. 4434–4438.
- [79] F. Zhao, H. Su, *Short-Term Load Forecasting Using Kalman Filter and Elman Neural Network*, in : Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on, 2007, pp. 1043–1047.
- [80] D. Chaturvedi, A. Sinha, O. Malik, *Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network*, International Journal of Electrical Power & Energy Systems 67 (0) (2015) 230 – 237.
- [81] S. Ramos, J. Soares, Z. Vale, S. Ramos, *Short-term load forecasting based on load profiling*, in : Power and Energy Society General Meeting (PES), 2013 IEEE, 2013, pp. 1–5.
- [82] N. Tang, D.-J. Zhang, *Application of a Load Forecasting Model Based on Improved Grey Neural Network in the Smart Grid*, Energy Procedia 12 (0) (2011) 180 – 184, the Proceedings of International Conference on Smart Grid and Clean Energy Technologies (ICSGCE 2011).
- [83] S. Kouhi, F. Keynia, *A new cascade {NN} based method to short-term load forecast in deregulated electricity market*, Energy Conversion and Management 71 (0) (2013) 76 – 83.
- [84] M. Ghofrani, M. Ghayekhloo, A. Arabali, A. Ghayekhloo, *A hybrid short-term load forecasting with a new input selection framework*, Energy 81 (0) (2015) 777 – 786.
- [85] N. Amjady, F. Keynia, *Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm*, Energy 34 (1) (2009) 46 – 57.
- [86] C.-H. Wu, G.-H. Tzeng, R.-H. Lin, *A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression*, Expert Systems with Applications 36 (3, Part 1) (2009) 4725 – 4735.
- [87] D. Niu, Y. Wang, D. D. Wu, *Power load forecasting using support vector machine and ant colony optimization*, Expert Systems with Applications 37 (3) (2010) 2531 – 2539.
- [88] W.-C. Hong, *Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model*, Energy Conversion and Management 50 (1) (2009) 105 – 117.

- [89] A. Escobar, L. Perez, *Application of support vector machines and ANFIS to the short-term load forecasting*, in : Transmission and Distribution Conference and Exposition : Latin America, 2008 IEEE/PES, 2008, pp. 1–5.
- [90] Q. Wu, *Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts*, Expert Systems with Applications 38 (1) (2011) 379 – 385.
- [91] X. Yang, J. Yuan, J. Yuan, H. Mao, *An improved {WM} method based on {PSO} for electric load forecasting*, Expert Systems with Applications 37 (12) (2010) 8036 – 8041.
- [92] R. Enayatifar, H. J. Sadaei, A. H. Abdullah, A. Gani, *Imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS–ICA) for short term load forecasting*, Energy Conversion and Management 76 (0) (2013) 1104 – 1116.
- [93] A. U. Haque, P. Mandal, J. Meng, R. L. Pineda, *Performance Evaluation of Different Optimization Algorithms for Power Demand Forecasting Applications in a Smart Grid Environment*, Procedia Computer Science 12 (0) (2012) 320 – 325, complex Adaptive Systems 2012.
- [94] G. Sudheer, A. Suseelatha, *Short term load forecasting using wavelet transform combined with Holt–Winters and weighted nearest neighbor models*, International Journal of Electrical Power & Energy Systems 64 (0) (2015) 340 – 346.
- [95] N. Amjady, F. Keynia, *Mid-term load forecasting of power systems by a new prediction method*, Energy Conversion and Management 49 (10) (2008) 2678 – 2687.
- [96] E. González-Romera, M. Jaramillo-Morán, D. Carmona-Fernández, *Monthly electric energy demand forecasting with neural networks and Fourier series*, Energy Conversion and Management 49 (11) (2008) 3135 – 3142, special Issue 3rd International Conference on Thermal Engineering : Theory and Applications.
- [97] P.-C. Chang, C.-Y. Fan, J.-J. Lin, *Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach*, International Journal of Electrical Power & Energy Systems 33 (1) (2011) 17 – 27.
- [98] A. Azadeh, S. Ghaderi, S. Sohrabkhani, *Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors*, Energy Conversion and Management 49 (8) (2008) 2272 – 2278.
- [99] S. Yu, Y.-M. Wei, K. Wang, *China's primary energy demands in 2020: Predictions from an MPSO–RBF estimation model*, Energy Conversion and Management 61 (0) (2012) 59 – 66.
- [100] M. Piltan, H. Shiri, S. Ghaderi, *Energy demand forecasting in Iranian metal industry using linear and nonlinear models based on evolutionary algorithms*, Energy Conversion and Management 58 (0) (2012) 1 – 9.

- [101] S. Chan, K. Tsui, H. Wu, Y. Hou, Y.-C. Wu, F. Wu, *Load/Price Forecasting and Managing Demand Response for Smart Grids: Methodologies and Challenges*, Signal Processing Magazine, IEEE 29 (5) (2012) 68–85.
- [102] N. Shrivastava, A. Khosravi, B. Panigrahi, *Prediction Interval Estimation of Electricity Prices Using PSO-Tuned Support Vector Machines*, Industrial Informatics, IEEE Transactions on 11 (2) (2015) 322–331.
- [103] P. Pany, S. Ghoshal, *Dynamic electricity price forecasting using local linear wavelet neural network*, Neural Computing and Applications (2015) 1–9.
- [104] M. Mahmoud, M. Jaidane-Saidane, J. Souissi, N. Hizaoui, *Modeling of the Load Duration Curve using the Asymmetric Generalized Gaussian distribution: Case of the Tunisian Power System*, in : Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE, 2008, pp. 1–7.
- [105] O. Mohamed, M. Jaidane-Saidane, J. Ezzine, J. Souissi, N. Hizaoui, *Variability of Predictability of the Daily Peak Load Using Lyapunov Exponent Approach : Case of Tunisian Power System*, in : Power Tech, 2007 IEEE Lausanne, 2007, pp. 1078–1083.
- [106] N. M. Bellaaj, L. Bouzidi, M. Elleuch, *ANN Based Prediction of Wind and Wind Energy*, International Journal on Energy Conversion (IRECON) 2 (2).
- [107] B. Lotfi, B. Najiba, M. El Euch, *Impact of wind power integration on operating reserves*, in : Electrotechnical Conference (MELECON), 2012 16th IEEE Mediterranean, 2012, pp. 67–71.
- [108] B. Lotfi, M. Mourad, M. Najiba, E. Mohamed, *Treatment methodology of erroneous and missing data in wind farm dataset*, in : Systems, Signals and Devices (SSD), 2011 8th International Multi-Conference on, 2011, pp. 1–6.
- [109] I. Sansa, S. Missaoui, Z. Boussada, N. Bellaaj, E. Ahmed, M. Orabi, *PV power forecasting using different Artificial Neural Networks strategies*, in : Green Energy, 2014 International Conference on, 2014, pp. 54–59.
- [110] M. Benrejeb, P. Borne, J. Haggège, *Les réseaux de neurones : Présentation et applications*, Édition TECHNIP, 2007.
- [111] B. E. Boser, I. M. Guyon, V. N. Vapnik, *A Training Algorithm for Optimal Margin Classifiers*, in : Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, ACM Press, 1992, pp. 144–152.
- [112] H. Frezza-Buet, *Machines à Vecteurs Supports*, Tech. rep., Supélec (November 2012).
- [113] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, V. Vapnik, *Support Vector Regression Machines*, in : NIPS, 1996, pp. 155–161.

- [114] L. Breiman, et al., *Classification and Regression Trees*, Chapman & Hall, New York, 1984.
- [115] L. Breiman, *Random Forests*, *Machine Learning* 45 (1) (2001) 5–32.
- [116] R. Genuer, *Forêts aléatoires : aspects théoriques, sélection de variables et applications*, Ph.D. thesis, Université Paris Sud - Paris XI (2010).
- [117] S. B. Daly, Y. Fajraoui, *Base de données STEG*, Centre National des Mouvements d'Énergie, Radès, 2014.
- [118] P. S. Georgilakis, *Technical challenges associated with the integration of wind power into power systems*, *Renewable and Sustainable Energy Reviews* 12 (3) (2008) 852 – 863.
- [119] M. M. Alam, S. Rehman, L. Al-Hadhrami, J. Meyer, *Extraction of the inherent nature of wind speed using wavelets and {FFT}*, *Energy for Sustainable Development* 22 (0) (2014) 34 – 47, wind Power Special Issue.
- [120] GWEC, *Global Wind Report Annual Market Update 2013*, Tech. rep., Global Wind Energy Council (April 2014).
- [121] R. B. D. HASSEN, *Rapport Annuel 2013*, Tech. rep., Société Tunisienne de l'Électricité et du Gaz (STEG) (2013).
- [122] M. Ayed, *Projet Éolien de Sidi Daoud : Étude d'impact sur l'environnement de la centrale électrique éolienne*, Tech. rep., Société Tunisienne de l'Électricité et du Gaz (STEG) (Novembre 2009).
- [123] W. C. Briceno Vicente, *Modélisation des réseaux de distribution sous incertitudes*, Theses, Université de Grenoble (Sep. 2012).
- [124] P. Pinson, C. Chevallier, G. Kariniotakis, *Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power*, *Power Systems, IEEE Transactions on* 22 (3) (2007) 1148–1156.
- [125] M. Lydia, S. S. Kumar, *Wind Farm Power Prediction: An Overview*, in : *The Fifth International Symposium on Computational Wind Engineering (CWE2010)*, Chapel Hill, North Carolina, USA, 2010.
- [126] U. Focken, M. Lange, K. Mönnich, H.-P. Waldl, H. G. Beyer, A. Luig, *Short-term prediction of the aggregated power output of wind farms-a statistical analysis of the reduction of the prediction error by spatial smoothing effects*, *Journal of Wind Engineering and Industrial Aerodynamics* 90 (3) (2002) 231 – 246.
- [127] C.-C. Chang, C.-J. Lin, *LIBSVM : A library for support vector machines*, *ACM Transactions on Intelligent Systems and Technology* 2 (2011) 27 :1–27 :27, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

- [128] S. Zhou, J. Yuan, Z. Song, J. Tang, L. Kang, *Wind Signal Forecasting Based on System Identification Toolbox of MATLAB*, in : Intelligent System Design and Engineering Applications (ISDEA), 2013 Third International Conference on, 2013, pp. 1614–1617.
- [129] F. Javed, N. Arshad, F. Wallin, I. Vassileva, E. Dahlquist, *Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting*, Applied Energy 96 (0) (2012) 150 – 160, smart Grids.
- [130] A. Kaur, H. T. Pedro, C. F. Coimbra, *Impact of onsite solar generation on system load demand forecast*, Energy Conversion and Management 75 (0) (2013) 701 – 709.
- [131] S. Fan, Y.-K. Wu, W.-J. Lee, C.-Y. Lee, *Comparative study on load forecasting technologies for different geographical distributed loads*, in : Power and Energy Society General Meeting, 2011 IEEE, 2011, pp. 1–8.
- [132] A. Liaw, M. Wiener, *Classification and Regression by randomForest*, R News 2 (3) (2002) 18–22.
- [133] H. Mori, E. Kurata, *An efficient kernel machine technique for short-term load forecasting under smart grid environment*, in : Power and Energy Society General Meeting, 2012 IEEE, 2012, pp. 1–4.
- [134] S. Kouhi, F. Keynia, S. N. Ravadanegh, *A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection*, International Journal of Electrical Power & Energy Systems 62 (0) (2014) 862 – 867.
- [135] A. Kaur, H. T. Pedro, C. F. Coimbra, *Ensemble re-forecasting methods for enhanced power load prediction*, Energy Conversion and Management 80 (0) (2014) 582 – 590.
- [136] S. Fan, L. Chen, W.-J. Lee, *Machine learning based switching model for electricity load forecasting*, Energy Conversion and Management 49 (6) (2008) 1331 – 1344.
- [137] S. Chakraborty, T. Ito, T. Senjyu, *Smart pricing scheme: A multi-layered scoring rule application*, Expert Systems with Applications 41 (8) (2014) 3726 – 3735.
- [138] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, N. A. Khan, *Load forecasting, dynamic pricing and {DSM} in smart grid: A review*, Renewable and Sustainable Energy Reviews 54 (2016) 1311 – 1322.
- [139] A. JAIDANE, M. SASSI, *La télécommande centralisée : outil d’optimisation de la demande électrique, Revue de l’Électricité et du Gaz REG*, Tech. rep., Société Tunisienne de l’Électricité et du Gaz (STEG) (février 2009).
- [140] R. Weron, *Electricity price forecasting: A review of the state-of-the-art with a look into the future*, International Journal of Forecasting 30 (4) (2014) 1030 – 1081.

- [141] S. Lechtenböhmer, R. Missaoui, *Étude Stratégique du Mix Énergétique pour la Production d'Électricité en Tunisie*, Tech. rep., Institut Wuppertal du Climat, de l'Environnement et de l'Énergie (Avril 2012).
- [142] H. Higgs, A. Worthington, *Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The Australian wholesale spot electricity market*, Energy Economics 30 (6) (2008) 3172 – 3185, technological Change and the Environment.
- [143] A. Ranganathan, *The levenberg-marquardt algorithm*, Tutorial on LM Algorithm (2004) 1–5.
- [144] M. Hagan, M. Menhaj, *Training feedforward networks with the Marquardt algorithm*, Neural Networks, IEEE Transactions on 5 (6) (1994) 989–993.
- [145] V. Torczon, *On the Convergence of Pattern Search Algorithms*, SIAM Journal on Optimization 7 (1) (1997) 1–25.

ANNEXE A L'algorithme Levenberg-Marquardt

L'algorithme de Levenberg-Marquardt (LM) est un des algorithmes d'optimisation numérique les plus utilisés. Il dépasse la descente de gradient simple en termes de performance. Il est utilisé pour l'apprentissage du réseau de neurones artificiels.

Le problème pour lequel l'algorithme LM fournit une solution est appelé minimisation non-linéaire par moindres carrés. Ça implique une fonction à minimiser sous la forme [143] :

$$f(x) = \frac{1}{2} \sum_{j=1}^m r_j^2(x) \quad (\text{A.1})$$

Où $x = (x_1, \dots, x_n)$ est un vecteur, et chaque r_j est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. On admet que $m \leq n$.

Pour avoir une forme plus simple, f peut être représentée par un vecteur $r : \mathbb{R}^n \rightarrow \mathbb{R}^m$ défini par :

$$r(x) = (r_1(x), r_2(x), \dots, r_m(x)) \quad (\text{A.2})$$

Maintenant, f peut être réécrite sous la forme $f(x) = \frac{1}{2} \|r(x)\|^2$. Les dérivées de f peuvent être écrites en utilisant la matrice jacobienne J de r par rapport à x , définie par $J(x) = \frac{\partial r_j}{\partial x_i}, 1 \leq j \leq m, 1 \leq i \leq n$.

Considérons d'abord le cas linéaire dans lequel chaque r_j est une fonction linéaire de x . La matrice jacobienne est alors constante et on peut représenter r comme un hyperplan à travers l'espace, de façon que f soit donnée par la forme quadratique $f(x) = \frac{1}{2} \|Jx + r(0)\|^2$. On peut aussi obtenir $\nabla f(x) = J^T(Jx + r)$ et $\nabla^2 f(x) = J^T J$. Le problème de minimisation est résolu en mettant $\nabla f(x) = 0$, d'où $x_{min} = -(J^T J)^{-1} J^T r$.

En revenant au cas non-linéaire, on a :

$$\nabla f(x) = \sum_{j=1}^m r_j(x) \nabla r_j(x) = J(x)^T r(x) \quad (\text{A.3})$$

$$\nabla^2 f(x) = J(x)^T J(x) + \sum_{j=1}^m r_j(x) \nabla^2 r_j(x) \quad (\text{A.4})$$

La propriété distinctive des problèmes des moindres carrés, c'est qu'en étant donné la matrice jacobienne J , il est possible d'obtenir directement la matrice hessienne $\nabla^2 f(x)$ s'il est possible

d'approcher les r_j par des fonctions linéaires (lorsque les $\nabla^2 r_j(x)$ sont petits) ou si les $r_j(x)$ eux-mêmes sont petits. La matrice hessienne dans ce cas devient :

$$\nabla^2 f(x) = J(x)^T J(x) \quad (\text{A.5})$$

Ce qui est la même chose pour le cas linéaire.

La descente de gradient est la méthode la plus simple et la plus intuitive pour trouver les minima d'une fonction. La mise à jour des paramètres est effectuée en ajoutant l'opposé du gradient à chaque étape (multiplié par une constante) :

$$x_{i+1} = x_i - \lambda \nabla f \quad (\text{A.6})$$

Cette méthode peut être améliorée en utilisant la dérivée seconde, autrement dit l'information de courbure. La démarche consiste à utiliser la méthode de Newton, c'est-à-dire développer le gradient de f en utilisant les séries de Taylor autour de la variable courante x_0 .

$$\nabla f(x) = \nabla f(x_0) + (x - x_0)^T \nabla^2 f(x_0) + o(x - x_0) \quad (\text{A.7})$$

En ignorant les termes d'ordre élevé (on suppose que f est quadratique autour de x_0), et en résolvant l'équation $\nabla f(x) = 0$, on obtient la règle de mise à jour de la méthode de Newton :

$$x_{i+1} = x_i - (\nabla^2 f(x_i))^{-1} \nabla f(x_i) \quad (\text{A.8})$$

Où x_0 et x ont été remplacés par x_i et x_{i+1} , respectivement.

Puisque la méthode de Newton suppose que f est quadratique, il n'est pas nécessaire de calculer la matrice hessienne avec exactitude. L'approximation de (A.5) peut alors être utilisée. Le principal avantage de cette méthode est sa convergence rapide. Cependant, cette convergence est très sensible au point initial (plus précisément, à la linéarité autour du point initial).

Levenberg propose une règle de mise à jour qui combine les avantages de la descente du gradient et de la méthode de Gauss-Newton :

$$x_{i+1} = x_i - (H + \lambda I)^{-1} \nabla f(x_i) \quad (\text{A.9})$$

Où H est la matrice hessienne évaluée à x_i . Avec l'approximation de (A.5), ça revient à dire :

$$x_{i+1} = x_i - (J(x_i)^T J(x_i) + \lambda I)^{-1} J(x_i)^T r(x_i) \quad (\text{A.10})$$

La règle de mise à jour est utilisée comme suit. Si l'erreur diminue après une mise à jour, ça implique que l'hypothèse de forme quadratique sur $f(x)$ est correcte et on réduit λ (généralement par un facteur de 10) afin de diminuer l'influence de la descente du gradient. Et si l'erreur s'élève, on a tendance à suivre plus le gradient et alors λ est multiplié par le même facteur. L'algorithme alors :

- effectue la mise à jour en utilisant la règle (A.10)
- évalue l'erreur
- si l'erreur augmente, alors la mise à jour est annulée (donc les poids reviennent à leurs valeurs précédentes), le facteur λ est multiplié par 10 (ou une autre valeur significative) et on essaye une nouvelle mise à jour
- si l'erreur diminue, alors la mise à jour est gardée et le facteur λ est diminué.

Marquardt a remplacé la matrice identité dans (A.9) par la diagonale de la matrice hessienne, donnant ainsi la forme finale de la règle de mise à jour de l'algorithme LM :

$$x_{i+1} = x_i - (H + \lambda \text{diag}(H))^{-1} \nabla f(x_i) \quad (\text{A.11})$$

L'algorithme LM n'est pas optimal mais plutôt heuristique. Néanmoins, il fonctionne très bien en pratique. La seule limitation est qu'il nécessite l'inversion de la matrice à chaque itération. Cet inconvénient est surmonté par l'utilisation des techniques intelligentes de pseudo-inversion. Donc, pour des valeurs de n n'excédant pas quelques centaines, il est généralement beaucoup plus rapide que la descente du gradient [143].

ANNEXE B L'apprentissage par rétropropagation

On considère un réseau feedforward à M couches. L'objectif de l'apprentissage est de faire apprendre à la machine les associations entre des paires d'entrée-sortie qu'on appelle communément des patrons. On va donc remplacer S_n par un ensemble de Q patrons $\{(p_1, t_1), (p_2, t_2), \dots, (p_Q, t_Q)\}$, où p_q est l'entrée imposée et t_q est la sortie désirée. Le système d'équations du réseau de neurones s'écrit alors sous la forme :

$$\underline{z}^{(0)} = \underline{p} \quad (\text{B.1})$$

$$\underline{z}^{(l+1)} = \underline{f}^{(l+1)}(W^{(l+1)}\underline{z}^{(l)} + \underline{w}_0^{(l+1)}), l = 0, 1, \dots, M - 1 \quad (\text{B.2})$$

Le nombre n_l est le nombre de neurones dans la couche l , n_0 désigne alors le nombre d'entrées et n_M est le nombre de sorties.

L'indice de performance de la machine est donné par :

$$V = \frac{1}{2} \sum_{q=1}^Q (\underline{t}_q - \underline{z}_q^{(M)})^T (\underline{t}_q - \underline{z}_q^{(M)}) = \frac{1}{2} \sum_{q=1}^Q \underline{e}_q^T \underline{e}_q \quad (\text{B.3})$$

Où $\underline{z}_q^{(M)}$ est la sortie du réseau lorsque la $q^{\text{ème}}$ entrée $\underline{p}_q$ est présentée, et $\underline{e}_q$ est l'erreur correspondante [144].

Rétropropagation par descente du gradient

La rétropropagation standard utilise l'algorithme de la plus profonde descente (ou descente du gradient). L'indice de performance est approché par :

$$\hat{V} = \frac{1}{2} \underline{e}_q^T \underline{e}_q \quad (\text{B.4})$$

Où la somme des carrés est remplacé par l'erreur carrée $\underline{e}_q^T \underline{e}_q$ d'un seul patron $(\underline{p}_q, \underline{t}_q)$. L'algorithme donne alors :

$$\Delta w_{ij}^{(l)} = -\alpha \frac{\partial \hat{V}}{\partial w_{ij}^{(l)}} \quad (\text{B.5})$$

$$\Delta w_{i0}^{(l)} = -\alpha \frac{\partial \hat{V}}{\partial w_{i0}^{(l)}} \quad (\text{B.6})$$

Où α est le taux d'apprentissage. On définit $\delta_i^{(l)}$ comme l'indice de sensibilité de l'indice de performance aux changements de l'entrée du neurone i de la couche l .

$$\delta_i^{(l)} \equiv \frac{\partial \hat{V}}{\partial y_i^{(l)}} \quad (\text{B.7})$$

Il convient alors, en utilisant (1.9), (B.4) et (B.7) que :

$$\frac{\partial \hat{V}}{\partial w_{ij}^{(l)}} = \frac{\partial \hat{V}}{\partial y_i^{(l)}} \frac{\partial y_i^{(l)}}{\partial w_{ij}^{(l)}} = \delta_i^{(l)} z_j^{(l-1)} \quad (\text{B.8})$$

$$\frac{\partial \hat{V}}{\partial w_{i0}^{(l)}} = \frac{\partial \hat{V}}{\partial y_i^{(l)}} \frac{\partial y_i^{(l)}}{\partial w_{i0}^{(l)}} = \delta_i^{(l)} \quad (\text{B.9})$$

Les sensibilités satisfont aussi cette relation de récurrence :

$$\underline{\delta}^{(l)} = \dot{F}^{(l)}(\underline{y}^{(l)}) W^{(l+1)T} \underline{\delta}^{(l+1)} \quad (\text{B.10})$$

Avec :

$$\dot{F}^{(l)}(\underline{y}^{(l)}) = \begin{bmatrix} \dot{f}^{(l)}(y_1^{(l)}) & 0 & \dots & 0 \\ 0 & \dot{f}^{(l)}(y_2^{(l)}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dot{f}^{(l)}(y_{n_l}^{(l)}) \end{bmatrix} \quad (\text{B.11})$$

Et :

$$\dot{f}^{(l)}(y) = \frac{df^{(l)}(y)}{dy} \quad (\text{B.12})$$

La relation de récurrence est initialisée à la dernière couche :

$$\underline{\delta}^{(M)} = -\dot{F}^{(M)}(\underline{y}^{(M)})(\underline{t}_q - \underline{z}_q) \quad (\text{B.13})$$

L'algorithme procède alors comme suit : on propage l'entrée à l'avant en utilisant (B.1)-(B.2), ensuite, on propage les sensibilités à l'arrière par (B.13) et (B.10), et finalement on met à jour les poids et les biais en utilisant (B.5), (B.6), (B.8) et (B.9) [144].

Rétropropagation Levenberg-Marquardt

On reprend l'algorithme de Levenberg-Marquardt de l'annexe A, mais utilisé cette fois pour la rétropropagation dans le réseau de neurones. On rappelle que la règle de mise à jour de cet algorithme

s'écrit sous la forme :

$$\Delta x = (J(x)^T J(x) + \lambda I)^{-1} J(x)^T r(x) \quad (\text{B.14})$$

Qui sert à minimiser une fonction de la forme :

$$f(x) = \frac{1}{2} \sum_{j=1}^m r_j^2(x) \quad (\text{B.15})$$

Dans le cas de la rétropropagation, f est remplacée par l'indice de performance $V = \frac{1}{2} \sum_{q=1}^Q \underline{e}_q^T \underline{e}_q$, r par l'erreur e et x par $\underline{x} = [w_{11}^{(1)} \quad w_{12}^{(1)} \quad \dots \quad w_{n_1 n_0}^{(1)} \quad w_{10}^{(1)} \quad \dots \quad w_{n_1 0}^{(1)} \quad w_{11}^{(2)} \quad \dots \quad w_{n_M 0}^{(M)}]^T$.

La matrice jacobienne J s'écrit alors sous la forme :

$$J = \begin{bmatrix} \frac{\partial e_1(\underline{x})}{\partial x_1} & \frac{\partial e_1(\underline{x})}{\partial x_2} & \dots & \frac{\partial e_1(\underline{x})}{\partial x_n} \\ \frac{\partial e_2(\underline{x})}{\partial x_1} & \frac{\partial e_2(\underline{x})}{\partial x_2} & \dots & \frac{\partial e_2(\underline{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_N(\underline{x})}{\partial x_1} & \frac{\partial e_N(\underline{x})}{\partial x_2} & \dots & \frac{\partial e_N(\underline{x})}{\partial x_n} \end{bmatrix} \quad (\text{B.16})$$

Où N est le nombre total d'erreurs $N = Q \times n_M$, et n est le nombre total de poids $n = \sum_{l=1}^M n_l (n_{l-1} + 1)$.

Les éléments de la matrice jacobienne sont sous la forme $\frac{\partial e_q(m)}{w_{ij}^{(l)}}$ avec $m = 1, \dots, n_M$ et $q = 1, \dots, Q$. Afin de pouvoir les calculer, on procède à l'algorithme de rétropropagation standard avec cette modification dans la dernière couche :

$$\Delta^{(M)} = -\dot{F}^M(\underline{y}^{(M)}) \quad (\text{B.17})$$

Ceci revient à dire que chaque colonne de la matrice (B.17) est un vecteur de sensibilité qui doit être rétropropagé à travers le réseau afin de fournir une ligne de la matrice jacobienne.

L'algorithme LM procède comme suit :

1– On présente toutes les entrées au réseau et on calcule toutes les sorties correspondantes (en utilisant (B.1) et (B.2)), et les erreurs $\underline{e}_q = \underline{t}_q - \underline{z}_q^{(M)}$. On calcule la somme des carrés des erreurs $V(\underline{x})$.

- 2– On calcule la matrice jacobienne en utilisant (B.17), (B.10), (B.8), (B.9) et (B.16).
- 3– On résout (B.14) afin d'obtenir $\Delta \underline{x}$. Une méthode de pseudo-inverse peut être utilisée.
- 4– On recalcule la somme des carrés des erreurs en utilisant $\underline{x} + \Delta \underline{x}$. Si cette nouvelle somme est inférieure à celle calculée à l'étape 1, alors on réduit λ par un facteur β (10 par exemple), on accepte l'itération $\underline{x} = \underline{x} + \Delta \underline{x}$, et on revient à l'étape 1. Si la nouvelle somme des carrés n'est pas réduite, on augmente la valeur de λ par le même facteur β et on revient à l'étape 3.
- 5– Cet algorithme est censé converger lorsque la norme du gradient $\nabla V(\underline{x}) = J(\underline{x})^T e(\underline{x})$ devient inférieure à une valeur prédéterminée, ou lorsque la somme des erreurs atteint un objectif fixé à l'avance [144].

ANNEXE C La recherche par motif

La recherche par motif est une autre méthode d'optimisation numérique. Elle servira dans cette thèse à optimiser les paramètres de la machine à vecteurs de support, utilisée pour la prévision éolienne.

On considère le problème de minimisation d'une fonction dérivable continuellement $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

On note $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$ et $\mathbb{N}$ les ensembles des nombres réels, rationnels, entiers et naturels respectivement. Toutes les normes sont euclidiennes. On définit $L(y) = \{x : f(x) \leq f(y)\}$, $C(y) = \{x : f(x) = f(y)\}$ et $X_* = \{\nabla f(x) = 0\}$.

Le motif

Afin de définir un motif, on a besoin de deux composants : une matrice de base et une matrice génératrice. La matrice de base peut être n'importe quelle matrice non-singulière $B \in \mathbb{R}^{n \times n}$. La matrice génératrice est une matrice $C_k \in \mathbb{Z}^{n \times p}$, où $p > 2n$. On partitionne la matrice génératrice en composantes :

$$C_k = [M_k \quad -M_k \quad L_k] = [\Gamma_k \quad L_k] \quad (\text{C.1})$$

Il faut que $M_k \in M \subset \mathbb{Z}^{n \times n}$, où M est un ensemble fini de matrice non-singulières, que $L_k \in \mathbb{Z}^{n \times (p-2n)}$ et que L_k contienne au moins une colonne, la colonne des zéros.

Un motif P_k est alors défini par les colonnes de la matrice $P_k = BC_k$. Pour plus de commodité, on utilise la partition de la matrice génératrice C_k donnée par (C.1) pour partitionner P_k comme suit :

$$P_k = BC_k = [BM_k \quad -BM_k \quad BL_k] = [B\Gamma_k \quad BL_k] \quad (\text{C.2})$$

Étant donné $\Delta_k \in \mathbb{R}$, $\Delta_k > 0$, on définit un pas d'essai s_k^i comme un vecteur de la forme :

$$s_k^i = \Delta_k BC_k^i \quad (\text{C.3})$$

Où les c_k^i représentent une colonne de $C_k = [c_k^1 \dots c_k^p]$. Notons que BC_k^i détermine la direction du pas, alors que Δ_k sert comme paramètre de longueur du pas. À l'itération k , on définit un point d'essai comme n'importe quel point de la forme $x_k^i = x_k + s_k^i$, où x_k est l'itération actuelle [145].

Les mouvements exploratoires

Les méthodes de recherche de motif procèdent par une série de mouvement exploratoires autour de

l'itération actuelle avant de déclarer une nouvelle itération et mettre à jour l'information associée. Ces mouvements peuvent être vus comme un échantillonnage de la fonction autour de l'itération courante x_k de manière déterministe bien définie, afin de rechercher une nouvelle itération $x_{k+1} = x_k + s_k$ pour laquelle la fonction objectif décroît. Les méthodes de recherche de motif individuelles sont distinguables, en partie, par la manière selon laquelle ces mouvements sont conduits. Afin d'assurer un large choix de mouvements exploratoires, et assurer en même temps la convergence de la méthode, deux conditions sont requises :

$$s_k \in \Delta_k P_k \equiv \Delta_k BC_k \equiv \Delta_k [B\Gamma_k \quad BL_k] \quad (\text{C.4})$$

$$\text{Si } \min\{f(x_k + y), y \in \Delta_k B\Gamma_k\} < f(x_k), \text{ alors } f(x_k + s_k) < f(x_k) \quad (\text{C.5})$$

(La notation $y \in A$ veut dire que le vecteur y est contenu dans l'ensemble des colonnes de la matrice A).

Le choix des mouvements exploratoires doit assurer deux choses :

1. La direction de n'importe quel pas s_k accepté à l'itération k est défini par le motif P_k et sa longueur est déterminé par Δ_k .
2. Si la simple décroissance de la valeur de la fonction objectif à l'itération courante peut être trouvée parmi n'importe quels $2n$ pas d'essai définis $\Delta_k B\Gamma_k$, alors les mouvements exploratoires doivent produire un pas s_k qui fournit aussi une simple décroissance de la valeur de la fonction à l'itération courante. En particulier, $f(x_k + s_k)$ n'a pas besoin d'être inférieure ou égale à $\min\{f(x_k + y), y \in \Delta_k B\Gamma_k\}$.

Donc, un algorithme légitime de mouvements exploratoires est celui qui devine lequel des pas définis par $\Delta_k P_k$ produira la simple décroissance et qui évalue alors la fonction uniquement suite à un pas pareil (et ce pas peut être contenu dans $\Delta_k BL_k$ au lieu de $\Delta_k B\Gamma_k$). À l'autre extrême, cet algorithme serait celui qui évalue tous les pas p définis par $\Delta_k P_k$ et renvoie le pas qui a produit la valeur la plus faible de la fonction objectif.

Ce sont les propriétés des mouvements exploratoires qui permettent de prouver que :

$$\lim_{k \rightarrow +\infty} \inf \|\nabla f(x_k)\| = 0 \quad (\text{C.6})$$

Même si on exige uniquement la simple diminution de f . Pour obtenir carrément $\lim_{k \rightarrow +\infty} \|\nabla f(x_k)\| = 0$, il faut placer des hypothèses plus fortes sur les mouvements exploratoires [145].

La recherche de motif généralisée

L'algorithme 1 spécifie la méthode de recherche de motif généralisée pour une minimisation non contrainte.

Algorithme 1 :

Soit $x_0 \in \mathbb{R}^n$ et soit $\Delta_0 > 0$

Pour $k = 0, 1, \dots$

- a) Calculer $f(x_k)$
- b) Déterminer un pas s_k en utilisant l'algorithme des mouvements exploratoires.
- c) Calculer $\rho_k = f(x_k) - f(x_k - s_k)$
- d) Si $\rho_k > 0$ alors $x_{k+1} = x_k + s_k$. Sinon $x_{k+1} = x_k$
- e) Mettre à jour C_k et Δ_k

Afin de définir une méthode particulière, il est nécessaire de spécifier la matrice de base B , la matrice génératrice C_k , les mouvements exploratoires à utiliser pour produire un pas s_k , et les algorithmes pour mettre à jour C_k et Δ_k [145].

Les mises à jour

L'algorithme 2 spécifie les exigences pour mettre à jour Δ_k . Le but de la mise à jour est de forcer $\rho_k > 0$. Une itération avec $\rho_k > 0$ est réussie, sinon elle est ratée. On rappelle qu'un pas est accepté s'il produit une décroissance simple de la fonction objectif f .

Algorithme 2 :

Soit $\tau \in \mathbb{Q}$, on note $\theta = \tau^{\omega_0}$ et $\gamma_k \in \Lambda = \{\tau^{\omega_1}, \dots, \tau^{\omega_L}\}$, où $\tau > 1$ et $\{\omega_0, \dots, \omega_L\} \subset \mathbb{Z}$, $L \equiv |\Lambda| < +\infty$, $\omega_0 < 0$, et $\omega_i \geq 0, i = 1, \dots, L$.

- a) Si $\rho_k \leq 0$ alors $\Delta_{k+1} = \theta \Delta_k$
- b) Si $\rho_k > 0$ alors $\Delta_{k+1} = \lambda_k \Delta_k$

Les conditions en θ et λ assurent que $0 < \theta < 1$ et $\lambda_i \geq 1$ pour tout $\lambda_i \in \Lambda$. Donc, si une itération est réussie, il est possible d'augmenter le paramètre de longueur du pas Δ_k , mais ce dernier n'a pas le droit d'augmenter.

L'algorithme de mise à jour de C_k dépend de la méthode de recherche de motif utilisée. Il est suffisant de choisir les colonnes de C_k de sorte à satisfaire (C.1) et les conditions établies sur les matrices $M_k \in M \subset \mathbb{Z}^{n \times n}$ et $L_k \in \mathbb{Z}^{n \times (p-2n)}$ [145].