

DÉTECTION EN TEMPS RÉEL DE L'INSTABILITÉ
TRANSITOIRE DES RÉSEAUX ÉLECTRIQUES AVEC
LES MESURES SYNCHRONISÉES DE PHASEURS

PAR

Hussein SUPRÊME

THÈSE PRÉSENTÉE À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DU
DOCTORAT EN GÉNIE
Ph. D.

MONTRÉAL, LE 26 AVRIL 2017

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

©Tous droits réservés

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

PRÉSENTATION DU JURY

CETTE THÈSE A ÉTÉ ÉVALUÉE

PAR UN JURY COMPOSÉ DE :

M. Louis-A. Dessaint, directeur de thèse
Département de génie électrique à l'École de technologie supérieure

M. Tony Wong, président du jury
Département de génie de la production automatisée à l'École de technologie supérieure

M. Pierre-Jean Lagacé, membre du jury
Département de génie électrique à l'École de technologie supérieure

Mme Annissa Henniche-Oussédik, membre du jury
Institut de recherche d'Hydro-Québec (IREQ)

M. Serge Lefebvre, examinateur externe indépendant
Institut de recherche d'Hydro-Québec (IREQ)

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

LE 19 AVRIL 2017

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens à remercier Dieu de m'avoir guidé et éclairé tout au long de ce projet. Sa bénédiction m'a toujours soutenu dans mes moments de remise en question les plus difficiles.

Je tiens à témoigner mes profondes et sincères reconnaissances envers le Professeur Louis-A. Dessaint, mon directeur de thèse, pour l'honneur qu'il m'a fait en dirigeant et en supervisant mes travaux de maîtrise et mes études doctorales. Il m'a ouvert bien grand les portes de son laboratoire de recherche et a toujours mis ses connaissances, ses savoirs, son savoir-faire, et son savoir-être à ma disposition. Grâce à lui, j'ai pu améliorer mon bagage théorique. Sa rigueur scientifique a su m'aider à prendre le recul nécessaire pour la réussite de ce travail. J'en profite également pour remercier le Groupe de recherche en électronique de puissance et commande industrielle (GRÉPCI) de m'avoir octroyé une bourse d'études. À tous les collègues que j'ai rencontrés dans ce laboratoire, je leur dis merci pour leurs différents commentaires et conseils.

Je souhaite aussi exprimer ma gratitude à M. Innocent Kamwa, Professeur associé aux Universités Mc-Gill et Laval et Chef expertise Réseaux électriques et mathématiques de l'Institut de Recherche d'Hydro-Québec (IREQ). Il m'a offert la chance inouïe de réaliser un stage doctoral dans les murs du prestigieux centre de recherche mondial qu'est l'IREQ. Malgré son horaire de temps chargé, il a toujours été disponible et disposé pour me guider dans mes travaux. Ses riches expériences m'ont permis d'avoir une vue globale et des réflexions profondes sur la problématique de mon sujet de recherche. J'ai également eu la possibilité de rencontrer des chercheurs dévoués qui ont su m'inspirer et me conseiller. Je remercie particulièrement Mme Annissa Henniche-Oussédik. Ses conseils m'ont été précieux et tout le temps qu'elle m'a accordé pour peaufiner mes travaux n'a pas de prix.

Finalement, je formule mes plus vifs remerciements envers mes parents Jacques André et Jeanne Chantale, ma conjointe Geneviève, mes frères Stéphane et Krezky, mon oncle Jean Khany et tous les autres membres de ma famille qui m'ont toujours soutenu ou glissé un mot

d'encouragement. Vous avez tous cru en moi et m'avez stimulé au jour le jour à donner le meilleur de moi-même. Aussi, mille mercis à tous mes amis et à toutes mes amies qui n'ont jamais cessé de m'appuyer et de me supporter.

Comme je l'ai déjà dit à chacun d'entre vous, ce travail constitue non seulement une réussite personnelle mais aussi une réussite collective pour vous tous qui m'avez épaulé d'une manière ou d'une autre.

DÉTECTION EN TEMPS RÉEL DE L'INSTABILITÉ TRANSITOIRE DES RÉSEAUX ÉLECTRIQUES AVEC LES MESURES SYNCHRONISÉES DE PHASEURS

Hussein SUPRÊME

RÉSUMÉ

Les synchrophaseurs également appelés *Phasor Measurement Units*, couplés d'un système GPS, permettent de mesurer la phase du courant plus de 30 fois par seconde, soit une fréquence 120 fois plus grande que celle des systèmes actuels. Ils donnent la possibilité de cartographier en temps réel les phases du réseau électrique ; ce qui de toute évidence, facilite leur contrôle, leur surveillance, leur compréhension et leur gestion optimale. Les applications des synchrophaseurs incluent le contrôle étendu, la validation du modèle système, la détermination des marges de stabilité, l'optimisation de la charge pour un système stable, la détection d'îlotage, l'enregistrement des perturbations à l'échelle du réseau et la visualisation de la réaction dynamique du réseau. Avec cette technologie en vogue, les opérateurs de réseaux disposent d'une quantité d'informations très importante à manipuler.

La présente thèse a pour objectif de proposer aux opérateurs de réseaux électriques une démarche systématique qui permet de prédire l'instabilité des réseaux électriques en temps réel en extrayant les signatures cachées dans la masse de données fournies par les synchrophaseurs. Étant donné que le forage de données dispose d'un ensemble d'algorithmes conduisant à la construction des modèles, on exploite dans ce travail la forêt aléatoire d'arbres de décision et le boosting. Ces algorithmes vont chercher un maximum de connaissances dissimulées dans les données.

On propose un nouveau concept pour l'étude de l'instabilité : le Centre de Puissance (COP). Cette notion est une extension du centre d'inertie et de l'approche de mesure de zone étendue et de système de contrôle de tension. Elle prend en compte les valeurs de puissances effectives du réseau déduites des mesures des synchrophaseurs. À chaque région du réseau, on associe un centre de puissance de vitesse. On constate une instabilité dans le cas où le COP de vitesse de deux régions quelconques ne cesse de s'éloigner. On a aussi proposé deux indices de détection de l'instabilité se basant sur les COP de vitesse et les mesures des synchrophaseurs. La divergence de ces indices confirme que le réseau a perdu son synchronisme. Contrairement au snapshot largement utilisé dans la littérature, on propose dans cette thèse d'utiliser une fenêtre glissante lors de l'analyse de l'instabilité. Cette dernière a comme avantages de favoriser une prise de décision en tout temps sans un détecteur de défaut au préalable et de mettre à jour la base de données sur l'état du réseau avec l'arrivée de nouvelles données sans perdre les informations contenues dans les données précédentes.

Pour évaluer l'importance des prédicteurs proposés dans la construction des modèles de forêt aléatoire et de boosting, on considère trois scénarios. Dans le premier, on retient comme prédicteurs les variables découlant du COP, les indices proposés et toutes les variables mesurées et estimées. On constate qu'une plus grande importance relative est accordée aux

prédicteurs directement liés au COP et aux mesures synchrophasées dans la construction des modèles de prédiction de l'instabilité en temps réel. Dans le deuxième scénario, l'emphase est mise uniquement sur les prédicteurs proposés et les mesures synchrophasées. Il résulte que les écarts entre l'exactitude, la fiabilité et la sécurité des modèles de prédiction obtenus dans les deux scénarios sont très faibles. De plus, les taux de mauvaise classification et de fausse alarme se rapprochent. On en déduit que les prédicteurs proposés constituent un bon compromis entre la faible perte de précision du modèle et la diminution de la taille de la base de données sur laquelle l'apprentissage est effectué. Le troisième scénario permet de généraliser et de démontrer que les prédicteurs proposés ne dépendent pas du réseau.

Finalement, on propose une nouvelle approche adaptative de mise à jour du modèle de prédiction. L'état du réseau est déterminé par une double vérification : une première se reposant sur le modèle de la forêt aléatoire initialement construit et la seconde se basant sur un boosting. À chaque boosting construit, on attribue un indicateur qui dépend du poids de chacun des arbres du modèle et des erreurs y relatives. Lors de la mise à jour, cet indicateur est comparé au précédent et autorise ou non l'adoption du nouveau modèle de prédiction pour les événements futurs.

Mots clés : Centre de puissance (COP), Synchrophaseurs, Prédiction de l'instabilité, Stabilité transitoire, Stabilité dynamique, Forage de données, Arbre de décision, Boosting, Forêt aléatoire, Évaluation de la sécurité dynamique (DSA).

DÉTECTION EN TEMPS RÉEL DE L'INSTABILITÉ TRANSITOIRE ET DYNAMIQUE DES RÉSEAUX ÉLECTRIQUES AVEC LES MESURES SYNCHRONISÉES DE PHASEURS

Hussein SUPRÊME

ABSTRACT

Synchrophasors also called Phasor Measurement Units, coupled with a GPS system, measure the phase of the current more than 30 times per second, or a frequency that is well over 120 times that of current systems. They give the possibility of mapping in real time the phases of the electrical network which obviously facilitates their control, monitoring, understanding and optimal management. The applications of synchrophasors include extended control, system model validation, determination of stability margins, load optimization for a stable system, islanding detection, network-wide perturbation recording and visualization of the dynamic of the network. With this technology in perpetual development, the network operators have a very important amount of information to manipulate.

The aim of this thesis is to propose to the network operators a systematic approach which allows predicting the instability of the electrical networks in real time by extracting the hidden signatures in the mass of data provided by the synchrophasors. Given that data mining is a set of algorithms leading to construct models from the data by seeking a maximum of concealed knowledge, we exploit in this work the random forest of decision trees and the boosting.

We propose a new concept for the study of instability: the Center of Power (COP). This notion is an extension of the center of inertia and the extended area measurement approach and voltage control system. It considers the actual network power values deduced from the phasor measurements. A COP speed is associated with each area. There is an instability in the case where the speed COP of any two areas keeps moving away. Two indices were also proposed for detecting instability based on speed COPs and phasor measurements. The divergence of these indices confirms that the network has lost its synchronism. Unlike snapshot widely used in the literature, it is proposed in this thesis to use a sliding window during the analysis of instability. The latter has the advantages of promoting decision-making at any time without a fault detector and updating the database on the state of the network with the arrival of new data without losing the information contained in the previous data.

To evaluate the importance of predictors proposed in the construction of the random forest and the boosting models, three scenarios are considered. In the first, the variables derived from the COP, the proposed indices and all the measured and estimated variables are selected as predictors. Increased relative importance is given to predictors directly related to COP and to phasor measurements in the construction of predictive models of real-time instability. So, in the second scenario, the emphasis is placed solely on the proposed predictors and the phasor measurements. It results that the differences between the accuracy, the reliability and the security of the prediction models obtained in both scenarios are very low. In addition, rates of misclassification and false alarm come closer. It can be deduced that the proposed predictors

represent a good compromise between the small loss of precision of the model and the reduction in the size of the database on which the learning is carried out. The third scenario makes it possible to generalize and demonstrate that the proposed predictors do not depend on the network.

Finally, a new adaptive approach is proposed to update the prediction model. The state of the network is determined by a double check: a first one based on the model of the random forest initially constructed and the second based on a boosting. At each boosting constructed, an indicator is assigned which depends on the weight of each of the trees of the model and related errors. During the update, this indicator is compared to the previous one and allows or denies the adoption of the new model for predicting the state of the network for future events.

Keyword: Center of Power (COP), Phasor Measurements Units (PMU), Instability Prediction, Transient Stability, Dynamic Stability, Data mining, Decision Trees, Random Forest, Dynamic security assessment (DSA).

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 CONTEXTE DE L'ÉTUDE	5
1.1 Motivations de la thèse	5
1.2 Problématique de la recherche	6
1.2.1 Objectifs de la thèse	7
1.2.2 Méthodologie	8
1.2.3 Contribution de la thèse à l'avancement des connaissances scientifiques et/ou technologiques	8
1.2.4 Originalité du travail	9
CHAPITRE 2 REVUE DE LITTÉRATURE	11
2.1 Introduction	11
2.2 Méthodes de prédiction de la stabilité transitoire	11
2.2.1 Méthodes temporelles liées aux méthodes d'intégration numériques	12
2.2.2 Méthodes directes	13
2.2.3 Méthodes modernes	17
2.3 Revue critique de la littérature	25
2.4 Conclusion	27
CHAPITRE 3 ÉTUDE DE LA STABILITÉ	29
3.1 Introduction	29
3.2 Définition et classification de la stabilité sur les réseaux électriques	29
3.2.1 Classification de la stabilité des réseaux électriques	30
3.2.2 Relation entre fiabilité, sécurité et stabilité des réseaux électriques	34
3.3 Fondements théoriques de la stabilité des réseaux électriques	35
3.3.1 Définition mathématique de la stabilité d'un système dynamique	36
3.3.2 Stabilité des réseaux électriques	38
3.4 Modèle mathématique lié à l'étude de la stabilité transitoire	39
3.5 Modes d'opération des réseaux électriques	40
3.6 Conclusion	42
CHAPITRE 4 VUE D'ENSEMBLE DES SYNCROPHASEURS	43
4.1 Introduction	43
4.2 Mesure des synchrophaseurs	44
4.3 Réseau de communication et placement des PMUs	46
4.4 Applications des mesures synchrophasées dans les réseaux électriques	49
4.5 PMU vs SCADA	54
4.6 Conclusion	56
CHAPITRE 5 FORAGE DE DONNÉES : FORêt ALÉATOIRE D'ARBRES DE DÉCISION ET BOOSTING	59

5.1	Introduction.....	59
5.2	Techniques inductives versus transductives	60
5.3	Arbres de décision.....	63
5.3.1	Principe des arbres de décision	63
5.3.2	Critère de segmentation	64
5.3.3	Taille des arbres, surapprentissage, élagage	65
5.4	Forage de données : Forêt aléatoire d'arbres de décision, Random Forest (RF).....	66
5.4.1	Principe général du Random Forest	67
5.4.2	Prédiction d'un RF	69
5.4.3	Évaluation de la performance des RF	70
5.5	Forage de données : Boosting.....	75
5.6	Description du logiciel R	78
5.7	Conclusion	81
 CHAPITRE 6 FONDEMENTS THÉORIQUES DE LA MÉTHODOLOGIE PROPOSÉE.....		83
6.1	Introduction.....	83
6.2	Concept du centre de puissance (Center of Power, COP)	85
6.2.1	Principes théoriques	85
6.2.2	Définition des nouveaux indices basés sur le COP et les mesures des synchrophaseurs	90
6.3	Pourquoi utiliser le centre de puissance ?	93
6.4	Description de la nouvelle méthodologie de prédiction en temps réel	96
6.4.1	Méthodologie globale de prédiction	96
6.4.2	Approche de la fenêtre glissante (sliding window) versus snapshot	98
6.4.3	Mise à jour automatique de la base de données et du modèle de prédiction	100
6.5	Conclusion	103
 CHAPITRE 7 RÉSULTATS ET VALIDATION		105
7.1	Introduction.....	105
7.2	Présentation des cas étudiés et formation de la base des données	105
7.3	Comportement des indices proposés.....	107
7.4	Analyse de l'importance et de l'efficacité des prédicteurs	114
7.5	Évaluation de la performance out-of-bag des forêts aléatoires.....	117
7.6	Évaluation de la performance du modèle de prédiction.....	119
7.7	Mise à jour automatique du modèle de prédiction.....	126
7.8	Discussion des résultats	129
 CONCLUSION		133
 RECOMMANDATIONS		137
ANNEXE I	CARACTÉRISTIQUES DES RÉSEAUX ÉLECTRIQUES TEST KUNDUR TWO-AREA ET NEW ENGLAND	139

ANNEXE II	CARACTÉRISTIQUES DU RÉSEAU ÉLECTRIQUE SIMPLIFIED 14-GENERATORS AUSTRALIAN TEST SYSTEM	149
ANNEXE III	EXEMPLES D'APPLICATIONS DES ARBRES DE DÉCISION ET DE LA FORêt ALÉATOIRE D'ARBRE DE DÉCISION	163
BIBLIOGRAPHIE		227

LISTE DES TABLEAUX

	Page
Tableau 4.1	Formulation des contraintes de placement optimal des PMUs.....
Tableau 4.2	Comparaison entre EMS/SCADA et PMU.....
Tableau 5.1	Matrice de confusion.....
Tableau 5.2	Packages dans R correspondant à des techniques de forages
Tableau 7.1	Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau Australien simplifié
Tableau 7.2	Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau New England
Tableau 7.3	Comparaison des résultats de performance des modèles issus de RF sur une fenêtre glissante avec des variables COP et COI comme prédicteurs – Réseaux Australien simplifié et New-England
Tableau 7.4	Suivi de la mise à jour de la base de données en fonction de l'indicateur Ω
Tableau 7.5	Temps d'exécution en minutes pour l'apprentissage du modèle

LISTE DES FIGURES

	Page
Figure 2.1	Familles de méthodes pour la prédition de la stabilité transitoire.....
Figure 2.2	12
Figure 2.2	Caractéristiques $P = f(\delta)$ avant, pendant et après défaut
Figure 3.1	15
Figure 3.1	Classification des phénomènes de stabilité sur les réseaux électriques
Figure 3.2	31
Figure 3.2	Illustration de la définition de la stabilité
Figure 3.3	37
Figure 3.3	Illustration de la définition de la stabilité asymptotique.....
Figure 3.4	38
Figure 3.4	Modèle simplifié d'une machine synchrone.....
Figure 3.5	39
Figure 3.5	Diagramme d'état de la sécurité du réseau électrique de DyLiacco.....
Figure 4.1	41
Figure 4.1	Convention de la représentation d'un synchrophaseur
Figure 4.2	45
Figure 4.2	Diagramme conceptuel d'un système de mesures de phaseurs synchronisées
Figure 4.3	46
Figure 4.3	Applications des PMUs dans les réseaux électriques
Figure 4.4	50
Figure 4.4	Représentation de la structure d'un EMS/SCADA et d'un WAMS
Figure 5.1	55
Figure 5.1	Méthodes inductives et transductives
Figure 5.2	61
Figure 5.2	Arbre de décision
Figure 5.3	64
Figure 5.3	Interprétation de l'out-of-bag dans le fonctionnement de la forêt aléatoire
Figure 5.4	70
Figure 5.4	Importance des variables.....
Figure 6.1	73
Figure 6.1	Représentation d'un réseau avec n régions et m générateurs
Figure 6.2	86
Figure 6.2	Variation du COP de vitesse pour chaque région du réseau test de Kundur (cas stable versus instable).....
Figure 6.3	89
Figure 6.3	Comportement des indices γCOP et κCOP (cas stables versus instables) – Réseau Test Kundur
	91

Figure 6.4	Comportement des angles du rotor des générateurs – Réseau Test Kundur	92
Figure 6.5	Description des séquences de calcul et des scénarios pour la prédiction de l'instabilité	98
Figure 6.6	Comportement de κCOP avec une fenêtre glissante pour des cas stables et instables Réseau Test Kundur	100
Figure 6.7	Dispositif de mise-à-jour périodique du modèle de prédiction de l'instabilité	101
Figure 7.1	Cadre d'apprentissage pour l'évaluation de la sécurité du réseau	106
Figure 7.2	Comparaison entre les valeurs réelles et celles d'une fenêtre glissante pour des cas stables et instable – Réseau Test New England	108
Figure 7.3	Comportement de γCOP et κCOP 6 s après l'effacement du défaut – Réseau Test New England	109
Figure 7.4	Comportement des COP de vitesse versus les angles du rotor des générateurs pour un cas instable – Réseau Test New England	110
Figure 7.5	Comportement des variables pour un cas instable sans (gauche) et avec (droite) des mesures correctives – Réseau New England	111
Figure 7.6	Courbe représentative des indicateurs d'instabilité pour une instabilité rapide sans (gauche) et avec (droite) des actions correctives – Réseau New England	112
Figure 7.7	Comportement de γCOP et κCOP 5 s après l'effacement du défaut – Réseau Simplifié Australien	113
Figure 7.8	Vue globale sur les indicateurs de détection de l'instabilité – Réseau Simplifié Australien	114
Figure 7.9	Importance des variables dans la construction de la forêt aléatoire pour le Scénario 1 – Réseau New England (droite) et Réseau Simplifié Australien (gauche)	115
Figure 7.10	Importance des variables dans la construction de la forêt aléatoire pour le Scénario 2 – Réseau New England (droite) et Réseau Simplifié Australien (gauche)	116
Figure 7.11	Importance des variables dans la construction de la forêt aléatoire pour le Scénario 3	117

Figure 7.12	Caractéristiques de convergence de l'OOB pour l'apprentissage du RF avec la fenêtre glissante pour les Scénarios 1 et 2 – Réseau New England (bas) et Réseau Simplifié Australien (haut).....	118
Figure 7.13	Caractéristiques de convergence de l'OOB pour l'apprentissage du RF avec la fenêtre glissante pour le Scénario 3	119
Figure 7.14	Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau Australien simplifié	121
Figure 7.15	Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau New England	123
Figure 7.16	Comparaison des résultats de performance des modèles issus de RF pour le Scénario 3	125
Figure 7.17	Comparaison des résultats de performance des modèles issus du Boosting proposé et du RF reconstruit suite à une mise à jour de la base de données – Scénario 3.....	127
Figure 7.18	Temps de calcul du modèle de prédiction lors de la mise à jour – Boosting proposé vs RF reconstruit.....	128

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

Adaboost	Adaptive Boosting
ANN	Artificial Neural Networks
Arcing	Adaptive Resampling and Combining
AVR	Automatic Voltage Regulator
BCU	Boundary of Stability Region Controlling
BPA	Bonneville Power Administration
CAE	Critère des aires égales
CART	Classification and Regression Tree
CCT	Critical clearing time
CHAID	CHi-squared Automatic Interaction Detector
COI	Center of Inertia
COP	Center of Power
CRAN	Comprehensive R Archive Network
CT	Clearing time
DOE	Department of Energy of United States of America
DP	Dot Product
DSA	Dynamic Security Assessment
DT	Decision Tree
EDF	Électricité de France
EEAC	Extended Equal Area Criterion
EMS	Energy Management System
EPRI	Electric Power Research Institute

FCMdd	Fuzzy C-medoid
FHRCNN	Fuzzy Hyper-Rectangular Composite Neural Network
GPS	Global Positioning System
GUI	Graphical User Interface
HVDC	High-Voltage Direct Current
i.e.	c'est-à-dire
IEEE	Institute of Electrical and Electronics Engineers
INLP	Integer NonLinear Programming
IP	Integer Programming
IQP	Integer Quadratic Programming
KNN	k -Nearest Neighbor
Matlab	Matrix Laboratory
MDA	Mean Decrease in Accuracy
MDG	Mean Decrease in Gini
MLP	Perceptron multicouche
n.d.a.	Notes de l'auteur
NASPI	North American SynchroPhasor Initiative
NERC	North American Electric Reliability Corporation
NRC	Natural Resources Canada
NRC	National Research Council Canada
OMIB	One machine infinite bus
OOB	Out-of-Bag
OPGW	Optical Ground Wire

OPP	Optimal PMU Placement
PDC	Phasor data concentrator
PDC	Phasor data concentrator
PEBS	Potential Energy Boundary Surface
PIDAS	Parameter Identification Data Acquisition System
PMU	Phasor Measurement Units
PSS	Power System Stabilizers
PST	Power System Toolbox
RF	Random Forest
RK4	Runge-Kutta ordre 4
SCADA	Supervisory Control and Data Acquisition
SEP	Stable Equilibrium Point
SimPowerSystems™	Simscape Power Systems™
SIPS	System Integrity Protection Schemes
SPS	Special Protection Systems
SPSS	Statistical Package for the Social Sciences
SVC	Static-Var Compensator
SVM	Support Vector Machine
SW	Sliding Window
TD	Time domain
TEF	Transient Energy Function
TS	Tabu Search
TSC-OPF	Transient Stability-Constrained Optimal Power Flow

UEP	Closest Unstable Equilibrium Point
UTC	Coordinated Universal Time
WAMS	Wide Area Measurement System
WSCC	Western Systems Coordinating Council

LISTE DES SYMBOLES ET UNITÉS DE MESURE

UNITÉS DE MESURE

A	Ampère
GHz	giga Hertz (10^6 Hz)
Go	giga Octet (10^6 octet)
Hz	Hertz
J	Joule
kV	kilo Volt (10^3 Volt)
m	mètre (unité de longueur)
ms	milliseconde (10^{-3} seconde)
N	Newton
N.m	Newton-mètre
rad	radian
rad/s	radian par seconde
s	seconde
V	Volt
VA	Volt-Ampère
W	Watt
Ω	Ohm

SYMBOLES

B_{ij}	Partie réelle de l'élément i-j de la matrice d'admittance
E'_d	Tension interne du générateur axe d
E_{fd}	Tension d'excitation
E'_q	Tension interne du générateur axe q
G_{ij}	Partie imaginaire de l'élément i-j de la matrice d'admittance
I_d	Courant du générateur axe d
I_q	Courant du générateur axe q
K_A	Gain de l'amplificateur
K_E	Gain de l'excitatrice
K_F	Gain du stabilisateur
M_C	Machines critiques
M_N	Machines non critiques
P_D	Puissance active demandée
P_G	Puissance active générée
P_a	Puissance d'accélération
P_e	Puissance électrique
P_m	Puissance mécanique
Q_D	Puissance réactive demandée

Q_G	Puissance réactive générée
R_f	Signal de retour de l'information
R_s	Résistance du stator
S_{EK1}	Première constante de la fonction de saturation
S_{EK2}	Deuxième constante de la fonction de saturation
T_A	Constante de temps de l'amplificateur
T_E	Constante de temps de l'excitatrice
T_F	Constante de temps du stabilisateur
T_a	Couple d'accélération
T'_{do}	Constante de temps axe d
T_e	Couple électrique
T_m	Couple mécanique
T'_{qo}	Constante de temps axe q
V_R	Sortie du régulateur de tension
V_i	Module de la tension à la barre i
Y_{bus}	Matrice d'admittance
Y_{ij}	Module de l'admittance i-j
$t_{cleared}$	Temps effacement de défaut
t_r	Temps de retour
t_u	Temps d'instabilité
x_d	Réactance du générateur axe d
x'_d	Réactance transitoire axe d
x_q	Réactance du générateur axe q
x'_q	Réactance transitoire axe q
δ_r	Angle de retour
δ_u	Angle d'instabilité
θ_i	Argument de la tension à la barre i
H	Constante d'inertie de la machine
V	Vecteur de tension
D	Constante d'amortissement
I	Vecteur de courant
J	Moment d'inertie
δ	Angle du générateur
η	Marge de stabilité
τ	Rapport de transformation
ω	Vitesse angulaire

INTRODUCTION

Le centre de contrôle, de surveillance, et d'acquisition de données est couramment utilisé par les opérateurs de réseaux électriques pour surveiller, contrôler et optimiser la production et la performance du réseau tout en assurant la sécurité des opérations. Durant ces dernières années, on observe que la demande en électricité ne fait qu'accroître à cause notamment de l'industrialisation, de l'urbanisation et de la croissance démographique. De plus, avec la libéralisation des marchés de l'électricité, on assiste à une forte augmentation des flux d'échanges transfrontaliers et à une pénétration des sources d'énergie renouvelables telles que les cellules photovoltaïques et éoliennes. Tous ces facteurs engendrent l'expansion considérable et rapide de la taille et de la complexité des réseaux. L'intégration de ces nombreux équipements de natures diverses soulève la problématique de la stabilité.

Par ailleurs, les réseaux électriques sont fréquemment soumis à des contingences pouvant causer un dysfonctionnement partiel ou total. Ces contingences sont de nature, très diversifiées : pannes de ligne, variations brusques de la charge, défauts de court-circuit, foudre etc... Elles peuvent entraîner, selon leur gravité, la perte de synchronisme du réseau, des dommages sur les différents équipements de protection et de transport... Ces situations néfastes augmentent le risque de pannes en cascade, ce qui peut affecter la fiabilité et la sécurité du réseau. À titre d'exemple récent, on se rappelle de la panne du 31 août 2014 en Tunisie due à la chute d'un câble de garde à fibres optiques (*Optical Ground Wire*, OPGW) après sa coupure par un coup de foudre ; et de celle au Japon, le 11 mars 2011, causée par le séisme de la côte Pacifique du Tōhoku de magnitude 9 sur l'échelle de Richter (Norio et al., 2011). Par conséquent, les fournisseurs d'électricité sont exposés à faire face à des coûts financiers énormes ; lesquels coûts vont avoir une répercussion directe sur les services publics et la société. Compte tenu uniquement des paramètres précités, il s'ensuit que la gestion des réseaux électriques devient de plus en plus complexe et que la prédition de la stabilité en temps réel est un enjeu primordial.

En effet, les opérateurs de réseaux électriques doivent utiliser des outils très performants pour prédire l'instabilité afin d'assurer la continuité du service, la sécurité et la fiabilité du réseau. De nos jours, avec l'émergence de la technologie des synchrophaseurs, les unités de mesure de phaseurs situées à des postes stratégiques sur le réseau peuvent fournir en temps réel des phaseurs de tension et de courant synchronisés. Grâce à la qualité supérieure de ces mesures, en l'occurrence leur précision et leur latence, elles permettent d'avoir une vue globale du réseau en temps réel. Ces données peuvent être intégrées dans un environnement opérationnel et exploitées afin de prédire rapidement si le réseau tend à devenir instable ou non.

L'objectif principal de l'étude de la prédiction de la stabilité transitoire est de surveiller le comportement du réseau suite à une contingence, la connaissance du temps critique de défaut n'est pas nécessaire (Liu et al., 1999a). De récentes recherches en lien avec l'intelligence artificielle ont démontré que des techniques comme les arbres de décision (Rovnyak et al., 1994b), les réseaux de neurones artificiels (Bahbah et Girgis, 2004), la logique floue (Kamwa, Grondin et Loud, 2001b), les apprentissages statistiques (Kamwa et Grondin, 2002b) et les machines à vecteur de support (Moulin et al., 2004) sont des approches prometteuses pour l'analyse de la stabilité transitoire en temps réel. En outre, la forêt aléatoire d'arbres de décision a été recommandée pour résoudre les problèmes difficiles de régression et de classification basés sur les données des synchrophaseurs pour l'évaluation de la sécurité dynamique (*Dynamic Stability Assessment, DSA*) (Suprême et al., 2016) – (Kaci et al., 2014b) – (Kamwa, Samantaray et Joos, 2012). Parmi les grandes méthodes d'apprentissage, (Hastie, Tibshirani et Friedman, 2003) a démontré que la forêt aléatoire est celle qui permet d'avoir le meilleur compromis entre la performance et le temps d'exécution. Sur un éventail de dix grandes méthodes d'apprentissage supervisées, un test réalisé par (Caruana et Niculescu-Mizil, 2006) sur des problèmes de classification et de régression place en deuxième position les forêts aléatoires pour leur performance pendant qu'elles fournissent une meilleure fiabilité de prédiction pour les cas étudiés.

Dans cette thèse, en exploitant les données fournies par les synchrophaseurs, on propose une nouvelle méthodologie basée sur les forêts aléatoires d'arbres de décision et le boosting pour

prédir l'instabilité des réseaux électriques. On introduit le nouveau concept de centre de puissance, auquel s'ajoutent deux indices de détection d'instabilité. Ces nouveaux indices très performants et faciles à calculer, servent de prédicteurs dans la construction des arbres de décision du boosting et de la forêt aléatoire. On propose également pour la prédition de l'instabilité, un algorithme basé sur une fenêtre glissante.

Ce manuscrit comporte sept chapitres, une conclusion, des recommandations, trois annexes et une bibliographie. Le premier chapitre présente le contexte de l'étude. On expose les motivations qui ont mené au choix du sujet de cette thèse. On aborde la problématique de recherche en mentionnant les objectifs de la thèse, la méthodologie utilisée et la contribution apportée. On fait ressortir également l'originalité de ce présent travail. Au deuxième chapitre, une revue de littérature assez exhaustive permet de cerner les différentes méthodes de prédition de la stabilité transitoire. On procède par la suite à une revue critique de la littérature. Le troisième chapitre est consacré à l'étude de la stabilité transitoire. On rapporte les définitions des termes utilisés et les fondements théoriques de la stabilité des réseaux électriques. Le modèle mathématique lié à l'étude de la stabilité transitoire est aussi vu. Une vue d'ensemble sur les synchrophaseurs est débattue dans le quatrième chapitre. On précise ce que ces infrastructures requièrent pour leur bon fonctionnement et on énonce plusieurs applications des mesures synchrophasées dans les réseaux électriques. Le sixième chapitre est réservé au forage des données. Le principe de fonctionnement des arbres de décision est détaillé ainsi que celui de la forêt aléatoire des arbres de décision et du boosting. On décrit aussi le logiciel **R** utilisé pour les différentes simulations. Le sixième chapitre concerne les fondements théoriques de la méthodologie proposée. On définit le concept de centre de puissance proposé pour l'étude de l'instabilité transitoire. On présente par la suite les deux indices servant à prédire l'instabilité. On fait ressortir les raisons qui justifient la bonne performance de ce nouveau concept. L'algorithme proposé est également décrit. Le septième chapitre est une présentation des différents résultats et de la validation de ceux-ci. En fin de chapitre, on procède à une discussion des résultats obtenus.

La conclusion fait une récapitulation notamment sur le concept de centre de puissance, la méthodologie proposée et les principaux résultats obtenus. Dans les recommandations, on met l'accent sur les perspectives liées à nos travaux de recherche. On examine différentes retombées que cette présente thèse peut avoir sur l'avancement des connaissances scientifiques et/ou technologiques.

CHAPITRE 1

CONTEXTE DE L'ÉTUDE

1.1 Motivations de la thèse

Les phénomènes de stabilité transitoire constituent un problème très ancien et fort débattu dans la littérature. Leur étude reste et demeure pourtant un sujet d'actualité et tout à fait pertinent à cause de l'évolution rapide des grands réseaux électriques et de l'apparition de nouvelles technologies de mesure, de protection et de contrôle. À preuve, les deux grandes pannes majeures du 30 et du 31 juillet 2012 ayant survenu au Nord et à l'Est de l'Inde ont affecté plus de 300 millions de personnes (Rampurkar et al., 2016). De plus, les coûts des dommages causés par la fameuse panne d'août 2003 qui a touché le nord-est américain et la côte sud canadienne sont estimés à plus de 6 billions de dollars américains en perte de revenu uniquement (Ray, 2003). Les impacts de ces perturbations étant désastreux sur l'environnement, la société et l'économie, il est d'une priorité absolue de fournir aux opérateurs de réseaux électriques les outils des plus performants pour prédire de tels événements en temps réel.

L'émergence des unités de mesure de phaseurs (*Phasor Measurement Units*, PMU) dans la surveillance en temps réel des conditions de fonctionnement du réseau porte à croire que ces dernières représentent une solution intéressante et viable pour assurer la fiabilité du réseau. En effet, plusieurs entreprises publiques de renommée à travers le monde ont manifesté ces dernières années leur intérêt grandissant pour ce type de technologies (Mauch et Foss, 2005). Entre autres, on peut citer Hydro-Québec (Kamwa et al., 2006), *American Electric Power* (Nuqui et al., 2001), Électricité de France (EDF) (De Boeck et al., 2016) et de nombreux fournisseurs du *Western Systems Coordinating Council* (WSCC) dont *Bonneville Power Administration* (BPA) (Huang et al., 2013) et *Southern California Edison Co* (Bhargava, 1999) – (Bhargava, Salazar et Unit, 2007). Plusieurs auteurs se sont intéressés à diverses facettes de l'analyse de la stabilité transitoire à partir des données provenant des PMUs allant de la prédiction en temps réel en s'intéressant à l'évolution de l'angle du rotor (Dasgupta et al.,

2015) – (Liu et al., 1999b) aux placements des PMUs (Tao et al., 2004). Même dans l'éventualité que les PMUs soient présents dans le futur à chaque jeu de barre en temps réel, la surveillance de l'état dynamique n'est seulement qu'une étape dans la compréhension des propriétés de stabilité dynamique du réseau. Dans le contexte de la stabilité de l'angle du rotor, on souhaite être en mesure de prédire l'état du réseau en temps réel suite à des perturbations bénignes ou graves. On pourrait ainsi déduire si oui ou non un générateur et/ou un groupe de générateurs peuvent potentiellement conduire le réseau à être instable.

1.2 Problématique de la recherche

En se reportant à la panne de 2003 évoquée précédemment, on a constaté que la perte éventuelle des générateurs et des lignes de transmission conduit rapidement à la perte de synchronisme entre les différentes régions du réseau électrique. Ce rapport (Ray, 2003) montre clairement que la perte de synchronisme s'est produite en quelques secondes, alors que les baisses de tension qui ont précédé ont duré quelques minutes. Une attention particulière doit être accordée à la stabilité transitoire car il s'agit d'un événement qui survient vite et qui a des répercussions désastreuses sur le réseau. De par leur nature, les réseaux électriques sont des systèmes dynamiques non linéaires de grande dimension. Il en découle évidemment que la correspondance entre les paramètres d'état et la stabilité transitoire est une relation non linéaire complexe.

De nos jours, les avancées de l'informatique ont favorisé un gain considérable dans le temps de calcul et de simulation. On est en mesure de traiter un grand flux de données et de trouver des informations significatives cachées dans les bases de données. Ces informations sont souvent codées dans des modèles qui peuvent être utilisés pour faire des prédictions. Les synchrophaseurs fournissent toute une panoplie de données informant sur l'état actuel du réseau. Les opérateurs et les gestionnaires de réseau sentent la nécessité de les exploiter au maximum. Le défi majeur est d'élaborer de nouvelles méthodes qui explorent les données en profondeur et proposent des outils pratiques d'analyse des données afin de mieux planifier la surveillance du réseau et d'assurer sa fiabilité et sa sécurité maximales.

En raison de toutes les incertitudes impliquées dans les opérations quotidiennes dans un centre de contrôle, le développement d'outils de surveillance de la stabilité de l'angle du rotor et de détection de l'instabilité en temps réel devient une tâche incontournable. Par conséquent, la problématique de recherche se formule autour de ces trois questions :

1. Peut-on établir de nouveaux critères de détection de l'instabilité ?
2. Comment peut-on extraire des connaissances de la mine d'informations se retrouvant dans toutes les données captées par les PMUs ?
3. Quels sont les meilleurs modèles pour prédire l'instabilité transitoire en temps réel ?

1.2.1 Objectifs de la thèse

L'objectif de cette thèse est de proposer aux opérateurs de réseaux électriques une démarche systématique qui permet de prédire l'instabilité des réseaux électriques en temps réel en allant chercher les signatures cachées dans les données fournies par les PMUs dans les conditions normales d'opération et les conditions d'urgence. Une telle démarche va faciliter le contrôle, la surveillance et le développement d'automatismes pour consolider la sécurité et la fiabilité du réseau.

Les objectifs spécifiques qui vont mener à l'atteinte de cet objectif principal sont :

1. De créer une base de données renfermant des points d'opération du réseau ;
2. D'investiguer la base de données à la recherche de relations entre les différentes variables afin d'établir des critères de prédiction de l'instabilité ;
3. D'identifier les paramètres apportant le plus d'informations dans l'évaluation de la stabilité afin de réduire la taille de la base de données ;
4. De proposer une nouvelle approche de prédiction de l'instabilité en temps réel se basant sur les données de l'instant de l'évaluation de l'état du réseau et non sur des photos instantanées (*snapshot*, anglais) du réseau ;
5. De prédire l'instabilité en fonction des prédicteurs retenus en utilisant la modélisation par la forêt aléatoire d'arbres de décision ;
6. De comparer les résultats avec d'autres approches ;

7. De proposer une autre approche de prédition de l'instabilité tenant compte de la mise à jour automatique de la base de données avec de nouveaux point d'opération.

Les objectifs ainsi fixés amènent à l'adoption de la méthodologie présentée dans la sous-section suivante.

1.2.2 Méthodologie

Les étapes présentées ci-après ont permis d'atteindre les objectifs fixés :

1. Création d'une base de données représentatives du problème posé ;
2. Exploration des données pour trouver des relations entre les différentes variables ;
3. Établissement de critères d'instabilité ;
4. Mise au point de modèles prédictifs se basant sur les forêts aléatoires et le boosting ;
5. Prédition en temps réel de l'état du réseau à partir des indices développés (stable ou instable).

Ces différentes étapes sont détaillées au Chapitre 6.

1.2.3 Contribution de la thèse à l'avancement des connaissances scientifiques et/ou technologiques

La présente thèse propose un nouveau concept pour l'étude de la stabilité transitoire en temps réel. En se référant aux informations fournies par les synchrophaseurs, ce concept permet de suivre le comportement dynamique effectif du réseau de transport. De ce concept, s'ensuivent plusieurs autres contributions listées ci-après :

1. Élaboration du nouveau concept de centre de puissance pour l'évaluation de la stabilité transitoire ;
2. Conception de nouveaux indices fondés sur le concept de centre de puissance pour la prédition de la stabilité en temps réel ;
3. Évaluation de la stabilité avec une technique de fenêtre glissante ;

4. Exploitation des techniques de fouille de données, notamment le *boosting* et la forêt aléatoire, pour proposer une solution à la prédition de la stabilité transitoire en temps réel :
 - a. Utilisation des données de trois réseaux test pour la conception et la validation de la méthode ;
 - b. Identification des prédicteurs les plus importants ;
 - c. Développement d'une nouvelle technique se reposant sur le *boosting* pour mettre à jour la base de données avec de nouveaux points d'opération et pour ainsi continuer à prédire la stabilité transitoire en temps réel.

1.2.4 Originalité du travail

Il est communément utilisé dans la littérature le centre d'inertie pour réaliser l'étude de la stabilité. L'originalité de cette thèse repose principalement sur la notion de centre de puissance introduite pour évaluer la stabilité transitoire. Contrairement à l'inertie qui est une valeur invariante dans le temps et difficilement accessible, le centre de puissance qui dépend de la puissance effective des machines permet de suivre la dynamique en temps réel du réseau grâce aux données des synchrophaseurs. Les indices en découlant considèrent l'apport réel de chacune des machines du réseau sur le maintien ou non du synchronisme. De plus, lors de l'évaluation de la stabilité, on utilise des données de l'instant de l'évaluation obtenues grâce à une technique de fenêtre glissante. Contrairement aux *snapshots* qui donnent une photo du réseau 150 ou 300 ms après le défaut, la fenêtre glissante ainsi proposée permet de décider en tout temps à partir des valeurs observées à l'instant en question. Finalement, une utilisation intelligente des forces de deux techniques de forage des données (forêt aléatoire et boosting) est faite pour favoriser la mise à jour automatique de la base de données et pour augmenter les performances du modèle de prédition.

CHAPITRE 2

REVUE DE LITTÉRATURE

2.1 Introduction

Maintenir la stabilité et la sécurité est une problématique inhérente à l'existence même d'un réseau électrique. Il faut en tout temps assurer la continuité dans la fourniture de l'électricité même en cas de défauts très sévères. Durant les 50 dernières années, les réseaux électriques ont énormément augmenté en taille. Il en résulte, en toute logique, une croissance accrue du nombre d'interconnexions. La conception d'outils efficaces pour exploiter les réseaux électriques devient de plus en plus délicate. De plus, avec l'avènement des synchrophaseurs, il est possible de montrer et de contrôler en temps réel le comportement du réseau (Bernabeu, Thorp et Centeno, 2012) – (Pal et al., 2013). Pour l'étude de la stabilité transitoire, de nombreuses techniques tenant compte des données en temps réel ont émergé pour une exploitation sécuritaire et économique du réseau électrique. Cette étude est pertinente tant au niveau de la planification et la gestion de grands réseaux qu'au niveau de dimensionnement d'équipement et de protection (Khan, Saeed et Lo, 2008).

Dans ce chapitre, on passe en revue les techniques les plus en populaires dans la littérature pour l'étude de la stabilité transitoire. On procède également à une analyse critique de la littérature.

2.2 Méthodes de prédition de la stabilité transitoire

Plusieurs méthodologies existent pour évaluer la stabilité transitoire. Elles se distinguent par les techniques de modélisation et les hypothèses adoptées. Il est à noter qu'une analyse de la stabilité transitoire va au-delà du maintien du synchronisme des machines synchrones après de graves perturbations. On examine également le comportement dynamique du réseau au complet et les phénomènes électromécaniques qui se produisent des fois plusieurs secondes après la durée du défaut.

La Figure 2.1 suggère de classer en trois grandes familles les méthodes pour l'analyse de la stabilité transitoire, lesquelles méthodes sont décrites dans cette section.

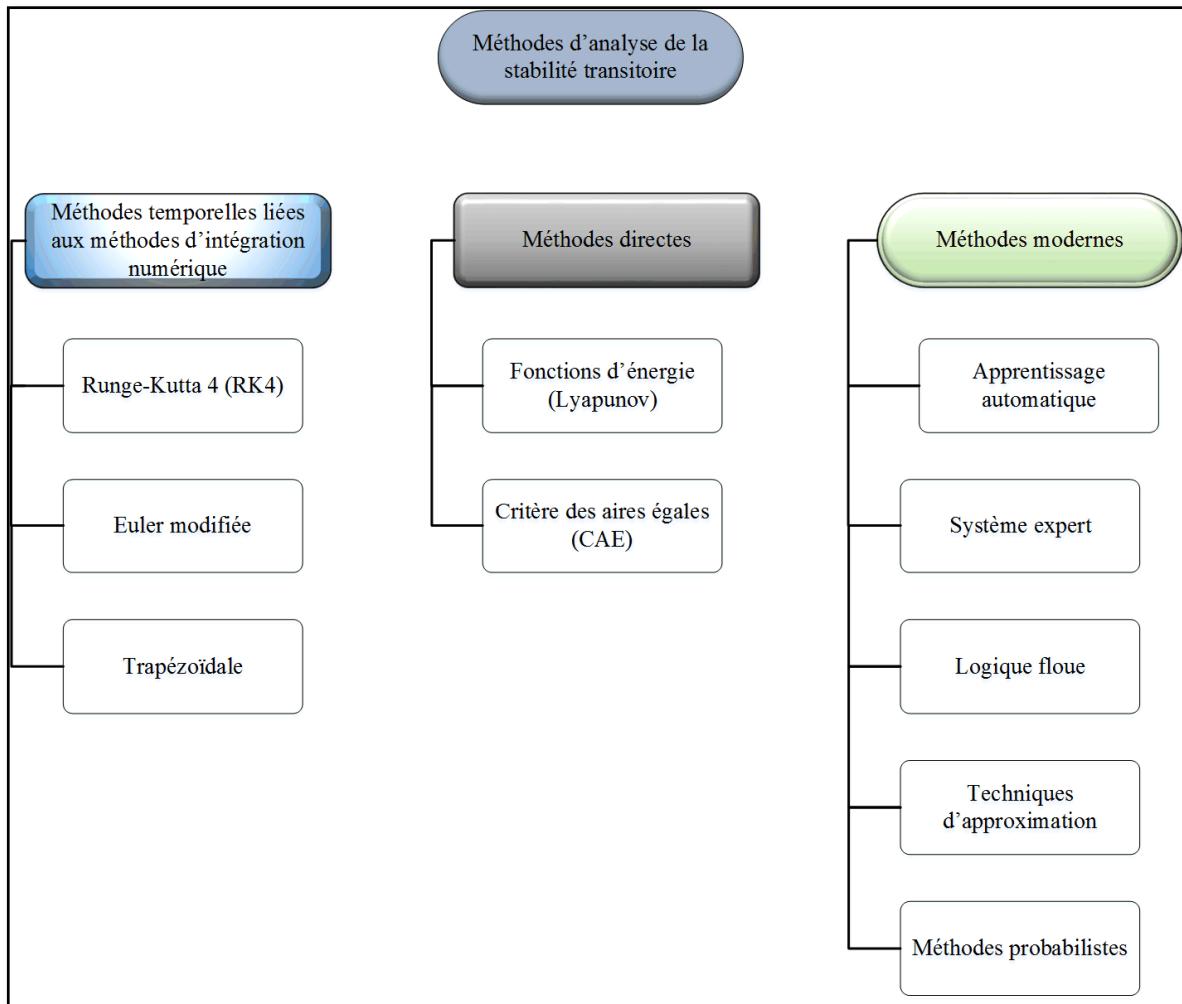


Figure 2.1 Familles de méthodes pour la prédiction de la stabilité transitoire

2.2.1 Méthodes temporelles liées aux méthodes d'intégration numériques

L'étude de la stabilité par les méthodes temporelles (*Time-Domain*, TD) liées aux méthodes d'intégration numérique consiste à résoudre dans le domaine du temps les équations différentielles non linéaires décrivant les variables d'état du système (Liu et Thorp, 2000), en l'occurrence le système d'équations démontré plus loin dans les équations (3.7) et (3.8). Ces

méthodes étaient surtout utilisées avant la percée des ordinateurs dans le monde de la recherche, on utilisait des techniques de résolution manuelle des équations différentielles (Park et Bancker, 1929). Pour un réseau à n machines, on a 2^n équations différentielles à résoudre simultanément. Pour pallier à ce problème, on utilise des méthodes numériques implicites et explicites à pas d'intégration constant et/ou variable (Stagg et El-Abiad, 1968). La méthode de Runge-Kutta d'ordre 4 (RK4) repose sur le principe d'itération. La première estimation de la solution est utilisée pour calculer une seconde estimation plus précise. On répète cette démarche à 4 reprises (Adibi, Hirsch et Jordan, 1974). La méthode d'Euler est une procédure numérique pour résoudre par approximation des équations différentielles du premier ordre avec une condition initiale. Elle est peu précise. La méthode modifiée d'Euler est une méthode explicite qui permet de réduire l'erreur globale en ajoutant deux étapes de prédiction et de correction dans l'approximation. La méthode trapézoïdale revient à estimer l'intégrale d'une fonction comme l'intégrale de son interpolation linéaire par intervalles en utilisant des trapèzes (Dommel et Sato, 1972). Pour l'analyse de la stabilité transitoire, (Apraez, 2012) a démontré que les méthodes de Runge-Kutta sont les plus efficaces. Toutes ces techniques ont comme point commun : la connaissance de l'état initial du réseau, le choix du pas de calcul et finalement l'application du calcul d'intégration. Ces approches sont pratiques quand on fait de l'analyse hors-ligne. Pour des opérations en temps réel, on rencontre de nombreuses difficultés (Milano, 2010) :

- a. Chaque défaut indistinctement doit être traité séparément ;
- b. Les conditions limites de stabilité sont obtenues pour chacune des configurations ;
- c. Il faut des pas de calcul très petits pour conserver la précision et la stabilité de la méthode numérique ; ce qui implique un très long temps de calcul pour un grand nombre d'équations différentielles non linéaires (Anderson et Anderson, 1995).

2.2.2 Méthodes directes

L'objectif principal des méthodes directes est de remplacer l'intégration numérique des équations représentant le réseau après le défaut par un critère de stabilité. On distingue généralement dans la littérature deux grandes classes :

1. Le critère des aires égales (CAE) est une méthode graphique simple et facile d'implantation. Pour déterminer si un réseau est stable après une contingence, on trace la courbe de puissance en fonction de l'angle du rotor pour les périodes avant, pendant et après le défaut (Kimbark, 1995). On pose comme hypothèse (Pavella, Ernst et Ruiz-Vega, 2012) que les machines synchrones ont une puissance mécanique constante et un amortissement négligeable et qu'elles sont représentées par des sources de tension constantes avec leur réactance transitoire. Les charges sont représentées par une impédance à caractéristiques constantes. Sur la Figure 2.2, l'aire A_{acc} représente l'énergie cinétique nette emmagasinée (énergie d'accélération) pendant le défaut dans le rotor. Par contre, l'aire A_{dec} représente l'énergie perdue pendant la décélération. Le réseau est stable après le défaut si l'aire de décélération est supérieure à l'aire d'accélération ; instable sinon (Athay, Podmore et Virmani, 1979) – (Xue, Cutsem et Ribbens-Pavella, 1988). Le CAE est limité car il nécessite une simplification du réseau en une machine connectée à une barre infinie (Pavella, Ernst et Ruiz-Vega, 2012). Cette contrainte fait en sorte que leur application n'est pas totalement efficace sur les grands réseaux électriques interconnectés. La variante généralisée du CAE (*Extended Equal Area Criterion*, EEAC) vient contrer cette problématique. Elle prend en compte les grands réseaux tout en effectuant un nombre très élevé de simulations. Elle va au-delà de la stabilité *first-swing* et permet d'analyser tout type d'instabilités susceptibles de survenir dans la pratique et de concevoir des moyens de les traiter (Xue et al., 1992).

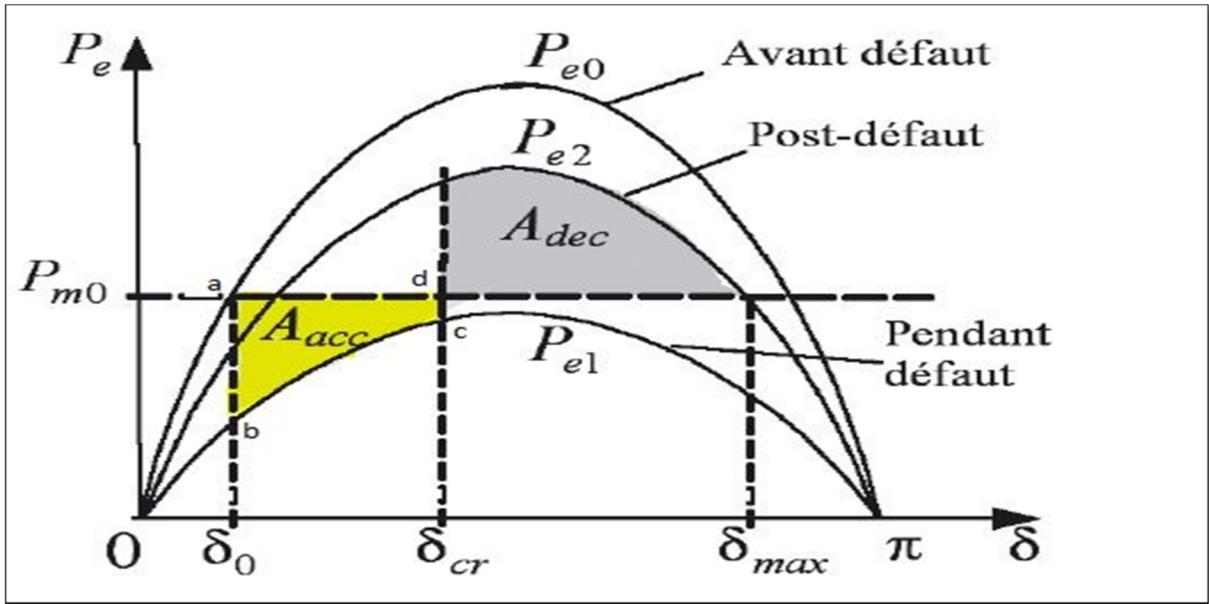


Figure 2.2 Caractéristiques $P = f(\delta)$ avant, pendant et après défaut
Adaptée de Xue et al. (1992, p. 3)

2. Le développement des fonctions d'énergie pour l'analyse de la stabilité transitoire ont débuté vers la fin des années 60 (El-Abiad et Nagappan, 1966). Ces méthodes reposent essentiellement sur les fonctions de Lyapunov. Soit V une fonction de Lyapunov qui assure la stabilité du réseau dans une région autour du point d'équilibre stable après défaut (*Stable Equilibrium Point*, SEP) et soit V_{cr} sa valeur à la limite de la région. Si V est inférieur à V_{cr} , pendant la période transitoire juste après le défaut, le réseau est stable (Ribbens-Pavella et Evans, 1985) – (Pai, 2012).

Ces méthodes sont classées en 4 catégories selon (Chiang, Wu et Varaiya, 1994) : *Closest Unstable Equilibrium Point* (UEP), *Controlling UEP*, *Potential Energy Boundary Surface* (PEBS), et *Boundary of Stability Region Controlling UEP* (BCU). Elles fonctionnent toutes selon le principe qui suit. Le point d'opération du réseau est considéré comme un point d'équilibre stable entouré par une région dans l'espace d'état, appelé la région de la stabilité opérationnelle du réseau. Si à l'instant initial du défaut, le point de fonctionnement du réseau est dans la région de stabilité, il finira par revenir au point de fonctionnement stable. Le temps requis par un défaut pour conduire le système en dehors de la région de stabilité est appelé le temps critique. D'autre part, si

l'état du réseau est initialisé en dehors de la région de stabilité, il restera en dehors de la région de stabilité et n'approchera jamais son point de fonctionnement stable.

Les méthodes des fonctions d'énergie présentent comme avantages la réduction du temps de calcul énorme pris avec les méthodes TD et le calcul de marges de stabilité capables de prédire l'état du réseau. Cependant, pour les réseaux avec plusieurs machines, il est toujours difficile de trouver la fonction de Lyapunov appropriée à moins d'utiliser un réseau simplifié (El Guindi et Mansour, 1982). De plus, la prédiction est valide uniquement si on se trouve à l'intérieur de la région délimitée comme étant stable (Llamas et al., 1995) ; dans le cas contraire, on ne peut rien avancer sur l'état du réseau (Llamas et De La Ree, 1993). La prédiction ne se fait que pour les instabilités *first-swing* ; pour les instabilités *multi-swing*, les méthodes de fonction d'énergie n'apportent aucune réponse. Pour surmonter ces contraintes, plusieurs pistes de solutions sont explorées dont la définition d'une fonction pseudo-Lyapunov traitée avec les méthodes TD (Fink et Trygar, 1979) et des techniques basées sur le fonctionnement en dehors de la région de stabilité BCU (Chiang, Wu et Varaiya, 1994).

Certaines méthodes dites hybrides combinent les méthodes directes et les méthodes TD, toutes deux dites des méthodes déterministes, pour l'analyse de la stabilité transitoire. Par exemple, pour déterminer le temps critique d'instabilité, (Xue et al., 2015) fait appel au CAE et à des techniques d'intégration numérique. (Pavella, Ernst et Ruiz-Vega, 2012) utilise une représentation d'une machine connectée à une barre infinie (*One-Machine Infinite Bus*, OMIB) pour déterminer la marge de stabilité et le temps critique. Dans cette approche, toutes les machines du réseau sont classées en deux groupes critique ou non-critique, puis on déduit un modèle simplifié d'une seule machine pour représenter tout le réseau.

2.2.3 Méthodes modernes

Les méthodes modernes englobent les métaheuristiques et toutes les méthodes basées sur l'apprentissage automatique, la logique floue, les approximations et les statistiques. L'apprentissage automatique (AA) a été introduit au début des années 70 (Liacco, 1974). À cette époque, des méthodes relatives à l'AA ne pouvaient pas être développées avec des résultats concluants à cause des technologies informatiques limitées et peu adaptées. Avec le développement des synchrophaseurs et de nombreux outils informatiques ces dernières années, de nouvelles techniques de prédiction de stabilité transitoire nécessitant peu de temps de calcul sont mises de l'avant.

Approximations de la trajectoire du rotor après défaut

Les méthodes basées sur la trajectoire du rotor après défaut utilisent les données recueillies par les synchrophaseurs (Rajapakse et al., 2010b). Un réseau équivalent réduit est déterminé en observant dans le temps les valeurs des angles du rotor après-défaut pendant une fraction de seconde, et en les regroupant en groupes cohérents (Liu et Thorp, 1995). Le système équivalent est intégré en utilisant la plus récente série de mesures de phaseurs comme conditions initiales. On procède aussi à une approximation de l'angle d'oscillation en considérant une fonction polynôme de la forme de l'équation (2.1). Les coefficients a_i sont obtenus en utilisant les données des instants précédents (Haque et Rahim, 1989). L'angle $\hat{\delta}(t)$ estimé correspond à la valeur prédite à l'instant t .

$$\hat{\delta}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \quad (2.1)$$

La trajectoire peut être aussi décrite en utilisant une fonction trigonométrique (équation 2.2) (Song, Bi et Yang, 2006) ou en se servant d'une auto-régression avec n points (équation 2.3) :

$$\hat{\delta}(t) = \sum_{n=0}^{\infty} a_n \cos(nt) + b_n \sin(nt) \quad (2.2)$$

$$\hat{\delta}(t) = \alpha_1\delta(t-1) + \alpha_2\delta(t-2) + \cdots + \alpha_n\delta(t-n) \quad (2.3)$$

L'extrapolation par un polynôme a l'avantage que la topologie du réseau n'est pas importante. L'approximation par une sinusoïde a l'avantage de bien prédire les cas stables, contrairement aux cas instables. Avec une auto-régression, les résultats ne sont intéressants que s'il n'y a pas de changement brusque de la variation de l'angle d'oscillation.

Certains auteurs proposent aussi de trouver d'abord une approximation de la vitesse du rotor par une interpolation avec la méthode de Newton (équation 2.4). De cette vitesse, on peut déduire par une intégration l'angle du rotor (Karady, Daoud et Mohamed, 2002) en appliquant l'équation (2.5) où T est la période, t_i les différents instants, $\hat{\delta}$ et $\hat{\omega}$ sont respectivement l'angle et la vitesse du rotor estimés.

$$\hat{\omega}(T + t_0) = \omega(t_2) + \alpha_1(T - t_2) + \alpha_2(T - t_1)(T - t_2) \quad (2.4)$$

$$\hat{\delta}(T + t_0) = \int_{t_0}^{T+t_0} \hat{\omega}(t) dt + \delta(t_0) \quad (2.5)$$

Détection de l'instabilité en temps réel avec des indices

Certains travaux (Kaci et al., 2014a) démontrent que les angles permettent une bonne prédiction de la puissance maximale et les limites de sécurité du réseau. Les méthodes de prédiction basées sur la détection de l'instabilité transitoire utilisent des critères de seuil comme limite. Une fois ce seuil atteint ou dépassé, le réseau est considéré instable. Deux de ces méthodes (Stanton et al., 1995) consistent à déterminer une valeur relative de l'angle du rotor ($\bar{\delta}_i$) (ou de la vitesse rotorique ($\bar{\omega}_i$)) des machines qui doit-être inférieure à un certain seuil après l'effacement du défaut (voir les équations 2.6 et 2.7).

$$\bar{\delta}_i = \delta_i - \frac{\sum(M_k \delta_k)}{\sum M_k} \leq \Delta \delta \quad (2.6)$$

$$\overline{\omega}_i = \omega_i - \frac{\sum(M_k \omega_k)}{\sum M_k} \leq \Delta \omega \quad (2.7)$$

Où δ_i (δ_k), ω_i (ω_k), M_k , sont respectivement l'angle, la vitesse du rotor et le moment d'inertie de la machine i (k).

Pour détecter une instabilité, on utilise aussi des fois le taux de variation des vitesses et/ou des angles des différentes machines (Roy, Sinha et Pradhan, 2013) – (Liancheng et Girgis, 1997). Les critères de valeur de seuil sont faciles à mettre en œuvre. Les inconvénients sont que les valeurs de seuil raisonnables sont difficiles à définir et les critères de seuil exigent un grand nombre de simulations.

Une autre méthode basée sur les synchrophaseurs s'articule autour des angles des générateurs, des vitesses angulaires et du taux de changement des vitesses angulaires par rapport aux angles des jeux de barres. On utilise la concavité ou la convexité de la surface décrite après-défaut par la courbe de l'angle en fonction de la vitesse. Pour une surface convexe, le système est considéré instable et pour une surface concave, le réseau est stable. En fonction des caractéristiques géométriques, on définit un indice Δk qui permet d'identifier l'instabilité. L'avantage de cette méthode est qu'elle est indépendante de la structure du réseau, de ses paramètres et de son modèle. Les équations suivantes tirées de (Xie et al., 2005) décrivent cet indice :

$$k_i(t+1) = \frac{\omega_i(t+1) - \omega_i(t)}{\delta_i(t+1) - \delta_i(t)} \quad (2.8)$$

$$\Delta k_i(t) = k_i(t+1) - k_i(t) > 0$$

Où $\delta_i(t+1)$ et $\omega_i(t+1)$ sont les valeurs de l'angle et de la vitesse angulaire de la machine i mesurées à l'instant $t+1$ et k_i , Δk_i des constantes décrivant l'indice.

Si Δk_i est toujours inférieur à zéro, le système est stable, tandis que le système perd la stabilité si Δk_i est supérieure à 0 immédiatement ou peu de temps après le temps d'effacement de défaut.

Le début de l'instabilité est détecté en observant la trajectoire post-défaut du réseau mono-machine équivalent à chaque pas de temps. Le mode de perturbation qui reflète le mieux la stabilité est déterminé et actualisé à chaque pas de temps d'échantillonnage. Contrairement au CAE (Suprême et Dessaint, 2012), cette méthode ne nécessite pas de calcul d'un point d'équilibre instable.

Se reportant à l'OMIB, une analyse géométrique du plan de trajectoire de phase est appliquée à la prédiction en temps réel de l'instabilité transitoire pour un système multi-machine. L'écart d'angle est utilisé comme un indice pour choisir la direction dominante de recherche de mode et les conditions d'instabilité pour arrêter la recherche de mode dominant. Le mode dominant approche donc l'instabilité réelle pas-à-pas. La prédiction de la trajectoire équivalente est réalisée uniquement lorsque la variation de paramètre peut affaiblir la stabilité du système. Et finalement, les caractéristiques géométriques de la trajectoire prédictive sont utilisées pour déterminer si le système est stable.

Méthodes probabilistes

Contrairement aux méthodes temporelles et directes qui proposent des solutions déterministes à l'analyse de la stabilité, les méthodes décrites dans cette sous-section abordent le problème sous un angle probabiliste. On tient compte des facteurs provoquant les défauts (type de défaut et son emplacement) et de l'état du réseau (sa configuration et sa topologie). En d'autres termes, on évalue la probabilité pour que le réseau reste stable lorsqu'une contingence spécifique se produit. Dans la littérature, on retrouve deux techniques principalement. La première technique recourt au théorème de probabilité conditionnelle pour évaluer analytiquement la probabilité de stabilité (Billinton et Aboreshaid, 1995) – (Billinton, Kuruganty et Carvalho, 1979). La seconde applique l'approche de simulation Monte-Carlo

(Chiodo, Gagliardi et Lauria, 1994) – (Anderson et Bose, 1983). L'objectif de ces méthodes est de déterminer des distributions de probabilité pour la stabilité du réseau ; l'aspect stochastique des conditions d'opération pré-défaut du réseau est traité dans (Anders, 1990) – (Yuan-Yih et Chung-Liang, 1988). Dans cet article (Anderson et Bose, 1983), on explore la nature des transformations nécessaires pour déterminer la probabilité de stabilité transitoire définie en fonction des marges de stabilité. Il est démontré d'une part, qu'il n'est pas évident de définir de telles transformations ; et d'autre part, qu'une simulation probabiliste utilisant la méthode de simulation de Monte-Carlo est réalisable. Toutefois la technique de simulation de Monte-Carlo appropriée pour l'analyse des événements complexes, peut nécessiter un temps de calcul considérable. Pour réduire le temps de calcul dans l'analyse de la stabilité transitoire, (Aboreshaid, Billinton et Fotuhi-Firuzabad, 1996) – (McCalley et al., 1997) proposent une méthode de bisection. De manière générale, ces méthodes sont pertinentes pour la planification (Vaahedi et al., 2000) mais inconvenables pour des applications en ligne. On peut néanmoins les combiner avec d'autres techniques telles que la reconnaissance des formes pour la prédiction en ligne.

Système expert

Dans cette approche, on requiert l'expertise des opérateurs de réseaux électriques. Leur connaissance est encodée dans des programmes suivant des règles d'inférence précises (Liu et Tomsovic, 1986). Cette méthode combine les méthodes TD (algorithme numérique) et les méthodes directes. Pour atteindre son objectif, elle utilise des données (écoulement de puissance, analyse de stabilité) acquises lors des simulations hors-ligne afin de déduire l'état du réseau (Fouad, Vekataraman et Davis, 1991). Ces analyses doivent être effectuées sur des conditions le plus diversifiées que possibles du réseau, i.e. topologies de réseau différentes, conditions de charge variées, type et localisation des défauts pluriels. Les résultats de ces analyses sont ensuite marqués avec certains attributs, à savoir l'état du réseau à l'étude. Des exemples d'applications concrètes se retrouvent dans ces publications (Christie, Talukdar et Nixon, 1990) – (Xue et al., 1989). L'avantage capital de ces méthodes est qu'elles relatent l'état réel du réseau car les opérateurs (experts) s'y connaissent bien de par leurs expériences.

Logique floue

La logique floue est une extension de la logique booléenne et constitue une généralisation de la théorie des ensembles classiques se fondant sur le « principe du tiers-exclus ». En introduisant la notion de degré dans la vérification d'une condition, permettant ainsi à une condition d'être dans un autre état que vrai ou faux, la logique floue confère une flexibilité très appréciable aux raisonnements qui l'utilisent, ce qui rend possible la prise en compte des imprécisions et des incertitudes (Yager et Zadeh, 2012). L'application de la théorie des ensembles flous dans l'évaluation de la stabilité transitoire a été rapportée pour la première fois dans cette publication (Souflis, Machias et Papadias, 1989). Les points d'opération du réseau sont classés dans l'un des six états suivants : très instable, instable critique, instable, très stable, stable critique, stable. Ils sont représentés à l'aide des valeurs d'appartenance floue dans un système de reconnaissance des formes floues. La méthode développée est applicable à tout réseau indépendamment de sa taille, de sa configuration, de la répartition de charge ou de sa topologie (Machias, 1989). Une approche similaire utilisant un indice d'instabilité en fonction de l'angle de rotor du générateur et de l'accélération, est proposée pour l'identification des machines instables dans les systèmes électriques multi-machines (Machias et Souflis, 1990). Ces publications (Kamwa, Pradhan et Joós, 2007) – (Kamwa et al., 2009a) recourant aux données des synchrophaseurs, démontrent l'applicabilité de l'algorithme de logique floue (*Fuzzy C-medoid*, FCMdd) sur les réseaux électriques (grands réseaux et réseaux tests) et sa capacité de partitionner convenablement la base de données. Il est aussi prouvé que les indices de degré de sévérité des contingences basés sur ces méthodes fonctionnent très bien pour des configurations et des profils de charges variés du réseau.

On retient aussi le réseau de fonction radiale basé sur la logique floue. Son algorithme d'apprentissage avec des entrées simples fonction d'angles de rotor du générateur après un défaut mesuré par les synchrophaseurs peut être facilement implémenté (Liu et al., 1999c). Une autre méthode s'intéresse aux trajectoires de l'angle du rotor juste après le défaut. À partir de simulations hors-ligne, l'état du réseau et le comportement des angles du rotor sont enregistrés sous différentes perturbations. La distance euclidienne entre les trajectoires de

mesure d'angle de rotor en temps réel et de la trajectoire stockée dans la base de données est calculée (Liu, Jiang et Cao, 2008). Les trajectoires d'angle de rotor sont prédites à l'aide d'une expression modifiée provenant de l'apprentissage du réseau (Liu et Lin, 2000).

À titre d'exemple concret, une méthode de prédiction de la stabilité transitoire se référant au réseau flou « hyper rectangulaire composite » de neurones (*Fuzzy Hyper-Rectangular Composite Neural Network*, FHRCNN) est proposée dans (Liu et al., 1999c). On utilise les données collectées à l'effacement du défaut. Il s'agit de trois valeurs consécutives de l'angle de chaque machine. Il est calculé par la suite deux vitesses et une accélération, ce qui fait un total de 6 prédicteurs (entrées) par générateur. Le critère d'instabilité utilisée consiste à déterminer si la différence entre les deux angles des machines est supérieure ou non à π radians dans la première seconde après l'effacement. L'écart de π radians est utilisé à titre illustratif mais il doit être calculé comme il dépend fortement des caractéristiques du réseau en présence.

Apprentissage automatique (*Machine learning*)

L'apprentissage automatique, champ de l'intelligence artificielle, regroupe trois grandes classes : l'apprentissage supervisé ou non supervisé (*supervised or non-supervised learning*), les réseaux de neurones artificiels (*Artificial Neural Networks*, ANN) et la reconnaissance des formes (*Recognition Pattern*). Les trois familles d'approches visent à extraire l'information essentielle contenue dans un ensemble de cas pré-analysés lors d'une phase d'apprentissage (Alpaydin, 2014). Des recherches récentes portant sur les arbres de décision (DT) (Rahmatian et al., 2017) – (Rovnyak et al., 1994b), les ANNs (Chih-Wen et al., 1999) – (Bahbah et Girgis, 2004), la théorie de l'apprentissage statistique (Li et Cao, 2005) – (Kamwa, Grondin et Loud, 2001a) et les machines à vecteurs de support avec des classificateurs ajoutés (SVM) (Moulin et al., 2004) – (Sharma et Tyagi, 2013) montrent que les techniques d'apprentissage supervisé sont des approches prometteuses pour l'analyse et la prédiction de la stabilité transitoire des réseaux électriques (Sobajic et Pao, 1989) – (Gomez et al., 2011).

Plusieurs techniques de prédiction de l'instabilité faisant appel aux ANNs sont traitées dans la littérature. Une première approche d'estimation de l'angle du rotor à partir des PMU en utilisant les réseaux de neurones artificiels (Del Angel, Glavic et Wehenkel, 2003) et (Del Angel et al., 2007) consiste à modéliser de manière détaillée la dynamique du réseau et des dispositifs d'acquisition des données. On utilise ensuite un simulateur hors ligne pour générer un ensemble de trajectoires du système représentatif du comportement du réseau avant, pendant et après défaut. La base de données ainsi créée doit être suffisamment grande pour tenir compte d'un large éventail de possibilités. Finalement, on se sert de l'apprentissage supervisé automatique pour former un modèle de boîte noire sous la forme d'un perceptron multicouche (MLP) pour estimer des valeurs moyennes d'angles de rotor basées sur les données des synchrophaseurs.

Les SVMs représentent des méthodes d'apprentissage émergentes qui résolvent théoriquement certains inconvénients des ANNs. Ces dernières sont surtout performantes dans la reconnaissance des formes. Une méthode de classification simple et rapide pour l'évaluation de la stabilité est présentée dans cette dernière référence (Kalyani et Swarup, 2011). Un autre classificateur est fait dans (Moulin et al., 2004) et sa performance est comparée à un perceptron multicouche (MLP). Une méthode de classification proposée dans (Rajapakse et al., 2010a) se basant sur la variation de la tension et des trajectoires d'angle du rotor préalablement simulées permet de prédire l'état de stabilité transitoire. Les tensions mesurées sont comparées avec des modèles pour évaluer la concordance entre les variations de tension mesurées et celles des bases de données. Les valeurs de similarité sont entrées au SVM formés pour faire le classement. D'autres méthodes de classification utilisant les synchrophaseurs existent également (Gomez et al., 2011).

Les arbres de décision utilisent un grand nombre de points de fonctionnement du réseau avec leurs critères de stabilité pour créer un modèle de prédiction de la stabilité transitoire (Li et al., 2014) – (Wehenkel et al., 1994). L'apprentissage est fait hors-ligne mais le modèle fonctionne en ligne (Rovnyak et al., 1994a). Ces méthodes sont amplement débattues dans les chapitres suivants.

La plupart des méthodes d'apprentissage automatique actuelles présentent de grands avantages mais connaissent quelques inconvénients inhérents limitant ainsi l'application pratique de l'évaluation de la stabilité transitoire. En effet, les ANNs ont une grande capacité de généralisation avec une rapidité d'exécution (AL-Masri et al., 2013) ; mais ils ont des problèmes de «surapprentissage» (Moulin et al., 2004). Les SVM rencontrent des difficultés dans la sélection des paramètres de prédiction à cause de leur incapacité à considérer pleinement les études spécifiques du réseau (topologie, niveau de charge, les différents états de fonctionnement) utilisées dans la pratique (Vapnik, 2000).

2.3 Revue critique de la littérature

Pour la prédiction de la stabilité transitoire, aucun intérêt particulier n'est porté sur le CCT (Bergen, 2000). Dans une telle démarche, on veut juste suivre l'évolution du réseau en cas de perturbation. La question clé de la prédiction de la stabilité transitoire est : les oscillations transitoires sont-elles convergentes ou divergentes ? (Glover, 2002)

Avec les méthodes numériques d'intégration, on rencontre de nombreuses difficultés d'implémentation. Elles exigent de longs calculs. En outre, chaque exécution de ces méthodes nécessite généralement une préparation minutieuse de gros fichiers de données d'entrée (Pai, 1989). Le concept d'indice développé à date ne fait mention que des conditions nécessaires pour l'atteinte de l'instabilité et non des conditions suffisantes (Guo et al., 2010) – (Kai, Lee et Pei, 2011). En approximant l'angle du rotor par un polynôme de degré n , le nombre de points dont on dispose n'est pas forcément suffisant pour décrire le comportement général de la variation de ce dernier (Hui et al., 2012). En cas de variation brusque, le critère établi par l'indice devient peu précis. L'approximation par des polynômes, par la méthode des moindres carrés et par des sinusoïdes est de moins en moins précise pour des cas instables. Le grand défi reste et demeure la quantité de données qui doit être utilisée pour l'évaluation.

Les méthodes découlant de l'apprentissage des machines sont de plus en plus exploitées. Ces dernières nécessitent des simulations hors-ligne. La validité du modèle de prédiction dépend

donc d'un échantillon qui doit être suffisamment représentatif de l'état d'opération du réseau. Les fonctions qui relient les données d'entrée avec celles de sortie sont extrêmement complexes et inconnues à priori (Krishna et Padiyar, 2000). Les arbres de décisions sont utilisés en apprentissage supervisé pour prédire la stabilité transitoire avec des données provenant des PMU (Rovnyak et al., 1994b) – (Qun et Rovnyak, 2011). Toutefois, on constate généralement dans les travaux récents (Rovnyak et al., 1994b) – (Kai et al., 2007) que si la réduction du taux d'erreur en prédiction est faible, la réduction de la taille de l'arbre est problématique. La précision des prévisions de l'arbre de décision considéré seul est donc généralement limitée pour la prédiction de la stabilité transitoire. Les réseaux de neurones de manière générale et ceux se basant sur la logique floue en particulier permettent de prédire la stabilité transitoire sans nécessiter beaucoup de données d'entrée (Amjadi, 2005). Cependant, dans les cas de non-linéarité dynamique, il est difficile de faire l'apprentissage de toutes les données d'entrée et de sortie pour les grands réseaux (Amjadi et Banihashemi, 2010) – (Amjadi et Majedi, 2007). Pour faire une interpolation, les méthodes se basant sur les ANN sont très bonnes ; par contre, pour l'extrapolation, elles posent problème (Kalyani et Swarup, 2011). D'où la difficulté de généraliser. Les SVM ont l'avantage d'être utilisables pour les grands réseaux (Moulin et al., 2002) – (Kamwa, Grondin et Loud, 2001a). Pour la prédiction de la stabilité transitoire, on peut malheureusement se limiter à certaines contingences bien précises, (Moulin et al., 2004) – (Zhonghong, Xiaoxin et Zhongxi, 2005). Ces méthodes sont de fait moins prises en compte dans les applications industrielles. Les SVMs peuvent être utilisés pour l'analyse de la stabilité transitoire dans les réseaux plus importants en raison de leur capacité implicite de sélection de caractéristiques. Cependant, ces réseaux neuronaux ont été habituellement utilisés pour des ensembles limités de contingences. Pour les nouveaux cas de contingences, une prédiction adéquate n'est pas garantie. D'autres applications d'ANN sont mentionnés dans (Mansour et al., 1997) pour la prédiction de la stabilité transitoire de grands réseaux, mais ces méthodes rencontrent les mêmes difficultés mentionnées précédemment.

L'avantage fondamental que présentent toutes ces méthodes est qu'elles peuvent être exploitées en parallèle en jouant sur les forces de l'une pour combler les faiblesses de l'autre. Les systèmes experts s'avèrent très importants car ils permettent d'ajouter des critères de

stabilité en se référant à des problématiques concrètes et réelles de fonctionnement du réseau électrique.

2.4 Conclusion

Dans ce chapitre, on a passé en revue les principales méthodes utilisées dans la littérature pour prédire la stabilité transitoire. On retrouve des méthodes heuristiques, des méthodes s'appuyant sur des simulations et d'autres fondées sur l'entraînement et l'apprentissage de bases de données. Elles cherchent toutes des relations entre les variables décrivant la dynamique du réseau. Les méthodes directes et temporelles constituent des méthodes très anciennes utilisant les fonctions d'énergie et des techniques de simulation. Les méthodes modernes, assez récentes comme leur nom l'indique, s'appuient sur les statistiques, les probabilités et l'apprentissage automatique. On a également présenté les forces et les faiblesses de chacune d'elles.

CHAPITRE 3

ÉTUDE DE LA STABILITÉ

3.1 Introduction

Depuis le début du XX^{ème} siècle, on reconnaît l'étude de la stabilité du réseau électrique comme une problématique importante et elle fait l'objet de sérieuses préoccupations pour les chercheurs (Steinmetz, 1920) – (AIEE, 1926). Historiquement, la stabilité transitoire a été le problème de stabilité dominante abordée sur la plupart des réseaux (Kundur et al., 2004) ; mais avec l'accroissement de l'exploitation des réseaux à leurs conditions limites, plusieurs autres formes de stabilité ont émergé : stabilité de tension, de fréquence, etc... Une compréhension claire des différents types de stabilité et de la façon dont ils sont liés entre eux s'impose pour une conception satisfaisante d'outils appropriés pour la prédition de l'instabilité en temps réel. En effet, dans cette thèse, on approche le problème du point de vue de l'*instabilité* qui est le fait de se retrouver hors de la région de stabilité du réseau.

Dans ce chapitre, on se propose de définir la stabilité de manière globale et dans le cadre spécifique des réseaux électriques tout en identifiant les différents types de stabilité généralement rencontrés sur ces derniers. Par la suite, on aborde les fondements théoriques de la stabilité ainsi que le modèle mathématique utilisé pour son évaluation. Finalement, on relate les différents modes d'opération des réseaux électriques.

3.2 Définition et classification de la stabilité sur les réseaux électriques

Sur les réseaux électriques, un ensemble de moyens sont mis en œuvre afin de maintenir les grandeurs de fréquence et de tension proches de leurs valeurs de consigne. La stabilité est une condition d'équilibre entre des forces opposées. Dépendamment de la topologie du réseau, de son état de fonctionnement et du type de perturbations, on peut avoir un certain déséquilibre, durable ou non dans le temps, entre les différentes forces opposées qui conduisent à des formes d'instabilité. L'identification des facteurs clés contribuant à l'instabilité et l'élaboration de

méthodes d'analyse de la stabilité passent inévitablement par une bonne catégorisation du type et de la nature de la stabilité (Kundur, Balu et Lauby, 1994). C'est dans cette optique qu'on relate dans cette section, les divers phénomènes de stabilité en les classifiant. En un second lieu, on décrit les relations existant entre la fiabilité, la sécurité et la stabilité des réseaux électriques.

3.2.1 Classification de la stabilité des réseaux électriques

Le problème de la classification de la stabilité des réseaux électriques est un problème très ancien. Dans ces publications (Crary, Herlitz et Favez, 1948) – (Committee, 1982), les premières bases ont été jetées pour classifier les phénomènes de stabilité. Toutefois, les définitions ne reflètent pas les besoins actuels des industries et ne prennent pas en compte des scénarios d'instabilité très répandus sur les réseaux actuels. Une classification adaptée au contexte des réseaux électriques modernes est proposée à la Figure 3.1 (Kundur, Balu et Lauby, 1994). Elle prend notamment en compte le temps pris pour évaluer la stabilité, la nature physique et le type de contingences appliquées sur le réseau. Elle est justifiée théoriquement par le concept de stabilité partielle (Vorotnikov, 2012) – (Rouche, Habets et Laloy, 1977).

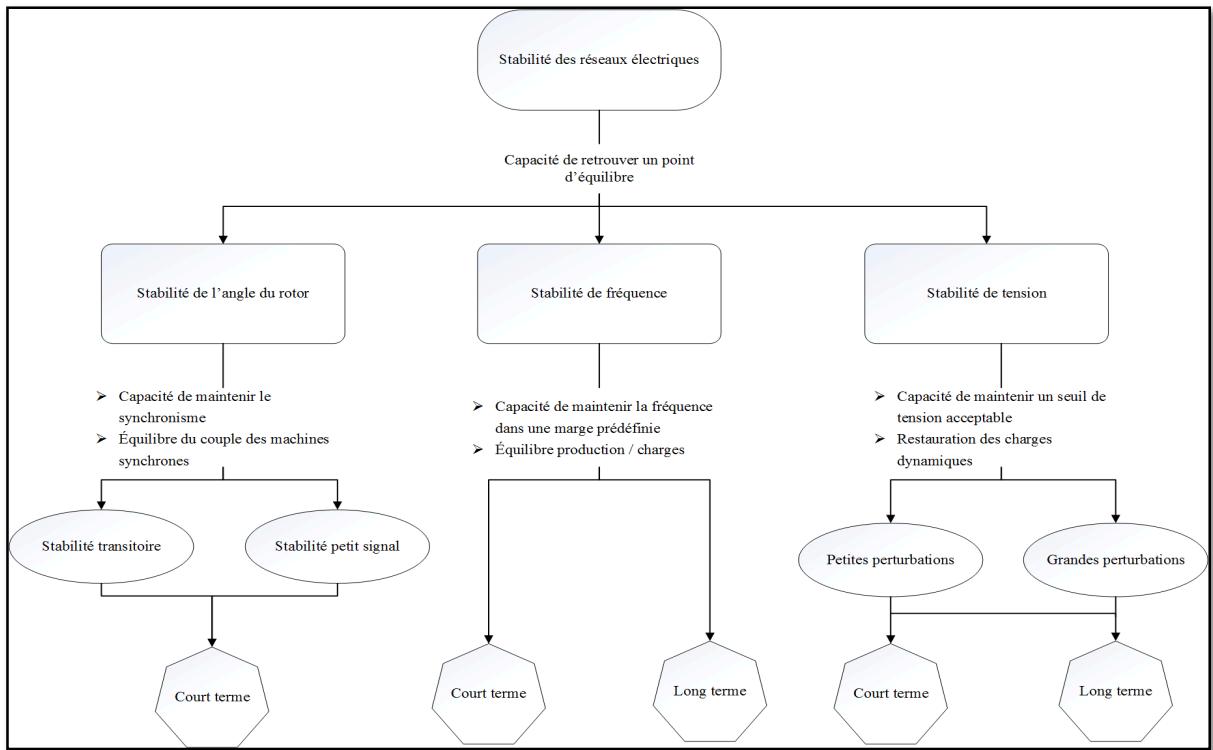


Figure 3.1 Classification des phénomènes de stabilité sur les réseaux électriques
Adaptée de Kundur, Balu et Lauby (1994, p. 36) et Suprême (2012, p. 12)

La stabilité de l'angle du rotor se réfère à la capacité des machines synchrones à garder leur synchronisme suite à une perturbation, i.e. à maintenir ou à rétablir l'équilibre entre le couple électromagnétique et le couple mécanique de chaque machine dans le réseau. L'instabilité qui en résulte se manifeste sous forme d'un accroissement de l'écart angulaire entre deux ou plusieurs machines jusqu'à la perte totale du synchronisme. Le problème de la stabilité du rotor conduit à l'étude des oscillations électromécaniques. Généralement, on analyse le comportement entre les puissances des machines synchrones en fonction de l'angle du rotor. En régime permanent, la vitesse reste constante et il y a un équilibre entre le couple mécanique et le couple électromagnétique de chaque machine. Si le réseau subit une contingence, l'équilibre est rompu et entraîne l'accélération ou la décélération des machines selon les lois du mouvement d'un corps en rotation. Si un générateur fonctionne plus vite que les autres, la position angulaire de son rotor par rapport à celle de la machine plus lente augmente. Une telle différence angulaire fait en sorte qu'une partie de la charge de la machine lente soit transférée à la machine rapide, ce qui tend à réduire sa vitesse et l'écart angulaire. Au-delà d'une certaine

limite, l'augmentation de l'écart angulaire est accompagnée d'une diminution du transfert d'énergie ; par conséquent, la position angulaire continue à augmenter. Il y a instabilité si le réseau ne peut pas absorber l'énergie cinétique correspondant à ces différences de vitesse du rotor (Pai, 2012). Le changement de couple électromagnétique d'une machine synchrone dû à une perturbation peut être résolu soit en synchronisant l'élément de couple avec l'écart angulaire de rotor soit en amortissant l'élément de couple avec l'écart de vitesse.

La stabilité du rotor est divisée en deux sous-catégories : la stabilité petit signal et la stabilité transitoire. La première porte sur la capacité du réseau à maintenir le synchronisme sous de faibles perturbations qui permettent la linéarisation des systèmes d'équations décrivant l'état du réseau (Paserba, 1996) – (Cigré, 1996). Sur les réseaux électriques actuels équipés de nombreux régulateurs de tension, ce type d'instabilité est essentiellement associé à un manque d'amortissement des oscillations (Rogers, 2012). Si ce problème d'amortissement ne concerne qu'un seul générateur ou une petite partie du réseau, on dit que le problème est local (*local plant mode oscillations*) ; mais s'il est repéré dans des interactions entre des grands groupes de générateurs, le problème est global (*interarea mode oscillations*). La seconde concerne la capacité du réseau à maintenir son synchronisme suite à de graves perturbations comme un court-circuit sur une ligne de transport. L'instabilité est surtout due au rapport non linéaire existant entre la puissance du générateur et l'angle du rotor (Pavella, Ernst et Ruiz-Vega, 2012). Elle peut être observée rapidement après le défaut (*first swing*) ou dans un laps de temps plus long (*multi swing*). La stabilité transitoire dépend à la fois des conditions initiales de fonctionnement du réseau, de l'emplacement du défaut, de sa sévérité, de sa durée, de la performance des systèmes de protection et des caractéristiques dynamiques des différents équipements. Ces deux types d'instabilité sont des phénomènes à court terme et ils sont étudiés dans les 10 – 20 secondes suivant le défaut ; sauf dans le cas du *first swing* on s'intéresse aux 3 – 5 premières secondes (Ernst et al., 2001).

Le terme de stabilité dynamique est aussi utilisé dans la littérature pour la stabilité de l'angle du rotor. Toutefois, certains auteurs en font usage pour parler de phénomènes divers. En Europe, ce terme renvoie uniquement à la stabilité transitoire (Machowski, Bialek et Bumby,

2011) – (Barbier, Carpentier et Saccomanno, 1978). En Amérique du Nord, ce terme fait référence à la stabilité petit signal avec des contrôles automatiques sur le réseau (Kundur, Balu et Lauby, 1994). Une définition que l'on adopte d'ailleurs dans le cadre de cette thèse.

La stabilité de tension repose sur la capacité du réseau à maintenir des tensions stables aux jeux de barre après avoir été soumis à une perturbation pour des conditions initiales de fonctionnement données. Il s'agit du maintien ou de la restauration de l'équilibre entre la charge et la production. Cette instabilité s'observe par le dépassement d'un certain seuil de tension (chute progressive) causé par la charge (Van Cutsem et Vournas, 1998). On utilise des fois le terme d'effondrement de tension (*voltage collapse*). Ce type d'instabilité est causé en grande partie par une demande de puissance réactive dépassant celle disponible sur le réseau lors d'une contingence (Van Cutsem et Vournas, 1998) – (Hill et al., 1993). La surtension peut aussi en être la cause (Van Cutsem et Mailhot, 1997). Les conséquences possibles d'une instabilité de tension sont la perte de charge dans une zone, le déclenchement des lignes de production et le mauvais fonctionnement d'appareils du système de protection. La stabilité de tension est catégorisée en petites perturbation et grandes perturbations. Les petites perturbations correspondent à la capacité du réseau à maintenir des tensions stables en réponse à de faibles contingences tel que des modifications incrémentielles de la charge. Cette forme d'instabilité est influencée par les caractéristiques des charges, des commandes continues et discrètes, etc... Dans de telles conditions, les équations du réseau peuvent être linéarisées pour évaluer la stabilité (Gao et al., 1996) – (Morison, Gao et Kundur, 1993). Les grandes perturbations se réfèrent à la capacité du réseau à garder des tensions constantes après de grands défauts comme la perte de production. Cette capacité est déterminée par les caractéristiques des charges et les interactions entre les commandes discrètes et continues et les systèmes de protection. Pour comprendre de tels événements, cela requiert l'analyse de la réponse non linéaire du réseau (Sastry, 2013) – (Zabaiou, Dessaint et Kamwa, 2014). Dans les deux cas, le temps d'étude du phénomène de stabilité peut aller de l'ordre de quelques secondes jusqu'à 10 minutes. C'est la raison pour laquelle on parle de phénomène à court terme et long terme dans la Figure 3.1.

La stabilité de fréquence se rapporte à la capacité du réseau de maintenir ou de rétablir la fréquence d'équilibre sur le réseau après une contingence grave due à un déséquilibre entre la production et la charge. L'instabilité résultant se révèle à travers des fluctuations de fréquence soutenues. Pour compenser ce déséquilibre, on peut utiliser l'énergie cinétique stockée dans les pièces tournantes des machines synchrones (Pai, 2012). Ainsi, les générateurs participant à la commande de fréquence régleront la puissance active fournie pour ramener la fréquence à des valeurs acceptables (Gómez-Expósito, Conejo et Cañizares, 2016). Mais, si l'écart de fréquence est trop grand, on peut assister à un effondrement complet du réseau (Anderson et Fouad, 2003).

Généralement, les problèmes de stabilité de fréquence sont associés à des réponses inadéquates en puissance des centrales, à une mauvaise coordination des systèmes de commande et de protection ou à un manque de production (Bevrani, 2009) – (Force, 1999). Dans les réseaux îlotés, on assiste aussi à une telle instabilité en cas de perte significative de charges ou de production (Hatziyargyriou, Karapidakis et Hatzifotis, 1998). Durant cette instabilité, on observe des variations au niveau de l'amplitude des tensions. Ces phénomènes peuvent durer d'une fraction de seconde (court terme) jusqu'à des minutes (long terme).

Il est important de distinguer les différentes formes de stabilité pour bien comprendre les causes sous-jacentes afin de concevoir des outils appropriés et adéquats pour une exploitation efficace et efficiente des réseaux électriques. Même si l'on s'intéresse à un cas spécifique de stabilité, il faut toujours garder en tête que l'objectif final est de garder la stabilité de manière globale sur tout le réseau.

3.2.2 Relation entre fiabilité, sécurité et stabilité des réseaux électriques

Dans cette sous-section, on tient à faire la différence entre la fiabilité, la sécurité et la stabilité des réseaux électriques car fort souvent, ces termes sont utilisés à tort. Ces clarifications vont aider à éviter toute confusion également par rapport aux métriques fiabilité et sécurité adoptées au Chapitre 5 pour évaluer la performance du modèle de prédiction.

La fiabilité d'un réseau électrique se rapporte à la probabilité d'un fonctionnement satisfaisant sur le long terme et dénote la capacité de fournir un service électrique adéquat continuellement avec d'éventuelles interruptions dans le temps (Billinton, 1970). La sécurité d'un réseau électrique se rapporte à sa capacité à survivre à des perturbations imminentées sans interruption de service. Elle est assimilable à la robustesse du réseau (Morison, Wang et Kundur, 2004) – (Fink et Carlsen, 1978). La stabilité, quant à elle, comme mentionné plus haut, correspond au maintien de l'équilibre entre la production et la consommation sur le réseau.

Dans la conception et le fonctionnement d'un réseau électrique, sa fiabilité est toujours l'objectif recherché. Pour être fiable, la sécurité doit être garantie quasiment en tout temps. Cette sécurité passe par la stabilité et la protection contre les autres risques qui ne sont pas classés comme des problèmes de stabilité : par exemple, des dommages à l'équipement comme une défaillance des câbles de transport et de distribution, une chute des pilonnes engendrée par l'accumulation de verglas (Gingras et al., 2000), etc... D'un autre côté, un système peut être stable suite à une contingence mais ne fonctionne pas dans des conditions de sécurité. Cela peut être entraîné par une surcharge de l'équipement ou des violations de tension dans la configuration post-défaut. L'analyse de stabilité fait donc partie intégrante de l'étude de la sécurité et de l'évaluation de la fiabilité du réseau. Dans cette thèse, on n'aborde que le volet de la stabilité.

3.3 Fondements théoriques de la stabilité des réseaux électriques

Dans cette section, on aborde les questions fondamentales liées à la définition de la stabilité du réseau électrique du point de vue théorique. L'étude classique de la stabilité d'un réseau électrique consiste à :

- a. Énoncer des hypothèses de modélisation et de formuler un modèle mathématique approprié pour l'étude des phénomènes dans le temps ;
- b. Sélectionner une définition convenable de la stabilité ;
- c. Simuler puis analyser la stabilité pour un ensemble de contingences ;
- d. Évaluer les résultats en fonction des hypothèses, puis les comparer avec la réalité.

3.3.1 Définition mathématique de la stabilité d'un système dynamique

On suppose que le modèle du réseau électrique est défini par un modèle d'état décrit par un ensemble d'équations différentielles de premier ordre condensé dans l'équation 3.1 :

$$\dot{x} = f(x, u) \quad (3.1)$$

On considère ce modèle d'état à partir d'un instant initial noté t_0 . L'état $x \in \mathbb{R}^n$ et l'entrée $u \in \mathbb{R}^m$ sont des fonctions vectorielles du temps que l'on note généralement $x(t)$ et $u(t)$. On omet dans cette thèse l'écriture de l'argument t sans risque de confusion. La dérivée de x est notée \dot{x} .

Un point x^* est un point d'équilibre si l'état du système à t_0 est x^* et $x(t) = x^*$ pour tout $t \geq t_0$ en absence de perturbations (Sastry, 2013). Par conséquent, on a $f(x^*, 0) = 0$ pour tout $t \geq t_0$. À l'aide d'un changement de variables, on peut déplacer l'origine pour la ramener au point d'équilibre :

$$\dot{x} = x^* + x_1 \quad (3.2)$$

D'où au point d'équilibre, on a :

$$x = 0 \quad (3.3)$$

Par définition, le point d'équilibre $x = 0$ est stable si, pour un réel positif ε donné, aussi petit qu'il soit, on peut choisir un autre réel $\delta(\varepsilon)$ positif également tel que (Khalil, 2002) :

$$\|x(t_0)\| < \delta(\varepsilon) \Rightarrow \|x(t)\| < \varepsilon, \forall t \geq t_0 \geq 0 \quad (3.4)$$

Sur la Figure 3.2, on décrit dans un espace vectoriel à deux dimensions le comportement des trajectoires dans le voisinage d'un point d'équilibre stable. En choisissant les conditions initiales dans un voisinage suffisamment petit d'une sphère de rayon δ , on peut forcer la trajectoire du système en tout temps $t \geq t_0$ à être située entièrement dans un cylindre donné de rayon ε . Cette stabilité est souvent dite stabilité au sens de Lyapunov.

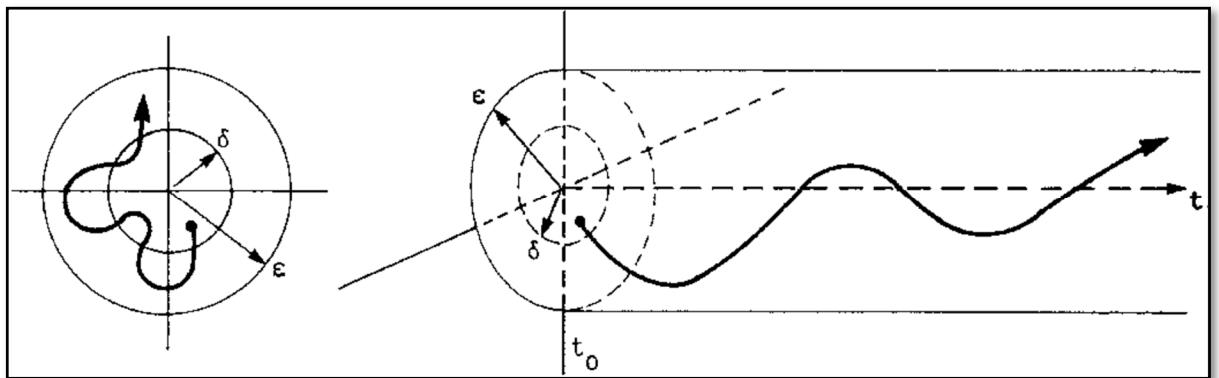


Figure 3.2 Illustration de la définition de la stabilité
Adaptée de Khalil (2002, p. 70)

Le point d'équilibre $x = 0$ est uniformément stable si, pour un réel positif ε donné, on a $\delta = \delta(\varepsilon) > 0$ indépendant de t_0 de sorte que l'équation (3.4) est satisfaite. Le point est instable s'il n'est pas stable (Khalil, 2002).

Le point d'équilibre est asymptotiquement stable si en plus d'être stable, il existe $\eta(t_0) > 0$ tel que :

$$\|x(t_0)\| < \eta(t_0) \Rightarrow x(t) \rightarrow 0 \text{ quand } t \rightarrow \infty \quad (3.5)$$

La stabilité signifie que les trajectoires ne changent pas trop sous des petites perturbations. La situation inverse, où une orbite à proximité (asymptotique) est repoussée de l'orbite donnée, est aussi intéressante. Comme montré à la Figure 3.3, de manière générale, perturber l'état initial dans certaines directions résulte dans la trajectoire approchant asymptotiquement celle donnée et dans d'autres directions à la trajectoire qui s'en éloigne.

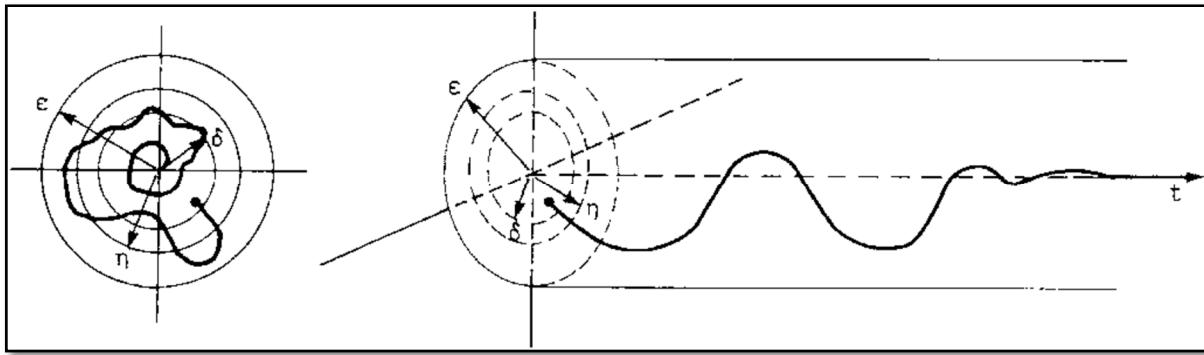


Figure 3.3 Illustration de la définition de la stabilité asymptotique
Adaptée de Fouad et Vittal (1994, p. 50)

3.3.2 Stabilité des réseaux électriques

En soumettant un réseau électrique à un défaut, on distingue trois phases distinctes (Kimbark, 1995). Une première phase qui se rapporte au moment précédent l'apparition du défaut (*prefault*) ; à cet instant, le réseau est considéré au repos. La deuxième phase est l'instant où le défaut survient (*fault-on*). On va avoir un temps d'effacement de défaut (*clearing time*, CT). La troisième phase est la phase après défaut (*post-fault*) pendant laquelle on étudie le comportement et la dynamique du réseau. Fort de ce constat, (Kundur, Balu et Lauby, 1994) propose une définition de la stabilité des réseaux électriques comme suit :

« *Power system stability may be broadly defined as that property of a power system that enables it to remain in a state of operating equilibrium under normal operating conditions and to regain an acceptable state of equilibrium after being subjected to a disturbance.* »

On distingue en outre un temps de préemption ne dépassant pas généralement 10 cycles pour remédier à la contingence et un temps critique (*critical clearing time*, CCT) au-delà duquel le réseau tombe irrévocablement instable.

3.4 Modèle mathématique lié à l'étude de la stabilité transitoire

Soit à la Figure 3.4 le modèle simplifié d'une machine synchrone représentée par une source sinusoïdale de tension E en série avec une résistance R_a et une réactance x'_d traversée par un courant I_t et branchée à une barre de tension E_t (Glover, Sarma et Overbye, 2012). Pour un réseau ayant plusieurs machines synchrones, chacune d'entre elles a exactement la même configuration vue à la Figure 3.4 (Fouad et Vittal, 1991).

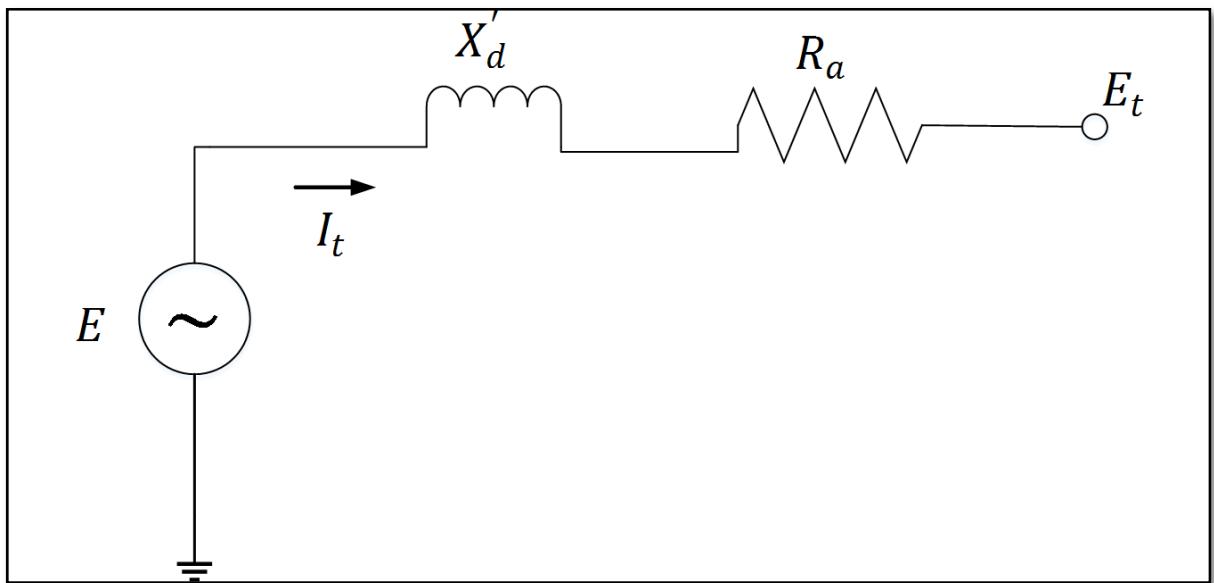


Figure 3.4 Modèle simplifié d'une machine synchrone
Adaptée de Kundur, Balu et Lauby (1994, p. 28)

Étudier la stabilité transitoire revient à résoudre les équations différentielles non linéaires représentant les machines synchrones et les équations algébriques d'écoulement de puissance représentant le réseau. L'équation (3.6) détermine la dynamique du rotor de l' $i^{\text{ème}}$ générateur du réseau. Elle est communément appelée l'équation d'oscillation per-unit.

$$2H_i \frac{d^2\delta_i(t)}{dt^2} = P_{m_{i_{p.u.}}}(t) - P_{e_{i_{p.u.}}}(t) = P_{a_{i_{p.u.}}}(t) \quad (3.6)$$

Dans cette équation, H_i désigne l'inertie normalisée de la machine, $P_{m_{i_{p.u.}}}$ la puissance mécanique fournie par le moteur et diminuée des pertes mécaniques, $P_{e_{i_{p.u.}}}$ la puissance électrique du générateur augmentée des pertes électriques et $P_{a_{i_{p.u.}}}$ la puissance d'accélération du générateur. Vu qu'il n'est pas facile de résoudre directement une équation différentielle du second ordre, cette équation est souvent décomposée en deux équations différentielles du premier ordre où ω_{syn} est la vitesse synchrone :

$$\begin{aligned} \frac{d\delta_i}{dt} &= \omega_i - \omega_{syn} \\ 2H_i \frac{d\omega_i}{dt} &= P_{m_{i_{p.u.}}} - P_{e_{i_{p.u.}}} = P_{a_{i_{p.u.}}} \end{aligned} \quad (3.7)$$

La puissance électrique de l' $i^{\text{ème}}$ générateur P_{e_i} est donnée par l'équation (3.8). Cette équation est démontrée et approfondie dans ce mémoire (Suprême, 2012). Dans cette équation, E_i désigne la force électromotrice dans la réactance de l' $i^{\text{ème}}$ générateur, G_{ij} et B_{ij} respectivement, la susceptance (partie réelle) et la conductance (partie imaginaire) de la matrice d'admittance réduite.

$$P_{e_i} = E_i^2 G_{ii} + \sum_{j=1}^n E_i E_j [B_{ij} \sin(\delta_i - \delta_j) + G_{ij} \cos(\delta_i - \delta_j)] \quad (3.8)$$

3.5 Modes d'opération des réseaux électriques

L'évaluation de la stabilité transitoire est cruciale dans la conception, la gestion et la planification des réseaux électriques (Pavella, Ernst et Ruiz-Vega, 2012). L'objectif est de savoir si après une contingence sur le réseau, les machines vont atteindre la fréquence synchrone pour un point d'équilibre stable de nouveaux angles de puissance, de nouvelles tensions aux jeux de barres et avec une nouvelle répartition de la puissance sans perdre le synchronisme (Glover, Sarma et Overbye, 2012). Une bonne évaluation de la stabilité transitoire passe inévitablement par l'identification du mode d'opération du réseau (Momoh, 2012) – (Grigsby, 2016) : l'état de prévention, l'état d'alerte et le mode curatif. Le choix de la

méthode d'analyse de la stabilité transitoire est fortement lié au mode d'opération. L'état de prévention réseau constitue l'état normal d'opération du réseau. En mode préventif, la demande des clients ou des utilisateurs est totalement satisfaite. L'objectif du contrôle effectué ou à effectuer en mode préventif est de continuer à satisfaire la demande sans interruption de services et ceci à moindres coûts. Ce mode d'opération est plutôt routinier, connu et bien maîtrisé par les opérateurs de réseaux. Des contingences pouvant provoquer une instabilité dans de telles conditions sont très rares. En cas de petits défauts, les opérateurs savent comment gérer l'équilibre. Le mode d'opération d'urgence est constaté quand des équipements de protection sont sérieusement sollicités ou endommagés. La tension ne peut être plus maintenue à une valeur raisonnable ou la fréquence diminue. Bref, on est en situation de pertes de synchronisme. L'objectif des actions de contrôle dans un tel cas est de remédier à cette perte en s'assurant d'une satisfaction maximale de la demande. Les considérations économiques sont secondaires. En ce qui a trait à l'état de réparation ou au mode curatif, on arrive à une situation de non-retour où certains consommateurs ne sont plus desservis. L'objectif est de passer d'un état de fonctionnement partiel à une satisfaction totale de la demande en minimisant le temps de réaction. Le temps de non-service doit être le plus petit que possible.

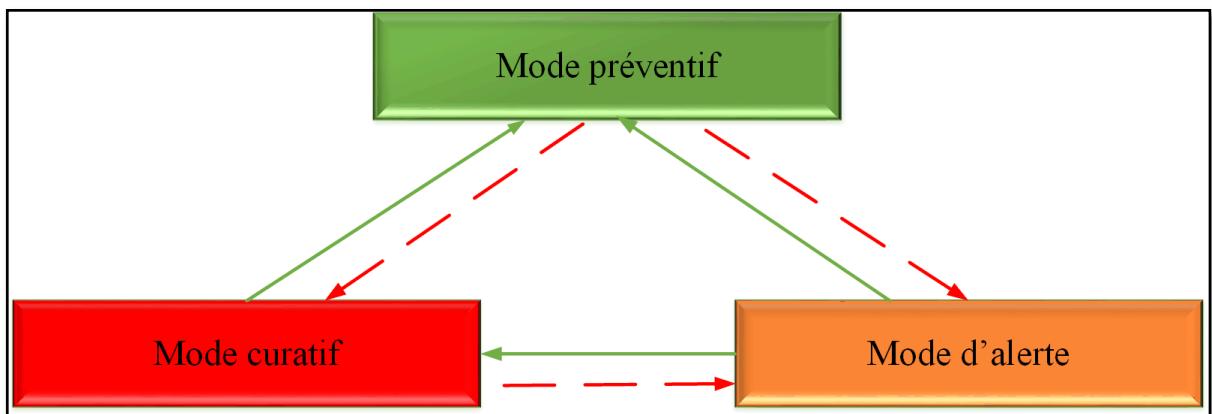


Figure 3.5 Diagramme d'état de la sécurité du réseau électrique de DyLiacco
Adaptée de Momoh (2012, p. 43)

Sur la Figure 3.5, on présente le diagramme d'état de la sécurité de DyLiacco relatif aux différents états de fonctionnement du réseau électrique. Les flèches pleines représentent un changement de mode d'opération dû à une action de contrôle prévu alors que les flèches en

pointillés indiquent un changement de mode de fonctionnement du réseau résultant d'imprévus ou d'actions de contrôle incorrects. L'objectif est de maintenir le réseau dans un mode de fonctionnement au préventif. Comme indiqué par les flèches, en mode préventif, une exploration du réseau peut être faite. On peut procéder à une analyse d'éventuelles contingences et créer une base de données avec ces informations pertinentes. On maximise du point de vue économique et on minimise les effets des contingences incertaines. En état d'alerte, on va toujours essayer de ramener le réseau à un mode préventif. On cherche toujours un compromis entre le mode préventif et les actions de contrôle potentielles. Certaines complexités peuvent se retrouver sur le réseau et faire en sorte qu'on ne soit plus en état d'alerte mais en mode curatif. On a pour mission et devoir de regagner l'état normal de fonctionnement du réseau.

3.6 Conclusion

Dans ce chapitre, on s'est penché sur la terminologie et le vocabulaire utilisés dans l'analyse et l'évaluation de la stabilité des réseaux électriques. On a défini la notion de stabilité dans un cadre général et aussi dans le contexte précis des réseaux électriques. Les différentes ramifications de la stabilité sont abordées. On a établi nettement la différence entre la sécurité, la fiabilité et la stabilité des réseaux électriques. Des concepts mathématiques et théoriques des machines synchrones liés à l'étude de la stabilité transitoire sont aussi fournis. Dans la dernière section, on a adressé les différents modes d'opération des réseaux électriques qui jouent un rôle important dans l'étude de la stabilité.

CHAPITRE 4

VUE D'ENSEMBLE DES SYNCROPHASEURS

4.1 Introduction

En 1893, Charles Proteus Steinmetz a présenté un document sur la description mathématique simplifiée des formes d'onde d'électricité en courant alternatif. Steinmetz a appelé sa représentation un phasor (Steinmetz, 1893). En s'inspirant des travaux de Steinmetz, Arun G. Phadke et James S. Thorp à Virginia Tech ont utilisé en 1988 la technique de calculs des phasors pour introduire la notion des PMUs sur les réseaux électriques. Les PMUs sont des dispositifs qui mesurent en temps réel les ondes électriques sur un réseau électrique. Ils sont synchronisés à une référence de temps absolu provenant du système de positionnement global (GPS, *Global Positioning System*). L'évolution des travaux sur cette technologie est relatée dans cet article (Phadke, 2002). Les PMUs sont considérés comme l'un des appareils de mesure les plus importants dans l'avenir des réseaux électriques (Delivery, 2013). Conscient de cette place que peut occuper cette technologie dans le futur, un premier document préparatoire de standard a été écrit en 1995 (IEEE Standard for Synchrophasors for Power Systems, 1995) et confirmé puis voté en 2005 (IEEE Standard for Synchrophasors for Power Systems, 2006). Un comité de travail formé en janvier 2011 a réalisé une mise à jour complète (Martin, 2011) de ce dernier en édifiant plus clairement les principes à respecter en utilisant des PMUs sur les réseaux électriques. Autre que la définition des différents concepts, il fournit une méthode de quantification des mesures et des tests de spécification de qualité à réaliser. Il définit également le format de transmission de données pour leur communication en temps réel. Cette dernière version est divisée en deux parties : la première couvrant la fourniture des mesures (IEEE Standard for Synchrophasor Measurements for Power Systems, 2011) et la seconde couvrant la communication des données (IEEE Standard for Synchrophasor Data Transfer for Power Systems, 2011). Dans l'ingénierie de puissance, les mesures obtenues des PMUs sont communément appelées des « synchrophaseurs ».

Dans ce chapitre, on expose une vue globale de cette technologie en tenant compte de ses avantages et de ses applications courantes sur les réseaux électriques.

4.2 Mesure des synchrophaseurs

Un phaseur est un nombre complexe qui représente à la fois l'amplitude et l'angle de phase des ondes sinusoïdales. Considérons la fonction sinusoïdale du signal de l'équation (4.1) :

$$x(t) = X_m \cos(\omega t + \phi) \quad (4.1)$$

Le phaseur correspondant est de la forme de l'équation (4.2) :

$$X \equiv \frac{X_m}{\sqrt{2}} e^{j\phi} = \frac{X_m}{\sqrt{2}} (\cos(\phi) + j \sin(\phi)) \quad (4.2)$$

Où l'amplitude est la valeur efficace du sinusoïde $\frac{X_m}{\sqrt{2}}$ et ϕ l'angle de phase. La valeur de ϕ dépend du temps, particulièrement de l'instant initial $t = 0$. Il est important de souligner que le phaseur est défini pour une fréquence angulaire ω ; l'estimation d'autres phaseurs doit être faite au même temps et à la même fréquence.

La représentation de synchrophaseur du signal $x(t)$ dans l'équation (4.1) est la valeur X dans l'équation (4.2) où ϕ est l'angle de phase instantané par rapport à une fonction cosinus de la fréquence nominale du réseau synchronisé avec le temps universel coordonné (*Coordinated Universal Time, UTC*). Suivant cette définition, ϕ est le décalage d'une fonction cosinus à la fréquence nominale du système synchronisé au temps universel. En utilisant des horloges synchronisées en tant que référence, par convention, la représentation de vecteur de phase d'un signal sinusoïdal constant par un PMU se fait tel qu'indiqué sur la Figure 4.1 et l'angle est noté positif dans le sens de l'aiguille d'une montre. On reproduit sur la Figure 4.1 (a) une sinusoïde de période T à des instants multiples de T_0 , où T_0 n'est pas un multiple de T .

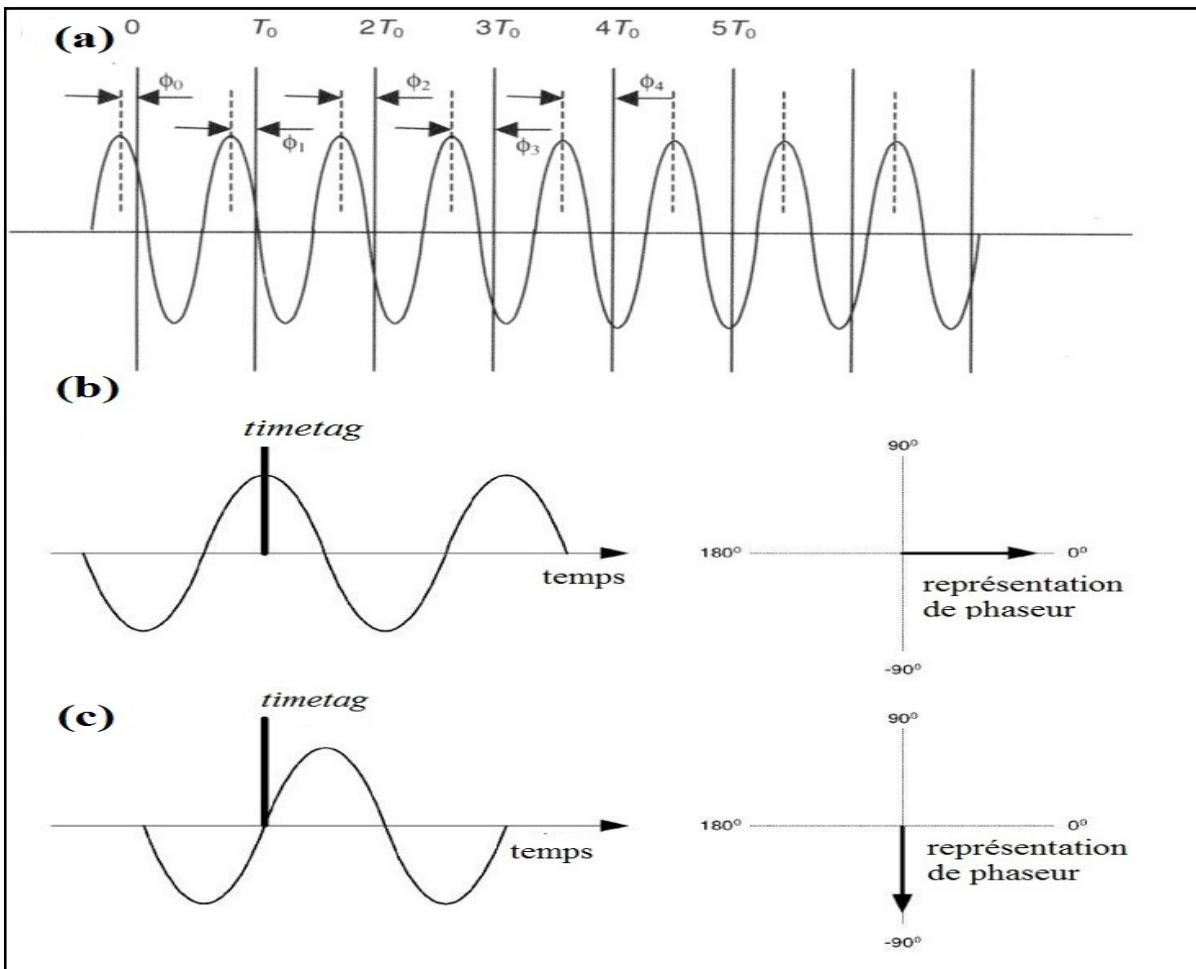


Figure 4.1 Convention de la représentation d'un synchrophasor
Adaptée de IEEE Standard for Synchrophasors for Power Systems (2006, p.7)

L'instant témoin (*timetag*) définit la référence pour la représentation de phasor de l'entrée sinusoïdale mesurée. La relation entre l'instant témoin et la représentation de vecteur de phase est telle que l'angle de phase du vecteur de phase soit égal à l'espacement angulaire entre l'instant témoin et le sommet de la sinusoïde. Par exemple, dans la Figure 4.1 (b), la crête du signal coïncide avec le *timetag* donnant la mesure de l'angle de 0° , tandis qu'à la Figure 4.1 (c), le signal passe par zéro au *timetag* donnant une mesure de -90° conforme à la norme du synchrophasor. L'angle de phasor pris seul n'a aucune signification. Toutefois, si tous les PMUs utilisent la même référence de temps, leurs mesures sont comparables, et les différences d'angle de phasor entre les phasors ont un sens exact.

4.3 Réseau de communication et placement des PMUs

Un système de mesure de zone étendue (*Wide Area Measurement System*, WAMS) est une infrastructure basée sur les nouvelles technologies de pointe qui permet d'exploiter les réseaux électriques dans leur capacité optimale (Momoh, 2012). En fournissant des données de mesure de phasor en temps réel, le WAMS facilite la surveillance, la compréhension et la gestion de grands réseaux électriques. Les PMUs constituent la composante principale d'un WAMS. Ils ont l'avantage de fournir de 30 jusqu'à 120 mesures de phasors par seconde (Ma, Makarov et Dong, 2010) avec une précision d'une microseconde ; ce qui permet une bonne observabilité et une vue globale du réseau (Borlase, 2013).

Comme montré à la Figure 4.2, les données de phasors synchronisées de courant, de tension et de fréquence prises par les PMUs branchés à des jeux de barres du réseau sont envoyées dans un concentrateur de données de phasor (*Phasor Data Concentrator*, PDC), situé dans des sous-stations ou des centres de contrôle en utilisant un réseau de communication en fibre optique dédiée. Toutes les informations sont stockées dans un centre principal (Super PDC) où on les traite et les interprète afin de détecter d'éventuelles anomalies sur le réseau.

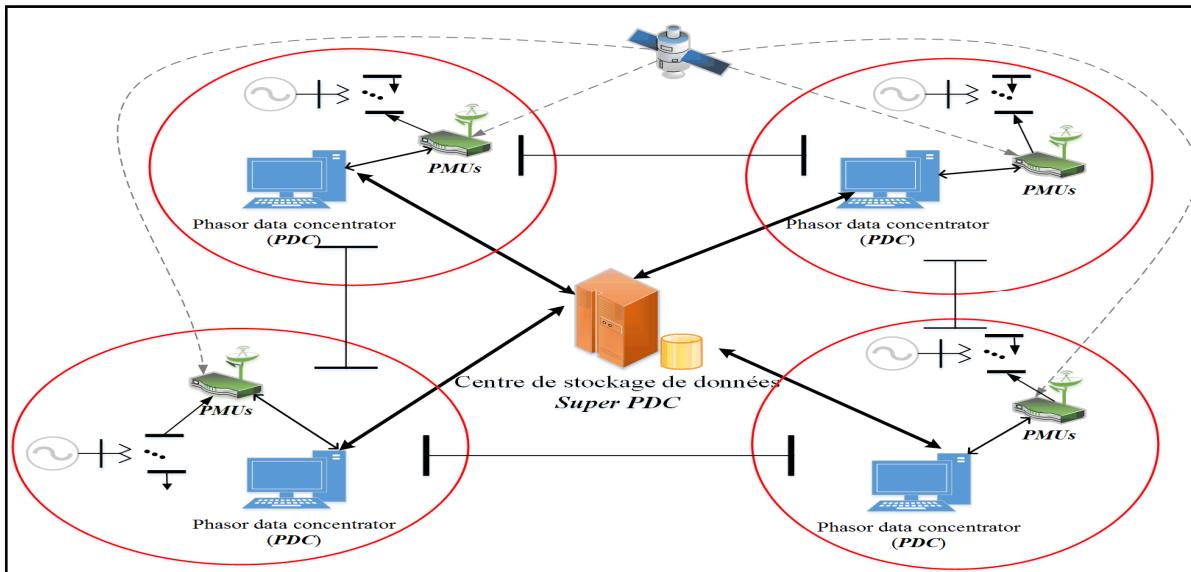


Figure 4.2 Diagramme conceptuel d'un système de mesures de phasors synchronisées

Les PMUs sont des instruments de mesure très coûteux dépendamment de l'utilisation qu'on veut en faire (NASPI, 2007) – (DOE, 2014). Plusieurs facteurs peuvent influer significativement sur leurs coûts d'acquisition et d'installation. Les plus importants, selon plusieurs opérateurs de réseaux, sont les installations des systèmes de communications, incluant leurs mises à jour fréquentes qui peuvent faire varier les coûts d'un facteur de 7 ; et les équipements assurant la cybersécurité (DOE, 2014). Il s'avère donc important de placer les PMUs à des endroits stratégiques du réseau. La formulation générale du problème de placement optimal des PMUs (*Optimal PMU Placement*, OPP) consiste en la minimisation du nombre de PMUs, noté n_{PMU} , et en leur localisation optimale, notée $S(n_{PMU})$, en veillant à ce que le réseau reste entièrement observable (Brueni et Heath, 2005). D'autres contraintes peuvent s'y ajouter pour avoir un modèle plus général comme stipulé dans l'équation (4.3) :

$$\min_{n_{PMU}} \left\{ \min_{S(n_{PMU})} G(n_{PMU}, S(n_{PMU})) \right\} \quad (4.3)$$

pour $f(n_{PMU}, S(n_{PMU})) = 1$

Où $G(n_{PMU}, S(n_{PMU}))$ est le nombre de jeux de barres non observables et $f(n_{PMU}, S(n_{PMU}))$ est une fonction logique d'évaluation multi-objectif donnée par l'équation 4.4 :

$$\begin{aligned} f(n_{PMU}, S(n_{PMU})) &= Obs(n_{PMU}, S(n_{PMU})) \quad (4.4) \\ &\wedge Obs(n_{PMU}, S(n_{PMU})) \vee Con(n_{PMU}, S(n_{PMU})) \end{aligned}$$

$Obs(\dots)$ est la fonction logique d'évaluation de l'observabilité du réseau, $Con(\dots)$ la fonction logique d'évaluation des contraintes, \wedge et \vee , respectivement les opérateurs logiques **ET** et **OU**. Comme $f(\dots)$ est un problème de satisfaisabilité d'une formule booléenne, il s'agit d'un problème NP-complet (Portier, 1998). Le problème n'a donc pas une solution unique. Le Tableau 4.1 regroupe plusieurs formulations de l'OPP avec des contraintes additionnelles et des publications y afférentes.

Tableau 4.1 Formulation des contraintes de placement optimal des PMUs

Formulations des contraintes	Publications
<i>Effect of Zero-Injection Buses</i>	(Aminifar et al., 2010) – (Gou, 2008)
<i>Effect of Conventional Measurements</i>	(Jiang, Vittal et Heydt, 2007) – (Kim et Poor, 2011)
<i>Single or Multiple Loss Contingency</i>	(Nuqui et al., 2001) – (Rakpenthai et al., 2007)
<i>Single Branch Outage Contingency</i>	(Dua et al., 2008) – (Kamwa et Grondin, 2002a)
<i>Contingency of Single Line Outage or Single PMU Loss</i>	(Emami et Abur, 2010)
<i>Effect of PMU Channel Capacity</i>	(Rakpenthai et al., 2007) – (Korkali et Abur, 2010)

Pour résoudre le problème OPP, deux méthodes principales sont privilégiées dans la littérature : 1) les modèles mathématiques qui cherchent à résoudre l'équation (4.4) en se référant à l'une des formulations précédentes et 2) les approches heuristiques qui tiennent compte, entre autres, de l'expérience, de l'apprentissage et des jugements d'experts responsables de l'ingénierie du réseau. Pour les modèles mathématiques, on peut procéder à une recherche exhaustive qui permet d'énumérer toutes les solutions possibles et de sélectionner celle qui satisfait aux contraintes de la fonction objectif de manière optimale en se basant sur la redondance des mesures (Chakrabarti et Kyriakides, 2008). On peut aussi procéder à une programmation en nombre entier (*Integer Programming*, IP) (Castillo et al., 2011) pour laquelle on considère des variables entières pour une fonction objectif et des contraintes linéaires (*Integer Linear Programming*, ILP) (Emami et Abur, 2010) – (Abbasy et Ismail, 2009), non-linéaires (*Integer NonLinear Programming*, INLP) (Theodorakatos, Manousakis et Korres, 2015) et quadratiques (*Integer Quadratic Programming*, IQP) (Chakrabarti, Kyriakides et Eliades, 2009). Parmi les algorithmes heuristiques, on peut citer : ceux se basant sur la recherche tabou (*Tabu Search*, TS) qui est une méthode d'optimisation gradient décroissante avec mémoire (Lee et El-Sharkawi, 2008) ; ceux qui se basent sur des

algorithmes génétiques où on incorpore des fonctions objectif concurrentes comme la minimisation du nombre de PMUs et la maximisation de la redondance des mesures (Milosevic et Begovic, 2003) ; ceux se basant sur un algorithme glouton (*Greedy Algorithm*) qui suit le principe de faire étape par étape un choix d'optimum local en réduisant le nombre de PMUs à placer jusqu'à arriver un optimum global (Kim et Poor, 2011) – (Kamwa et Grondin, 2002b).

4.4 Applications des mesures synchrophasées dans les réseaux électriques

En 2005, on a lancé le *North American SynchroPhasor Initiative* (NASPI) qui a pour vision d'améliorer la surveillance, le contrôle et la fiabilité des réseaux électriques à travers les WAMS. Sa mission est de créer une infrastructure de PMUs robuste, sécurisée et très facile d'accès pour le réseau électrique interconnecté nord-américain (NASPI, 2016). Ainsi, les opérateurs de réseaux seront dotés d'outils d'analyse et de planification puissants dans la réalisation de leurs tâches. En s'appuyant sur la mission et la vision de NASPI, on est capable d'affirmer que les mesures synchrophasées peuvent être utilisées pour réaliser des tâches en temps réel (*online*) ou hors-ligne (*offline*). Comme suggéré à la Figure 4.3, on divise en trois grandes classes les différentes applications des PMUs : la surveillance, le contrôle et la protection. On soumet également une liste non exhaustive d'applications pour la prise de décision opérationnelle sur les réseaux électriques.

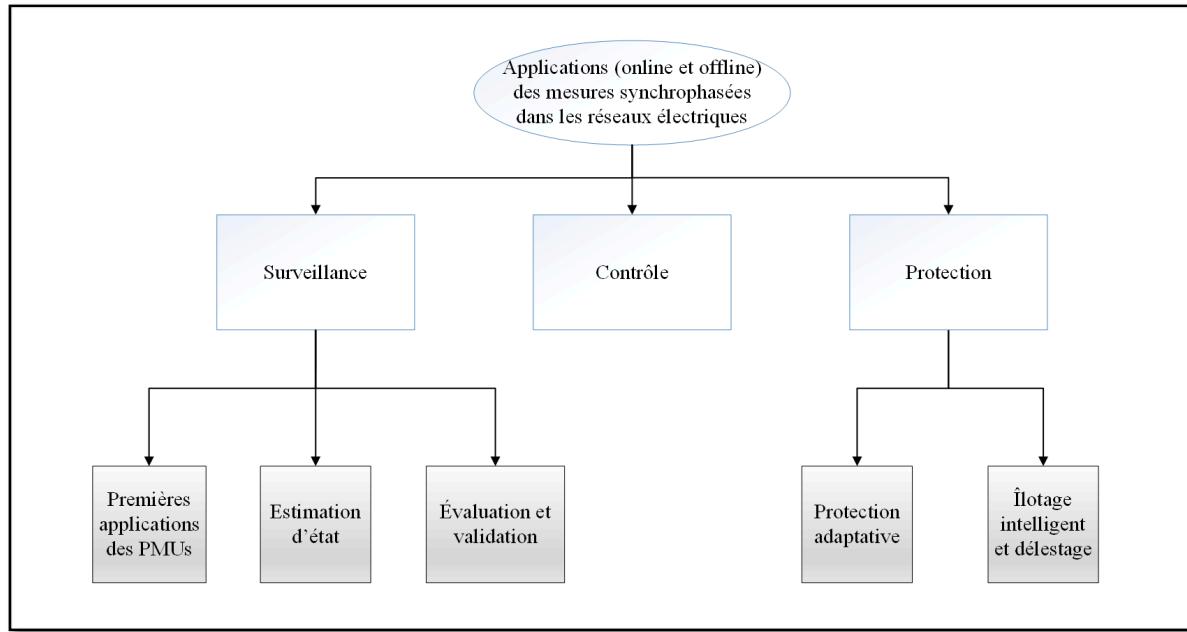


Figure 4.3 Applications des PMUs dans les réseaux électriques

Surveillance des réseaux électriques

A. Premières applications des PMUs

Avec la première génération de PMUs commercialisée, il était possible de faire uniquement des analyses post-défaits, vu qu'un nombre très restreint était installé sur le réseau. Les premières mesures enregistrées par des PMUs dans un WAMS ont été réalisées en 1992 dans le cadre du projet *Parameter Identification Data Acquisition System* (PIDAS) lancé par l'*Electric Power Research Institute* (EPRI). Les tests consistaient en l'ouverture et la fermeture des lignes de 500 kV de Klondike et Bonaire dans la station Scherer en Géorgie. Des PMUs déclenchés manuellement ont été placés à ladite station et à 5 autres stations en Géorgie, à Floride et au Tennessee (Burnett et al., 1994). Des résultats positifs ont permis de valider le modèle du système en comparant les valeurs mesurées par les PMUs aux valeurs estimées. Suite à ces tests, on a procédé à un plus grand déploiement sur le réseau américain. Ces derniers fonctionnaient comme un enregistreur digital de perturbations. Avec le temps, on a ajouté des fonctionnalités de protection et de contrôle. Lors des grandes pannes

survenues sur la côte ouest américaine en 1996 et au nord-est américain en 2003, les données des PMUs ont été déterminantes dans l'analyse post-mortem rapide et précise des défauts (De La Ree et al., 2010). En effet, l'une des recommandations de la *United States – Canada Task Force* (DOE et NRC) sur la panne de 2003 a été de :

« *R28. Require use of time-synchronized data recorders.*

In its requirements of February 10, 2004, NERC (North American Electric Reliability Corporation, n.d.a.) directed the Regional Reliability Councils to define, within one year, regional criteria for the application of synchronized recording devices in key power plants and substations. The Task Force supported this recommendation strongly, but recommended a broader approach with four components. »

B. Estimation d'état

L'estimation d'état est une fonction d'application très courante dans les centres de contrôle et de conduite des réseaux électriques (*Energy Management System*, EMS). Plusieurs algorithmes d'estimateurs d'état (Phadke et Thorp, 2008) utilisent les écoulements de puissance (puissance active et réactive) et les injections de puissance sur les lignes fournis par le *Supervisory Control and Data Acquisition* (SCADA) pour estimer tous les angles et les amplitudes de tension de jeux de barres du réseau. Avant les PMUs, on avait des mesures non synchronisées qui ne relataient pas forcément l'état actuel et réel du réseau. De manière générale, l'estimation d'état est un problème itératif non linéaire résolu en utilisant la technique des moindres carrés pondérés (Smith et al., 1991). L'avènement de la technologie des mesures synchrophasées a le potentiel de transformer le problème itératif d'estimation d'état non linéaire en une simple résolution linéaire en considérant à la fois les mesures des PMUs et les données du SCADA (Zhou et al., 2006). Les PMUs ont donc permis à améliorer la qualité des mesures et la performance des estimateurs d'état. Toutefois, la problématique de la certitude (confiance vs fiabilité) des données des PMUs est toujours posée dans la littérature (Bernabeu, Thorp et Centeno, 2012) – (Horowitz, Phadke et Thorpe, 1988).

C. Évaluation et validation

L'implémentation des PMUs donne la possibilité d'identifier les erreurs dans les paramètres de modélisation du réseau et de calculer des estimations améliorées (Zhang et al., 2015) – (Nayak et al., 2016). Les données des PMUs étant réputées plus précises, on peut s'en servir pour déterminer par exemple les matrices d'admittance, d'impédances et les données de prises exactes des transformateurs. D'ailleurs, en Europe, une telle application est commercialisée pour calculer l'impédance des lignes de transmission (Zamora-C et al., 2012).

Contrôle des réseaux électriques

Avant l'introduction des PMUs, le contrôle se faisait localement sur les réseaux. Les générateurs, par exemple, étaient contrôlés uniquement par des signaux locaux. Grâce aux PMUs, la prise de décision opérationnelle se fait instantanément et de manière globale en intégrant des données complètes du WAMS. On peut plus facilement s'attaquer à des problèmes de stabilité transitoire (Suprême et al., 2016), d'oscillations électromécaniques (Kamwa, Beland et Mcnabb, 2006) et à certains phénomènes de surcharge. Par exemple, à partir des mesures, un contrôleur *High-Voltage Direct Current* (HVDC) peut être appelé à amortir les oscillations électromécaniques entre deux régions très éloignées du réseau (He et Liu, 2014) ; ou encore des stabilisateurs (*Power System Stabilizers*, PSS) peuvent améliorer la stabilité du réseau lorsqu'il est couplé à un régulateur de tension automatique (*Automatic Voltage Regulator*, AVR) (Kamwa, Samantaray et Joos, 2013) – (Kamwa, Samantaray et Joos, 2013). On peut aussi intégrer au système de protection du réseau (*System Integrity Protection Schemes*, SIPS) des contrôleurs logiques (Zhang et Bose, 2008) résultant d'un apprentissage hors-ligne des données (Kamwa, Grondin et Loud, 2001b) – (Moulin et al., 2004).

Protection des réseaux électriques

A. Protection adaptative

Évidemment, certaines pannes d'électricité sont inévitables ; il faut élaborer des stratégies pour rétablir le courant le plus rapidement possible et à coût minimal après de tels évènements. Même si des stratégies de restauration pré-calculées obtenues à partir des études de simulation existent, elles sont souvent inadéquates parce que l'état réel du système est tout à fait différent de celui supposé dans les études de planification. En temps réel, les PMUs offrent une excellente occasion de déterminer une stratégie de restauration qui prend en compte l'état actuel et réel du WAMS (Bretas et Phadke, 2003). Une application concrète des PMUs a été faite dans la restauration du réseau lors de la méga panne européenne du 4 novembre 2006 (ERGEG, 2007) – (Khoi;, Novosel; et Synder, 2007).

Il est aussi admis dans la littérature qu'un groupe de générateurs, appelé générateurs critiques (Savulescu, 2014), peut causer l'effondrement complet du réseau. Les mesures synchrophasées peuvent être utilisées pour suivre leur comportement et ipso facto alimenter des algorithmes précis en données afin d'alerter de potentielles anomalies. Dans cet ordre d'idée, au Chapitre 6 de cette thèse, on a développé une nouvelle notion et proposé de nouveaux indices de prédiction de l'instabilité en temps réel basés sur les PMUs.

B. Îlotage intelligent et délestage

L'îlotage est réalisé en utilisant les SIPS (Phadke et Thorp, 2009). De manière générale, il est conçu sur la base des données de simulation. Par conséquent, la performance peut ne pas être optimale pour l'état réel du réseau. Pour pallier à ce manque de précision dans les données, on peut faire appel aux PMUs pour déterminer plus nettement si le réseau se dirige vers un état instable (déttection et localisation de défaut) et si une séparation de réseau est nécessaire pour éviter une panne d'électricité (problématique abordée dans le cadre de cette thèse). Ce type de problème est adressé dans cet article (Kamwa, Grondin

et Loud, 2001b) où les auteurs proposent des indicateurs de classement de la sévérité des contingences. On peut aussi déterminer les limites optimales de l’îlotage en fonction des conditions actuelles du réseau. Par exemple, établir quels groupes de générateurs seront séparés en raison de la perte de synchronisme et la façon d'équilibrer de manière optimale la charge et la production dans chaque île formée par des groupes et des charges de générateurs cohérents. EPRI a développé un outil (EPRI, 2016) capable de calculer des indices de marges de stabilité de tension en temps réel pour une région donnée ou pour un jeu de barre spécifique en utilisant les données provenant des PMUs. Ceci permet entre autres de faire des études de stabilité de tension hors-ligne après certaines perturbations.

De même, il est possible de formuler une stratégie qui traiterait la question de délestage avant que la fréquence commence à se dégrader, ou avant que la tension commence une plongée vers l’instabilité (Begovic et al., 2005) – (Madani, Novosel et King, 2014).

4.5 PMU vs SCADA

Afin d'aider les opérateurs de réseau de transport à surveiller, contrôler et optimiser les performances des systèmes de production et de transport, on utilise une suite d'applications collectées dans un système de gestion de l'énergie (EMS). Les fonctions de surveillance et de contrôle pour les EMS sont souvent fournies par le SCADA. Les données PMU des WAMS coexistent en parallèle avec les données de l'EMS/SCADA comme le montre la Figure 4.4 (Ekanayake, 2012). Les données PMUs peuvent être utilisées dans les applications EMS pour améliorer la surveillance du réseau.

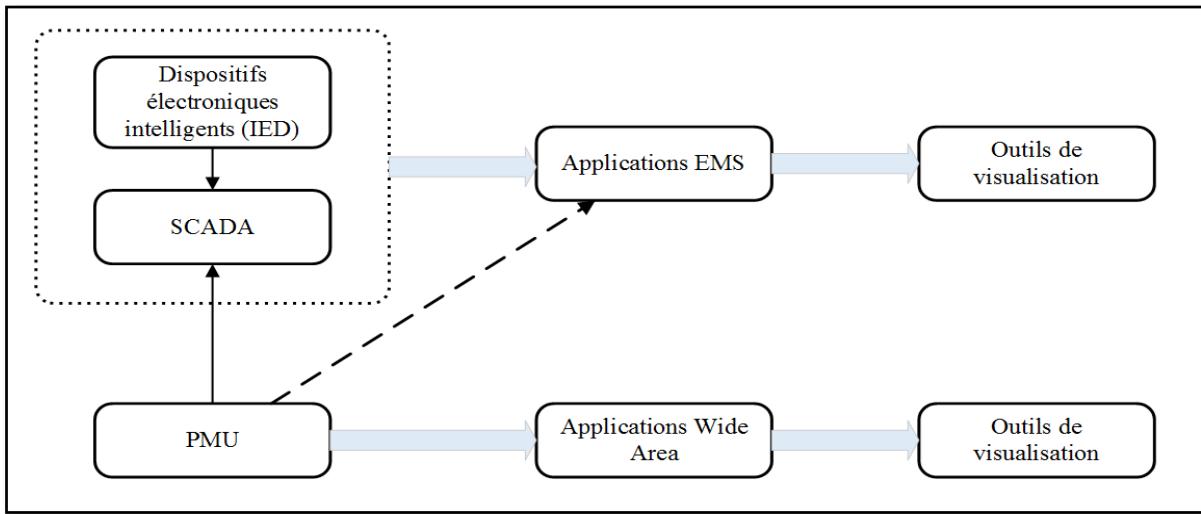


Figure 4.4 Représentation de la structure d'un EMS/SCADA et d'un WAMS
Adaptée de Ekanayake (2012, p. 174)

Le SCADA mesure l'amplitude des tensions mais ne tient pas compte des angles de phase correspondants (NASPI, 2016). Pour les courants alternatifs, ces angles sont importants à connaître dans la mesure où ils permettent de faire des calculs d'écoulement de puissance et de limites de stabilité. L'EMS/SCADA se base sur l'analyse d'écoulement de puissance quasi-constant, ce qui ne correspond pas fidèlement à la réalité ; et de ce fait, ne favorise pas la surveillance de phénomènes transitoires. En outre, l'un des principaux problèmes avec l'estimation de l'état actuel de l'EMS/SCADA et des fonctions connexes telles que l'analyse de contingence est la nécessité d'un modèle pour les mesures non locales. Ceci est encore un problème non résolu, même avec un grand centre de communication. Les PMUs peuvent aider à surmonter en grande partie ce problème puisque la disponibilité de mesures de phaseurs synchronisées à partir des zones globales peut faire la modélisation du réseau global de manière plus précise et plus simple en adoptant un modèle d'ordre réduit. De plus, étant donné que les mesures de phaseurs donnent des informations sur un laps de temps de l'ordre des millisecondes, ils peuvent aider à la détection rapide de la séparation du système et les pannes possibles (Savulescu, 2009). Le Tableau 4.2 résume la comparaison entre l'EMS/SCADA et les PMUs.

Tableau 4.2 Comparaison entre EMS/SCADA et PMU

Attributs	EMS/SCADA	PMU
<i>Mesures</i>	Analogiques	Digitales
<i>Résolution</i>	2 – 4 échantillons par cycle	Plus de 60 échantillons par cycle
<i>Observabilité</i>	Régime permanent	Régime dynamique et transitoire
<i>Surveillance</i>	Local	<i>Wide-Area</i>
<i>Mesures des angles de phase</i>	Non	Oui

Les données des PMUs sont archivées et peuvent également être envoyées à la base de données EMS/SCADA comme mentionné dans (Atanackovic et al., 2008) – (Novosel et al., 2008). La référence (Zhou et al., 2006) expose une méthode intéressante pour inclure les données de phasateurs dans l'EMS/SCADA. Malgré les nombreux avantages des PMUs par rapport au SCADA, par mesure de précaution, la sauvegarde des données SCADA est fortement recommandée (Patel, 2010). En effet, une diversité de sources de données aide à préserver la continuité du service au cas où l'une d'entre elles serait endommagée.

4.6 Conclusion

Dans ce chapitre, on a présenté la technologie des synchrophaseurs ainsi que ses principales applications. La quantité de données fournies par l'EMS/SCADA est restreinte et les analyses peuvent être faites uniquement localement. Grâce aux mesures synchronisées, les PMUs permettent d'avoir une vue globale du comportement du réseau en temps réel. Ils ne sont pas bénéfiques uniquement pour l'étude de la stabilité transitoire ; ils permettent de détecter d'éventuelles séparations au niveau du réseau en temps réel et d'analyser en boucle fermée les oscillations des machines afin de sauvegarder la stabilité dudit réseau. Les techniciens et les ingénieurs sont donc mieux outillés dans leur planification pour détecter, anticiper et corriger certaines anomalies sur le réseau suite à des contingences sévères. L'analyse post-défaut devient plus rapide et complète. Le principal obstacle de cette technologie demeure les coûts

d'investissement qui sont très élevés pour son déploiement en un premier temps et pour sa maintenance par la suite. Toutefois, les grands opérateurs de réseaux électriques à travers le monde croient que les PMUs constituent l'avenir et qu'il faut minimalement les placer aux endroits stratégiques du réseau pour un bon contrôle, une meilleure surveillance et protection.

CHAPITRE 5

FORAGE DE DONNÉES : FORÊT ALÉATOIRE D'ARBRES DE DÉCISION ET BOOSTING

5.1 Introduction

Dans les centres de contrôle des réseaux électriques, on est submergé par le volume de données collectées et stockées. Dans la base de données, plusieurs variables interviennent : les variables explicatives ou prédicteurs qui sont les variables utilisées pour construire le modèle et la variable cible qui est la variable dont on veut connaître la valeur. Les premières sont utilisées « en entrée » alors que la seconde est une « variable de sortie » attribuée à l'état du réseau stable ou instable dans cette thèse-ci. Les ingénieurs et opérateurs doivent être en mesure d'extraire, d'analyser et de restituer les informations pertinentes de ces données pour avoir une vision complète et unifiée du réseau ; ce qui, évidemment, facilite l'aide à la décision. Une telle aide passe par la compréhension du comportement réel et de sa modélisation. Le forage de données propose des solutions à des problèmes de ce type.

Par définition, le forage de données (*data mining*) (Tufféry, 2005) « est l'ensemble des méthodes et techniques destinées à l'exploration et l'analyse de (souvent grandes) bases de données informatiques, de façon automatique ou semi-automatique, en vue de détecter dans ces données des règles, des associations, des tendances inconnues ou cachées, des structures particulières restituant l'essentiel de l'information utile tout en réduisant la quantité de données. » Pour bien comprendre l'enjeu du forage de données, on peut se référer à la réponse donnée par Dr Arno Penzias, lauréat du prix Nobel de Physique en 1978 et ancien scientifique en chef chez Bell Labs, suite à la question :

« *What will be the killer applications in the corporation?* »

« *Data mining will become much more important and companies will throw away nothing about their customers because it will be so valuable. If you are not doing this, you are out of business* »

Dans la pratique, le forage de données est de deux ordres : soit descriptif, soit prédictif (Fayyad, Piatetsky-Shapiro et Smyth, 1996). Le forage de données descriptif se concentre sur la recherche de modèles décrivant les données naturellement interprétables par l'humain alors que le forage de données prédictif consiste à utiliser des variables dans la base de données pour prédire les valeurs inconnues des autres variables d'intérêt. La prédiction ne doit surtout pas être confondue avec la prévision, laquelle consiste à évaluer la valeur d'une variable à l'instant t sachant ses valeurs aux instants précédents.

Dans ce chapitre, on fait ressortir la différence entre les deux grandes catégoriques de techniques de prédiction. Par la suite, on décrit les processus d'obtention d'un arbre de décision, d'une forêt aléatoire d'arbres de décision et d'un boosting. Finalement, on expose le logiciel utilisé dans le cadre de l'implémentation de la forêt aléatoire et du boosting pour atteindre les objectifs de recherche.

5.2 Techniques inductives versus transductives

Les techniques de forage de données sont plus complexes que la statistique descriptive élémentaire (Aggarwal, 2015). Elles s'appuient sur des outils d'intelligence artificielle, la théorie de l'information, la théorie de l'apprentissage, la statistique inférentielle et l'analyse de données classique (analyse factorielle, classification automatique, analyse discriminante, etc...). Les techniques de prédiction sont de deux types : inductives et transductives. Les techniques transductives comprennent une seule étape (potentiellement réitérée) au cours de laquelle chaque variable est directement prédite par référence aux autres variables déjà prédites. À priori, il n'y a ni élaboration d'un modèle ni détermination de paramètres (Tuffery et Riesco, 2011). Par contre, pour les techniques inductives, une phase d'apprentissage, qui est une phase inductive, permet d'élaborer un modèle à partir de données historiques. Ce dernier résume les relations entre les variables et peut être ensuite appliqué à de nouvelles données pour en déduire une prédiction (phase déductive). La Figure 5.1 résume le procédé de ces deux méthodes.

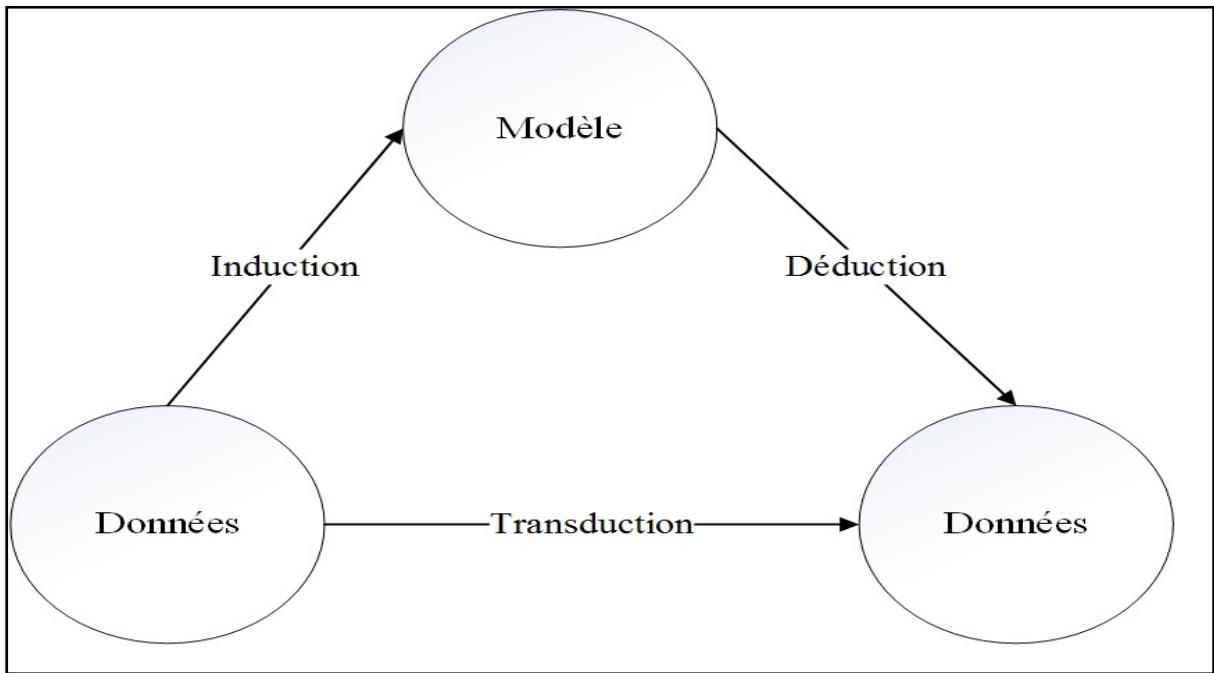


Figure 5.1 Méthodes inductives et transductives

Adaptée de Tuffery et Riesco (2011, p. 304)

Les deux méthodes transductives les plus connues sont la méthode de Parzen-Rosenblatt ou méthode de l'estimation par noyau et la méthode des k -plus proches voisins (*k-Nearest Neighbor*, KNN). L'estimation par noyau est une méthode non-paramétrique d'estimation de la densité de probabilité d'une variable aléatoire se basant sur un échantillon d'une population statistique et permettant d'estimer la densité en tout point du support (Simonoff, 2012) – (Devroye et Gyorfi, 1985). Comme avantages de cette méthode, on peut souligner qu'aucune supposition quant à la distribution des données n'est requise et que la convergence est assurée si le nombre d'exemples est suffisant (Simonoff, 2012) – (Silverman, 1986). L'inconvénient essentiel est que le modèle croît proportionnellement au nombre de variables (Duda, Hart et Stork, 2001). Le KNN est une méthode dans laquelle chaque variable est prédite en regardant parmi les variables déjà prédites, la classe des k variables qui sont ses plus proches voisins au sens de la distance euclidienne ou d'une autre (Witten, Hall et Frank, 2011). Cet algorithme est intéressant pour un faible nombre de variables où chaque classe est représentée par plusieurs prototypes et quand la répartition des classes dans l'espace euclidien est irrégulière (par exemple, reconnaissance de chiffres manuscrits ou d'images satellites) (Chirici et al., 2016).

Cependant, étant une méthode gourmande en place mémoire et sensible aux variables non pertinentes et corrélées, le KNN présente comme désavantage la manipulation d'un certain nombre de variables déjà prédites pour toute nouvelle prédiction (Keller, Gray et Givens, 1985).

Puisqu'il est plus simple de résumer dans un modèle l'information contenue dans les données pour pouvoir contrôler et appliquer rapidement ce modèle à de nouvelles données, les méthodes inductives sont généralement préférées aux méthodes transductives. L'algorithme suivant indique comment se déroule la prédiction dans une méthode inductive (Stéphane, 2012) :

Algorithme 5.1 Différentes étapes d'un algorithme de prédiction

1. Apprentissage : Étape effectuée avec un échantillon de variables tirées aléatoirement dans la base de données à modéliser et dont on connaît la prédiction ;
2. Test : Étape réalisée pour vérifier le modèle obtenu en apprentissage sur un autre échantillon de variables tirées aléatoirement de la même base de données que l'échantillon d'apprentissage et dont on connaît la prédiction. Ici, on sélectionne le meilleur des modèles élaborés dans l'étape 1 en évitant le biais que provoquerait le test sur le même échantillon que l'apprentissage ;
3. Validation : Étape où sur un troisième échantillon dont on connaît la prédiction, on mesure les performances du meilleur modèle sélectionné dans les étapes 1 et 2 ;
4. Application : Étape consistant à appliquer le modèle obtenu à l'ensemble de la base de données et/ou à de nouveaux jeux de données.

On souhaite qu'une technique de prédiction réunisse les qualités décrites ci-après (Larose et Larose, 2014) :

- a. Précise avec un faible taux d'erreur ;
- b. Robuste i.e. aisément généralisable à d'autres échantillons et le moins sensible que possible à des fluctuations aléatoires des variables et des données manquantes ;
- c. Concis avec le moins de règles simples que possibles ;
- d. Rapide.

Si toutes ces qualités ne peuvent pas être obtenues en même temps, on cherche à trouver un compromis entre celles qui sont les plus importantes dans le cadre de l'étude.

5.3 Arbres de décision

La technique d'arbre de décision (*decision tree*, DT) est l'une des méthodes d'extraction de données les plus intuitives car elle s'adapte bien avec les données hétérogènes, les données manquantes et les effets non-linéaires. De plus, grâce à sa robustesse, cette technique fournit des règles explicites, facilement interprétables et compréhensibles pour la classification. Les DT se rapprochent aussi bien des méthodes prédictives que des méthodes descriptives car ils créent une classification en segmentant la base de données à laquelle ils sont appliqués. Ils ne fournissent pas toujours la meilleure performance ; ils constituent néanmoins un bon compromis entre la simplicité d'explication et la performance.

5.3.1 Principe des arbres de décision

L'arbre de décision, comme technique de classification supervisée, permet de classer un ensemble de données décrites par des variables qualitatives et quantitatives en des classes les plus homogènes que possibles. Il est généralement présenté à l'envers comme le montre la Figure 5.2, avec la racine en haut et les feuilles au fond. En partant de la racine, l'arbre se sépare du tronc unique en deux ou plusieurs branches, aussi appelées nœuds. Le nœud est obtenu en choisissant la variable qui, par ses catégories, fournit la meilleure séparation des individus dans chaque classe contenant chacun la plus grande proportion possible des individus dans une seule classe. Cette opération est répétée sur chaque sous-ensemble obtenu de manière récursive jusqu'à ce qu'aucune autre séparation d'individus ne soit possible ou souhaitable dépendamment du type d'arbre. Le nœud terminal ainsi obtenu, est appelé une feuille. À nos jours, il existe essentiellement trois familles d'algorithmes de construction de DT : les arbres Quinlan (Kohavi et Quinlan, 2002a) – (Quinlan, 1986). les arbres CART (*Classification and Regression Tree*) (Breiman et al., 1999) et les arbres CHAID (*CHi-squared Automatic Interaction Detector*) (Kass, 1980).

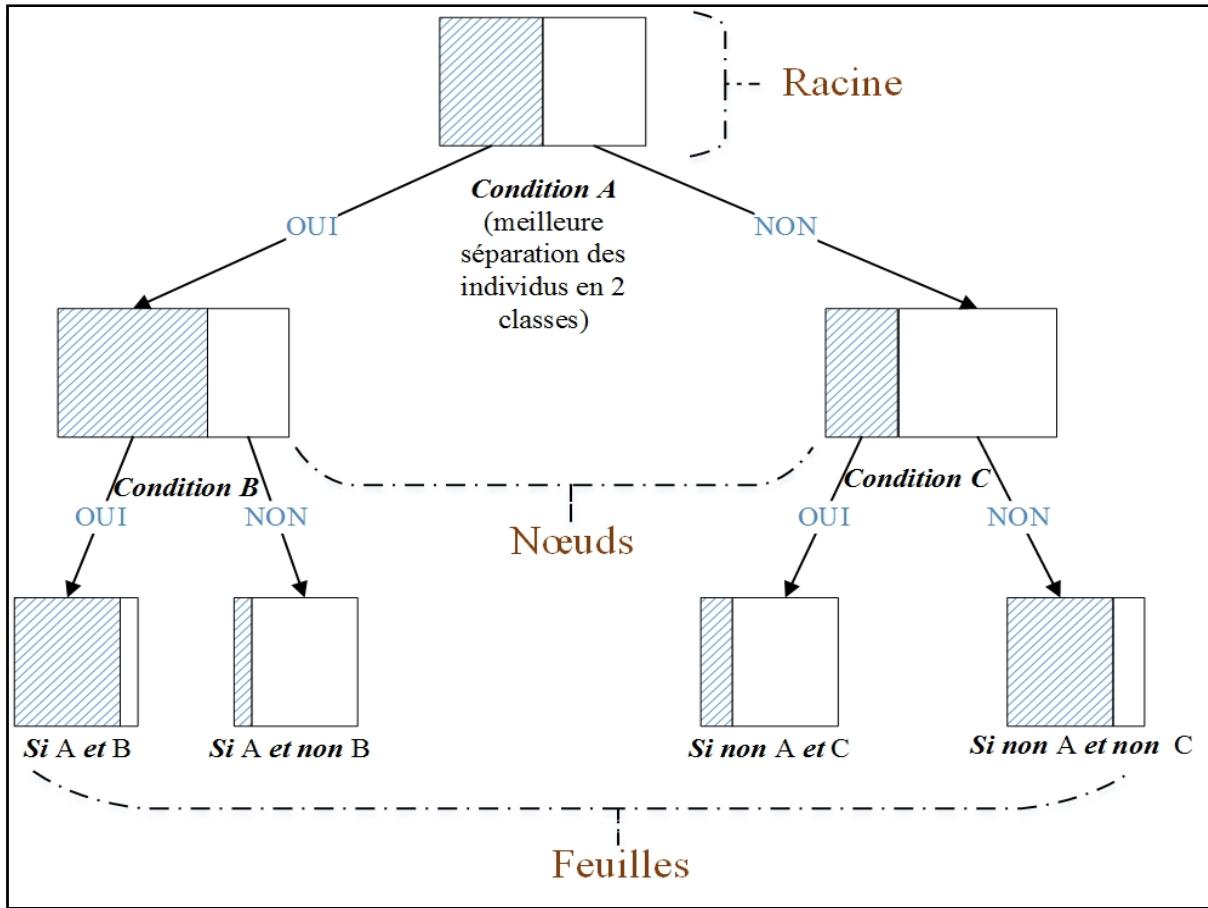


Figure 5.2 Arbre de décision

5.3.2 Critère de segmentation

L'ensemble d'apprentissage est de la forme d'un ensemble d'instances T , comme relaté dans l'équation (5.1), où Y désigne la variable cible et X le vecteur des m prédicteurs :

$$T = \{(X, Y)\} \quad (5.1)$$

avec $X = (x_1, x_2, \dots, x_m)$

Comme mentionné précédemment, les DT sont construits en divisant l'arbre de la racine vers les feuilles. À chaque étape, il faut choisir une variable d'entrée qui réalise le partage optimal de l'ensemble d'objets. Pour avoir la variable de séparation sur le nœud, les algorithmes évaluent les différents prédicteurs possibles et sélectionnent ceux qui maximisent un critère

donné. Les critères les plus utilisés sont le critère du χ^2 (Nisbet, Miner et Elder IV, 2009) utilisé dans l'arbre CHAID, le critère de Twoing et celui de Twoing ordonné (Breiman et al., 1999) tous deux utilisés dans CART, l'entropie (Quinlan, 1986) elle-même employée dans les DT Quinlan et le coefficient de Gini (CART) qui est adopté dans cette thèse.

Le coefficient de Gini, noté $Indice_{Gini}$, mesure avec quelle fréquence un élément aléatoire de l'ensemble serait mal classé si son étiquette était sélectionnée aléatoirement depuis la distribution des étiquettes dans le sous-ensemble. Il peut être calculé à partir de l'équation (5.2) en sommant la probabilité pour chaque élément d'être choisi, multipliée par la probabilité qu'il soit mal classé.

$$Indice_{Gini} = \sum_{i=1}^m f_i(1 - f_i) = 1 - \sum_{i=1}^m f_i^2 \quad (5.2)$$

Où les $f_i, i=1 \text{ à } m$ sont les fréquences relatives dans les nœuds des m prédicteurs. Plus les classes sont uniformément distribuées dans un nœud, plus $Indice_{Gini}$ est élevé ; plus le nœud est pur (homogénéité de l'ensemble des prédicteurs), plus son indice de $Indice_{Gini}$ est bas. En d'autres termes, l'indice de Gini mesure la probabilité que deux individus choisis aléatoirement et avec remise dans un nœud, appartiennent à deux classes différentes.

5.3.3 Taille des arbres, surapprentissage, élagage

Selon (Breiman et al., 1999), les performances d'un DT reposent principalement sur la détermination de sa taille. Cette dernière a tendance à croître avec le nombre d'observations dans la base d'apprentissage (Oates et Jensen, 1997). En pratique, dans la construction d'un arbre, on ne cherche pas à avoir un arbre dont les feuilles correspondent à des sous-ensembles parfaitement homogènes par rapport à la variable cible. En effet, le défi est de saisir l'information utile sur la structure statistique de la population, en excluant les caractéristiques spécifiques à l'échantillon étudié. Les arbres ont tendance à produire un classifieur trop complexe, corrélé exagérément aux données ; il s'agit là du phénomène de surapprentissage (*overfitting*). Pour éviter tout surapprentissage et avoir la taille optimale de l'arbre, on procède

à un élagage. Lors de ce processus, on perd en précision mais on gagne en prédition (Breslow et Aha, 1997). Il existe deux techniques d'élagage : le pré-élagage et le post-élagage. Le premier revient à arrêter quand il y a une classe majoritaire dans le nœud ou à arrêter de diviser un nœud quand la pureté des points non dominants est non parfaite mais suffisante. Pour cela, on utilise un seuil pour détecter une classe dominante. L'inconvénient avec le pré-élagage est que l'arrêt de la construction de l'arbre peut donner un arbre sous-optimal. Le second, apparu avec la méthode CART (Breiman et al., 1999), consiste à finir la construction de l'arbre le plus profond que possible. On le simplifie en remontant des feuilles vers la racine en utilisant des critères de qualité qui mesurent un compromis entre l'erreur obtenue et la complexité de l'arbre (Wallace et Patrick, 1993). On fait appel à un ensemble de validation pour mesurer l'erreur à chaque nœud (Quinlan, 2014) – (Kohavi et Quinlan, 2002b).

Un arbre pris seul a de nombreuses limitations. La procédure d'apprentissage étant statique, les arbres ne sont pas prévus pour apprendre de manière incrémentale de nouvelles instances qui s'ajouteraient à l'ensemble d'entraînement (Caron, 2011). De plus, ils sont sensibles au bruit et ont une forte tendance à sur-apprendre les données (i.e. à apprendre à la fois les relations entre les données et le bruit présent dans l'ensemble d'apprentissage). Ajouter à cela, le problème d'optimisation global est NP-complet (Hyafil et Rivest, 1976) pour de nombreux critères d'optimalité, ce qui conduit à l'emploi de nombreux heuristiques. Pour pallier à ces faiblesses, une des solutions revient à multiplier aléatoirement le nombre de DT comme débattu dans la section suivante.

5.4 Forage de données : Forêt aléatoire d'arbres de décision, *Random Forest (RF)*

Une forêt aléatoire d'arbre de décision est un ensemble d'arbres de décision binaire randomisés. Le côté aléatoire rend les forêts plus générales et plus efficaces qu'un seul arbre de décision. Cependant, l'interprétation du RF est plus difficile que celle d'un DT pris seul. Il existe plusieurs modèles de forêts aléatoires dépendamment de la manière dont la randomisation a été incorporée dans la construction des arbres. À titre d'exemple, on peut citer :

- a. Le *Random Subspace* (Tin Kam, 1998) qui procède à une sélection à chaque nœud de M variables de manière aléatoire et, parmi celles-ci, choisit celle qui minimise un certain critère ;
- b. Le *Random Select Split* (Dietterich, 2000) qui sélectionne les M meilleures séparations (position de coupure réalisée aléatoire) et qui en sélectionne une parmi celles-ci aléatoirement ;
- c. Le *Tree Bagging* (Breiman, 1996a) qui introduit de la randomisation dans l'échantillon initial en sélectionnant certains points plutôt que d'autres et en laissant l'arbre se construire le plus profond que possible ;
- d. Le *Random Forest* qui est une combinaison du CART, du *bagging* et du *subspace*. On a adopté cette technique dans la présente thèse et il fait l'objet d'une description complète dans la sous-section suivante.

5.4.1 Principe général du *Random Forest*

Adele Cutler et Leo Breiman initient formellement en 2001 la notion du *Random Forest* (Breiman et Cutler, 2001). Il s'agit au fait d'une adaptation du *bagging* (Breiman, 1996a) qui est un méta-algorithme combinant le *bootstrap* et l'agrégation, d'où le nom (*Bootstrap AGGregatING*). On appelle *bootstrap* (Efron et Tibshirani, 1994) – (Efron et Efron, 1982) d'un échantillon S comportant n éléments, l'échantillon obtenu par un tirage réalisé uniformément au hasard et avec remise de n éléments parmi n . Dans un échantillon *bootstrap*, chaque élément peut ne pas être tiré ou tiré plusieurs fois. La probabilité Γ pour qu'un élément soit tiré est donnée par l'équation 5.3 :

$$\Gamma = 1 - \left(1 - \frac{1}{n}\right)^n \quad (5.3)$$

où $\Gamma \rightarrow 1 - e^{-1} \approx 0.632$ quand $n \rightarrow \infty$

L'agrégation pour sa part, produit plusieurs *bootstraps* que chacun d'entre eux utilise pour entraîner un prédicteur. En définitive, le *bagging* permet de diminuer la variance mais non le

biais d'un modèle, et contribue à la correction du manque de robustesse des classifieurs instables des DT.

Un RF est une collection de B arbres $\{T_1(X), T_2(X), \dots, T_B(X)\}$ où $X = \{x_1, x_2, \dots, x_p\}$ est un vecteur de p colonnes de prédicteurs. L'ensemble produit B sorties $\{\hat{Y}_1 = T_1(X), \hat{Y}_2 = T_2(X), \dots, \hat{Y}_B = T_B(X)\}$, \hat{Y}_B ($b = 1, \dots, B$) étant la prédiction du B – ème arbre. Les sorties de tous les arbres sont agrégées pour produire une prédiction finale \hat{Y} , laquelle prédiction est la classe prédite par la majorité des arbres. Le *bootstrap* permet d'introduire le côté aléatoire avec remise dans le processus. Pour la phase d'apprentissage, il revient à considérer un jeu de données $D = \{(X_1, C_1), (X_2, C_2), \dots, (X_n, C_n)\}$ avec $(X_i, i = 1, \dots, n)$ un vecteur de prédicteurs et C_i la variable cible. L'algorithme suivant affiche les différentes étapes franchies lors de l'apprentissage du RF (Liaw et Wiener, 2002) :

Algorithme 5.2 Vue algorithmique de la construction d'un RF

1. Construire un échantillon *bootstrap* D_b à partir de D ;
2. Construire pour chaque échantillon D_b , un arbre T_b conformément à l'algorithme du CART (Breiman et al., 1999) en apportant les modifications suivantes :
 - a. Sélectionner aléatoirement n_{try} variables parmi les p prédicteurs avec ($n_{try} < p$);
 - b. Choisir à chaque nœud la variable permettant une meilleure séparation parmi les n_{try} descripteurs. Noter si $n_{try} = p$, il s'agit d'un *bagging* (Breiman, 1996a). Généralement, on choisit la valeur empirique $n_{try} = \sqrt{p}$;
 - c. Diviser le nœud en deux nœuds filles ;
 - d. Construire l'arbre afin d'atteindre sa taille maximale sans élagage ;
3. Répéter les procédures précédentes un grand nombre de fois pour obtenir un nombre suffisant d'arbres de décision ;
4. Prédire les nouvelles données en agrégant les prédictions des B arbres.

Quoique le RF soit un outil relativement récent dans le forage de données, plusieurs travaux académiques (Strobl, Malley et Tutz, 2009) – (Friedman, Hastie et Tibshirani, 2001)

reconnaissent et vantent ses forces. Il s'avère très robuste pour faire face au bruit et très efficace, en particulier quand le nombre de prédicteurs est très élevé. Contrairement au DT qui teste la séparation des nœuds sur tous les prédicteurs et procède à un élagage systématique, le RF tire avantage du fait que chaque arbre construit atteint sa profondeur maximale et que seulement un nombre restreint de ses prédicteurs sélectionnés aléatoirement sont choisis lors de la séparation des nœuds (Maimon et Rokach, 2005) comme rapporté dans l'Algorithme 5.2.

5.4.2 Prédiction d'un RF

La prédiction d'un RF est la combinaison des prédictions des différents arbres prises une-à-une. Pour un problème de classification, la classe ayant obtenu le plus grand nombre de votes suite à l'agrégation est la prédiction du RF, comme mentionné dans l'Algorithme 5.2. Il s'ensuit que $\hat{C}_{RF}^B(X)$, la prédiction du RF en fonction de la classe prédite du k – ème arbre $\hat{C}_k(X)$, est donnée par l'équation (5.4) (Breiman et Cutler, 2001) :

$$\hat{C}_{RF}^B(X) = \text{majorité de votes } \{\hat{C}_k(X), k = 1, \dots, B\} \quad (5.4)$$

Avec l'équation (5.5), on trouve les probabilités de prédiction $\hat{\Gamma}_{RF}^B$ en moyennant les fréquences relatives des classes de tous les arbres individuels ($\hat{\Gamma}_{T_k}$ probabilité associée à l'arbre T_k dans la forêt aléatoire) :

$$\hat{\Gamma}_{RF}^B(C_{RF}^B(X)) = \frac{1}{B} \sum_1^B \hat{\Gamma}_{T_k} \hat{C}_k(X) \quad (5.5)$$

Avec des arbres non élagués, un des avantages des RF, on aboutit à des estimations de probabilités plus précises (Liang, Zhang et Yan, 2006). En effet, en associant les estimations de probabilités d'arbres individuels à des variables aléatoires ayant chacun comme variance σ^2 , la variance de la moyenne de l'équation (5.5) est $\frac{\sigma^2}{B}$ (Friedman, Hastie et Tibshirani, 2001). En fin de compte, les prédictions des RF ne sont plus le fruit d'un raisonnement, mais avec sa

double randomisation constitue un consensus de raisonnements potentiellement très différents (Tuffery et Riesco, 2011).

5.4.3 Évaluation de la performance des RF

Estimation de l'erreur Out-of-Bag (OOB)

Il est bon de rappeler que lors de la phase d'apprentissage, on utilise dans le RF les deux tiers de données de la base initiale pour construire les arbres du modèle (James et al., 2013). Ces données sont vues comme étant dans un sac (*bag*). Le tiers de données restant, dites hors du sac (*out-of-bag*), sert en toute logique à tester la performance du modèle ainsi trouvé. La Figure 5.3 reproduit ces phases.

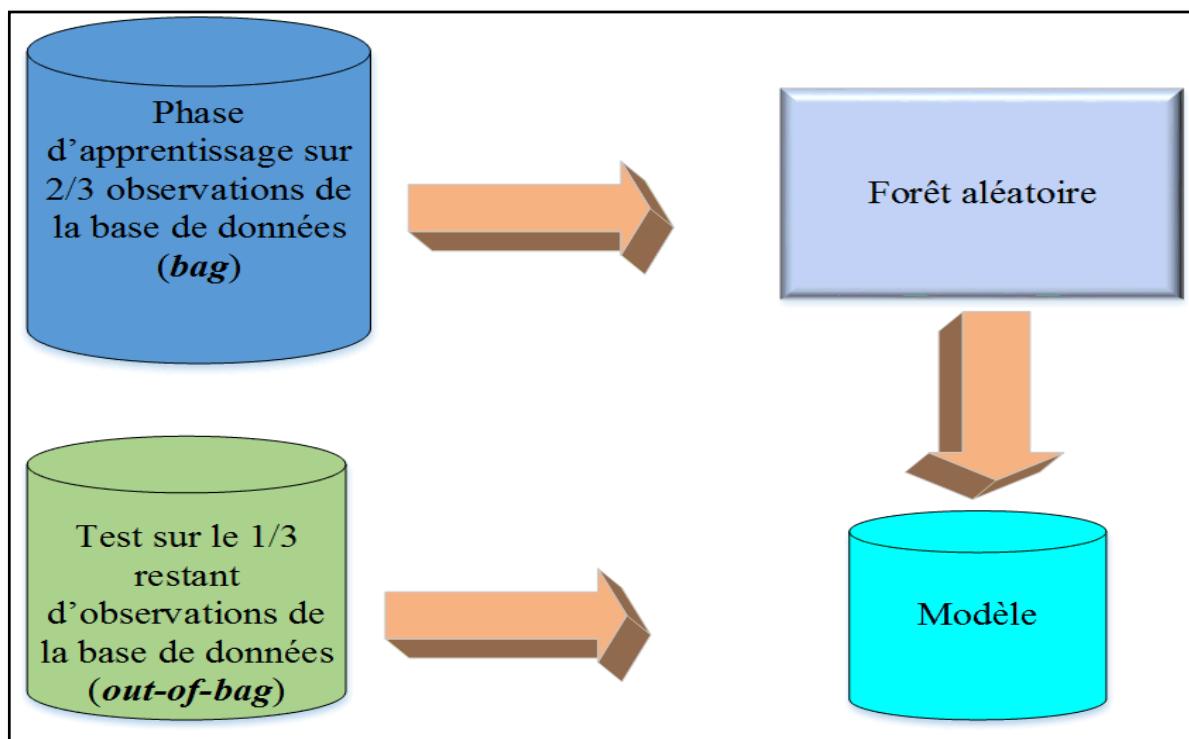


Figure 5.3 Interprétation de l'*out-of-bag* dans le fonctionnement de la forêt aléatoire

L'estimation de l'erreur OOB représente une manière plus simple d'estimer l'erreur commise sur les données n'ayant pas participé à l'apprentissage, ipso facto à la construction des arbres du modèle. Cette estimation « impartiale » d'erreur notée OOB_{error} et définie dans l'équation (5.6), suggère à quel point les réponses sont en erreur lorsque le modèle résultant est appliqué à de nouvelles observations (Breiman et Cutler, 2001).

$$OOB_{error} = \frac{1}{n} \sum_{i=1}^n I(\hat{Y}_{OOB}(X_i) \neq Y_i) \quad (5.6)$$

Où $\hat{Y}_{OOB}(X_i)$ est l'ensemble prédit par agrégation et $I(***)$ une fonction caractéristique. L'approche OOB pour estimer l'erreur de test est particulièrement pratique lors de la manipulation de grands ensembles de données pour lesquels la validation croisée serait techniquement laborieuse. Notez que grâce à l' OOB_{error} , on peut facilement déduire combien d'arbres sont suffisants pour prédire correctement.

Importance relative des variables

L'importance d'une variable est un concept généralement difficile à définir car elle peut exprimer des interactions entre les différentes variables. L'arbre de décision est connu pour sa capacité à sélectionner certains prédicteurs, dits « importants », et à ignorer d'autres, souvent inutiles. De plus, le DT fournit un modèle explicite décrivant la relation entre ses prédicteurs et ses variables cibles. Naturellement, étant un ensemble d'arbres, le RF hérite de la possibilité de sélectionner des prédicteurs « importants ». Cependant, comme il ne produit pas un modèle explicite, la relation entre les différents prédicteurs est cachée (Svetnik et al., 2003). Il existe tout de même deux algorithmes (Liaw et Wiener, 2002) permettant de savoir comment chaque prédicteur contribue à la précision du modèle découlant du RF. Il s'agit de la diminution moyenne du coefficient de Gini (*mean decrease in Gini coefficient, MDG*) et de la diminution moyenne de la précision (*mean decrease in accuracy, MDA*). Ces deux algorithmes sont implantés dans le logiciel **R** décrit dans la section 5.6 de ce présent chapitre. Le premier est une mesure locale qui évalue la contribution de chaque variable à garder une certaine homogénéité au niveau des nœuds et des feuilles. Plus précisément, à chaque nœud, on

sauvegarde la diminution de l'indice d'impureté Gini pour la variable X_i ayant participé dans la division du nœud en deux nœuds filles. On moyenne toutes les diminutions de l'impureté de Gini ainsi obtenues pour avoir le MDG.

Comme rappelé à la Figure 5.3, les observations OOB forment un test naturel fixé pour chaque arbre. Le second algorithme assimile l'importance de la variable à une diminution moyenne de la précision en se référant aux observations OOB. Le MDA mesure alors l'impact de chaque variable sur la précision du modèle. L'idée est de permuter les valeurs de chaque variable et de mesurer à quel point la permutation diminue la précision du modèle. Le calcul du MDA se fait conformément à l'Algorithme 5.3.

Algorithme 5.3 Étapes du calcul de l'importance des variables selon le MDA

Pour les échantillons $bootstrap b = 1, \dots, B$

1. Identifier les observations OOB notées $D_{b_{OOB}} = D \setminus D_b$;
2. Prédire l'appartenance de classe pour $D_{b_{OOB}}$ en utilisant l'arbre T_b et additionner le nombre de fois que la prédiction de l'arbre est correcte ;
3. Pour les variables indépendantes $j = 1, \dots, p$
 - a. Permuter les valeurs des variables indépendantes X_j dans $D_{b_{OOB}}$;
 - b. Se servir de l'arbre T_b pour prédire la classe pour $D_{b_{OOB}}$ en permutant les valeurs des X_j , puis additionner le nombre de fois que la prédiction est correcte pour $D_{b_{OOB}}$;
 - c. Soustraire le nombre de votes pour les classes correctement prédites pour les données permutes dans l'OOB du nombre de votes de données bien classées dans l'OOB.

En cherchant l'importance de chacune des variables, on calcule l'augmentation de l'erreur de prédiction si les valeurs de la variable en question sont permutes pour les observations OOB.

¹ Il s'agit ici de la notation « sauf » employé usuellement dans le langage des ensembles en mathématiques.

Cette mesure est faite pour chaque arbre puis moyennée sur la base de données au complet. Ce dernier résultat est divisé par l'écart-type de l'ensemble des données. Cette démarche est illustrée dans la Figure 5.4. Dans les deux cas, pour un nombre fixe d'arbres, des valeurs plus élevées indiquent que la variable est relativement plus importante.

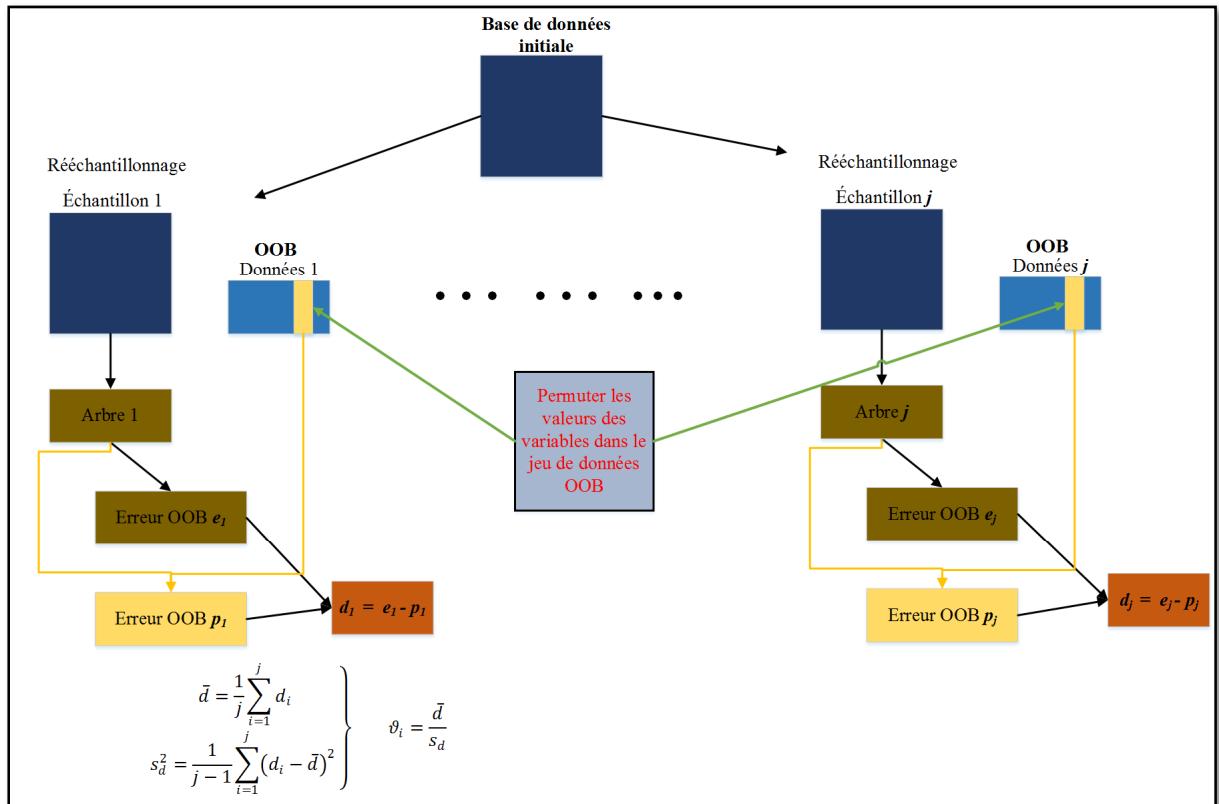


Figure 5.4 Importance des variables

Matrice de confusion et métriques statistiques d'évaluation de performance

Pour évaluer la performance et mesurer la qualité du modèle obtenu de la forêt aléatoire, on se sert de la matrice de confusion. Cette dernière consiste à confronter les valeurs observées de la variable dépendante Y aux valeurs prédites \hat{Y} fournies par le modèle. Une matrice de confusion pour un problème à deux classes ($Y = \{stable, instable\}$) prend la forme du Tableau 5.1. Chaque colonne correspond au nombre d'occurrences d'une classe estimée tandis que chaque

ligne renvoie au nombre d'occurrences d'une classe réelle où les données utilisées pour chacun de ces groupes sont différentes.

Tableau 5.1 Matrice de confusion

Classe réelle	Classe prédictive	
	Instable	Stable
Instable	α_u	β_{ds}
Stable	β_{du}	α_s

Sur la diagonale principale de la matrice, on retrouve les observations correctement classées où α_u est le nombre d'observations classées « instable », α_s représente le nombre d'observations « stable » et le nombre d'observations correctement classées est $\alpha = \alpha_u + \alpha_s$. Les observations « instable » classées « stable » sont notées β_{ds} et les observations « stable » classées « instable » sont notées β_{du} . Le nombre total d'observations mal classées valent $\beta = \beta_{du} + \beta_{ds}$. Conséquemment, le nombre total d'observations est $\alpha + \beta$. De cette matrice, on en déduit une série de métriques statistiques pour évaluer la performance du modèle. La première métrique est l'exactitude (*Accuracy*) qui est l'estimation de la justesse d'une observation. En d'autres termes, elle exprime dans quelle mesure le résultat est proche de la « vraie » valeur. Elle est donnée par la formule de l'équation (5.7) (Powers, 2011).

$$Accuracy = \frac{\alpha_s + \alpha_u}{\alpha + \beta} \quad (5.7)$$

La deuxième métrique prise en compte est la fiabilité (*Reliability*). Il est question de la cohérence globale de la mesure i.e. en effectuant la même expérience dans les mêmes conditions, on doit obtenir les mêmes résultats : reproductibilité de l'expérience. Cette mesure se trouve à partir de l'équation (5.8).

$$Reliability = \frac{\alpha_u - \beta_{ds}}{\alpha_u} \quad (5.8)$$

La troisième métrique est la sécurité (*Security*). Elle est une mesure de l'optimalité du modèle et se calcule comme mentionné dans l'équation (5.9) (Suprême et al., 2016).

$$Security = \frac{\alpha_s - \beta_{du}}{\alpha_s} \quad (5.9)$$

Les deux dernières métriques sont la mauvaise détection (*MisDetection*) et la fausse alarme (*False Alarm*). La mauvaise détection (équation 5.10) est la proportion de cas stables qui ont été mal classés alors que la fausse alarme (équation 5.11) représente le pourcentage de cas instables qui ont été mal classés (Suprême et al., 2016) – (Kamwa, Samantaray et Joos, 2009).

$$MisDetection = \frac{\beta_{ds}}{\alpha + \beta} \quad (5.10)$$

$$FalseAlarm = \frac{\beta_{du}}{\alpha + \beta} \quad (5.11)$$

À partir des définitions des métriques précédentes et des équations (5.7) à (5.11), on peut aisément déduire que le modèle recherché doit avoir une exactitude, une fiabilité et une sécurité très élevées avoisinant les 100% mais un taux de fausse alarme et de mauvaise détection très minime se rapprochant le plus possible du 0%.

5.5 Forage de données : *Boosting*

Le *boosting* est un méta-algorithme qui cherche à réduire principalement le biais et la variance du classifieur de base dans un processus d'apprentissage (Breiman, 1996b). L'idée générale d'un tel algorithme est d'utiliser plusieurs fois un apprenant faible (*weak learner*) sur des distributions différentes, puis de combiner les différentes hypothèses obtenues au sein d'un vote. Il tient son origine dans la question suivante (Kearns, 1988) – (Kearns et Valiant, 1994) :

« *Can a set of weak learners create a single strong learner? A weak learner is defined to be a classifier which is only slightly correlated with the true classification.*

In contrast, a strong learner is a classifier that is arbitrarily well-correlated with the true classification. »

Pour une base de données quelconque, le *boosting* construit plusieurs modèles en utilisant un autre algorithme d'apprentissage, l'arbre de décision est généralement utilisé. Son principe est d'attribuer un poids à chaque observation, puis de solliciter l'apprenant faible sur ces observations et ces poids pour obtenir une hypothèse qui classe correctement les observations de poids forts. Ensuite, les poids des observations sont modifiés en fonction de cette hypothèse i.e. les poids des observations correctement classées sont diminués et ceux des mal classées augmentés. Le processus est itéré jusqu'à l'obtention d'un modèle final dit adaptatif car le modèle précédent est toujours pris en compte.

L'une des faiblesses principales du *boosting* est sa sensibilité au bruit (Opitz et Maclin, 1999). Deux des algorithmes les plus utilisés en *boosting* sont *Adaboost* (*Adaptive Boosting*) (Freund et Schapire, 1996) et *Arcing* (*Adaptive Resampling and Combining*) (Breiman, 1998). Le *boosting* a certes des caractéristiques similaires au *bagging*. Cependant, dans sa construction on n'utilise pas des échantillons *bootstrap* et en étant un algorithme séquentiel le *boosting* ne peut pas être parallélisé. Le *boosting* ressemble aussi au RF puisqu'un ensemble de modèles basés sur des DT est formé pour déduire le meilleur des modèles. Ce qui le distingue principalement du RF est le fait que les arbres sont construits les uns les après les autres. Au fur et à mesure, le modèle est raffiné en tenant compte des arbres précédents. Les observations mal classées sont boostées pour qu'elles aient davantage d'importance vis-à-vis de l'apprenant faible dans le prochain tour afin qu'il pallie le manque.

On considère le même jeu de données que précédemment en modifiant la sortie tel que suit $Y = \{1, -1\}$. La sortie 1 correspond à l'état « Stable » alors qu'on assimile à la sortie -1 l'état « Instable ». Le pseudocode de l'Algorithme 5.4 donne les différentes étapes lors de l'implémentation de l'*Adaboost* (Schapire, 1999). L'idée principale de l'algorithme est de maintenir une distribution de poids sur les données d'apprentissage. Cette dernière est notée $D_t(i)$ où i est l'observation en question et t l'itération. Initialement, tous les poids sont supposés égaux, mais à chaque itération, les poids des observations mal classées sont

augmentés dans le but de forcer l'apprenant faible à se concentrer sur les observations difficiles à classer dans le jeu de données. L'apprenant faible doit trouver une hypothèse faible convenable à la distribution D_t . Cette hypothèse h_t est de la forme de l'équation 5.12 :

$$h_t : X \rightarrow \{-1, 1\} \quad (5.12)$$

La qualité de l'hypothèse faible est mesurée par son erreur ϵ_t dans l'équation (5.13) par rapport à la distribution D_t sur laquelle l'apprenant faible a effectué son apprentissage (Schapire, 1990). Dans cette notation, $Pr(\cdot)$ la probabilité que h_t fasse un mauvais classement de l'observation x_i choisi aléatoirement suivant la distribution D_t .

$$\epsilon_t = \Pr_{i \sim D_t} [h_t(x_i) \neq y_i] = \sum_{i: h_t(x_i) \neq y_i} D_t(i) \quad (5.13)$$

Le choix du paramètre α_t de l'algorithme peut être compris comme une importance assignée à h_t . Remarquez que pour $\epsilon_t \leq \frac{1}{2}$, on a $\alpha_t \geq 0$. Cet ajustement dans la formule permet de conserver une certaine généralité ; plus α_t est grand, plus ϵ_t est petit. La mise à jour de la distribution D_t conduit tel que souhaité, à une augmentation des poids des observations mal classées par h_t et à une diminution des poids de celles correctement classées. Finalement, l'hypothèse de sortie H est un vote à la majorité pondérée des T hypothèses faibles où α_t est le poids attribué à h_t .

Algorithme 5.4 Pseudocode du boosting selon Adaboost

Soit T le nombre d'itérations à effectuer. Initialiser $D_1(i) = \frac{1}{p}$.

Pour $t = 1, \dots, T$:

1. Entraîner l'apprenant faible en utilisant la distribution D_t ;
2. Obtenir une hypothèse faible h_t telle que décrite dans l'équation (5.11) avec une erreur ϵ_t comme indiqué dans l'équation (5.12) ;
3. Choisir α_t tel que :

Algorithme 5.4 Pseudocode du boosting selon Adaboost

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \quad (5.14)$$

4. Mettre à jour D_t tel que :

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & \text{si } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{si } h_t(x_i) \neq y_i \end{cases} \quad (5.15)$$

Où Z_t est un facteur de normalisation

5. Hypothèse finale :

$$H(x) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x)) \quad (5.16)$$

Il est à noter qu'à la sortie, on dispose du poids de chaque arbre dans le modèle ainsi que de l'erreur de classement associée à chacun de ces poids dans le classement. Ces données vont être exploitées plus loin dans cette thèse pour une prédition en temps réel de la stabilité du réseau électrique. Les métriques statistiques définies dans la section précédente demeurent adéquates pour évaluer la performance du modèle.

Somme toute, le *boosting* est une technique additive d'agrégation qui permet dans certains cas (base de données trop dense) d'améliorer de façon significative la qualité des prédictions, notamment en termes de robustesse. Néanmoins, on perd l'interprétation facile des arbres de décision. Pour des bases de données trop denses, le temps de calcul augmente rapidement et le stockage de toutes les données devient problématique.

5.6 Description du logiciel R

Les différents réseaux tests auxquels on a eu recours dans cette thèse sont implémentés dans *Simscape Power Systems™* de Matlab®. De nombreuses simulations ont mené à la construction d'une base de données. Pour exploiter toutes les informations contenues dans cette dernière et

pour simuler les techniques proposées dans cette thèse, on a utilisé le logiciel **R** (Ihaka et Gentleman, 1996). Ce logiciel de statistique est à la fois un langage informatique et un environnement de travail. **R** est un clone du logiciel *S-plus* fondé sur le langage de programmation orienté objet *S* et développé par *AT&T Bell Laboratories* (Becker, Chambers et Wilks, 1988). Il offre la possibilité de manipuler des données, de faire de laborieux calculs, de tracer et d'afficher des graphiques et de faire des analyses statistiques sur des données.

R étant un logiciel gratuit, libre et à code source ouvert (*open-source*), il est facile d'accès à tous et connaît des évolutions et des améliorations en permanence : nouvelles intégrations, nouvelles fonctionnalités, nouvelles méthodes d'analyse, etc... C'est un logiciel multiplateforme fonctionnant sous UNIX et Linux, Microsoft Windows et Mac OS. Au prime à bord, **R** est plus difficile à apprivoiser que certains logiciels du marché tels que *Statistical Package for the Social Sciences* (SPSS) et Minitab (Muenchen, 2011). Il faut donc passer du temps à apprendre la syntaxe et les commandes. Pour bien manipuler **R**, il faut maîtriser préalablement les concepts et les méthodes statistiques. Selon ses besoins, un utilisateur peut développer ses propres outils sur **R** pour faire des analyses très sophistiquées sur les données (de Micheaux, Drouilhet et Liquet, 2011).

R est aussi un langage de programmation simple et efficace dans l'intégration des conditions, des boucles et de la récursivité. Le traitement des tableaux et des matrices se fait aisément et il possède une bonne capacité de stockage des données (Muenchen et Hilbe, 2010). On peut y effectuer les opérations mathématiques classiques. Avec sa vaste et cohérente collection de procédures statistiques, on peut analyser les données et les représenter à partir de graphiques évolués. Le logiciel a un système de documentation intégré très bien conçu et un site web dédié *Comprehensive R Archive Network* (CRAN) <http://cran.r-project.org/manuals.html> et <http://cran.r-project.org/other-docs.html>. Une autre force de **R** est la possibilité d'appeler directement des fichiers Excel et des fonctions Matlab®.

Plusieurs techniques statistiques classiques et avancées sont implémentées dans **R**, parmi lesquelles on peut citer : la statistique descriptive, les tests d'hypothèse, l'analyse de variance, les méthodes de régression linéaire simple et multiple, etc... Toutes ces méthodes sont

regroupées dans des modules externes appelés *packages* qu'il faut télécharger selon les besoins. Tous les *packages* sont bien documentés quant au bagage théorique sur lesquels ils reposent et la façon de les manipuler <https://cran.r-project.org/>. Pour le forage des données, on dispose de plusieurs algorithmes dont ceux cités dans le Tableau 5.2 avec les packages dans lesquels ils se retrouvent.

Tableau 5.2 Packages dans **R** correspondant à des techniques de forages

Techniques de forage de données	<i>Packages</i> correspondant dans R
Analyse discriminatoire linéaire (<i>Linear Discriminant Analysis</i>)	lda
Régression logistique (<i>LogisticRegression</i>)	glm
Arbres de décision (<i>Decision Trees</i>)	rpart, wsrpart
Forêts aléatoires (<i>Random Forest</i>)	randomForest, wsrf
Boosting	ada, adaboost, adabag
Bagging	adabag
Réseaux de neurones (<i>Neural Networks</i>)	nnet
<i>Support Vector Machines</i>	kernlab

Dans le cadre de cette thèse, on réalise de manière séquentielle les tâches suivantes avec le logiciel **R** :

1. Charger le fichier contenant la base de données ;
2. Définir le rôle des variables (prédicteurs ou variables d'entrée versus variable cible ou sortie) ;
3. Diviser la base de données en trois échantillons :
 - a. Un échantillon d'apprentissage pour construire le modèle ;
 - b. Un échantillon de validation ou d'évaluation pour estimer l'erreur du modèle ;
 - c. Un échantillon de test pour évaluer l'erreur de généralisation du modèle final.
4. Analyser les données avec des approches de statistiques descriptives et des graphes ;
5. Construire les modèles prédictifs ;
6. Évaluer leur performance ;

7. Appliquer les modèles ainsi obtenus sur de nouveaux jeux de données.

L’interface graphique de **R** (*Graphical User Interface*, GUI) étant très limitée et peu conviviale, on préfère utiliser l’interface Rattle pratique pour l’apprentissage statistique (Williams, 2011) et l’interface RCommander (*package Rcmdr*) qui permet d’effectuer des analyses graphiques et statistiques usuelles. À l’Annexe III, on propose une démonstration simple des possibilités d’analyse de données et de graphiques dans **R** sur une base de données relativement facile à appréhender.

5.7 Conclusion

En disposant d’une grande quantité de données, l’exploration de données, communément appelée forage de données (*data mining*), a comme objectif principal d’extraire de la connaissance sous forme de modèles. Ces modèles sont construits à partir d’algorithmes se reposant sur des méthodes statistiques, l’intelligence artificielle et l’informatique. Dans le cadre de cette thèse, on retient la structure des arbres de décision comme outils d’aide à la décision. Comme un arbre pris seul a une forte tendance à sur-apprendre les données et est particulièrement sensible au bruit, on a préféré combiner un ensemble d’arbres dans un processus de *boosting* et de forêt aléatoire d’arbres de décision pour être en mesure de prédire en temps réel l’état du réseau électrique. Pour évaluer la performance du modèle trouvé, on a défini une série de métriques statistiques. On a également présenté dans ce chapitre, le logiciel *open-source R* à partir duquel l’implémentation des modèles a été réalisée.

CHAPITRE 6

FONDEMENTS THÉORIQUES DE LA MÉTHODOLOGIE PROPOSÉE

6.1 Introduction

L'approche du centre d'inertie (*Center of Inertia*, COI) est une approche initiée depuis longtemps dans l'analyse de la stabilité. Dans cet article (Athay, Podmore et Virmani, 1979), les auteurs décrivent les fondements théoriques sur lesquels elle repose. Cette approche se base essentiellement sur les limites et les faiblesses des fonctions d'énergie et des fonctions de Lyapunov débattues dans la revue de littérature. Il est présenté une méthode analytique pouvant potentiellement fournir une perspective plus large du problème de stabilité transitoire et pouvant simuler étape par étape, des cas particuliers de défauts avec des modèles détaillés de réseaux. La principale conclusion de cette étude est :

« ... the study is that the basic approach developed is practical, sufficiently accurate and could be applied to realistic problems in power system planning and operation. »

En se référant aux équations (3.7) et (3.8), on définit l'angle de référence δ_0 de l'équation (6.1) et on déduit l'équation (6.2) :

$$\delta_0 = \frac{1}{H_T} \sum_{i=1}^n H_i \delta_i \quad (6.1)$$

avec $H_T = \sum_{i=1}^n H_i$

$$2H_T \dot{\omega}_0 = \sum_{i=1}^n P_{m_i} - 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n G_{ij} \cos(\delta_{ij}) = P_a \quad (6.2)$$

Ces transformations relatives à l'angle de référence offrent une interprétation physique pour la formulation du problème de stabilité transitoire. Les dynamiques du centre d'inertie sont régies par l'équation (6.2). En définissant des angles θ_i et des vitesses $\tilde{\omega}_i$ par rapport à cette référence, on obtient :

$$2H_{T_i}\tilde{\omega}_i = P_{m_i} - P_{e_i} - \frac{H_i}{H_T}P_a \quad (6.3)$$

avec $\tilde{\omega}_i = \omega_i - \omega_0$; $\theta_i = \delta_i - \delta_0$ et $\dot{\theta}_i = \tilde{\omega}_i$; $i = 1, \dots, n$

On fait remarquer que les variables du centre d'inertie satisfont les contraintes :

$$\sum_{i=1}^n H_i \theta_i = \sum_{i=1}^n H_i \tilde{\omega}_i = 0 \quad (6.4)$$

La fonction d'énergie (*Transient Function Energy*, TEF) post-défaut V est donnée par l'équation (6.5) que l'on obtient en multipliant chacune des équations d'accélération relative par la vitesse relative correspondante et en intégrant la somme des équations obtenues à partir d'une limite inférieure fixe du point d'équilibre stable (indiqué par l'exposant "s") à une limite supérieure.

$$V = \sum_{i=1}^n H_i \tilde{\omega}_i^2 - \sum_{i=1}^n (P_{m_i} - E_i^2 G_{ii}) (\theta_i - \theta_i^s) - \sum_{i=1}^{n-1} \sum_{j=i+1}^n E_i E_j \left[B_{ij} (\cos(\theta_{ij}) \cos(\theta_{ij}^s)) - \int_{\theta_i^s + \theta_j^s}^{\theta_i + \theta_j} G_{ij} \cos \theta_{ij} d(\theta_i + \theta_j) \right] \quad (6.5)$$

On peut identifier tous les termes de l'équation (6.5) comme suit. Le premier terme est la variation totale de l'énergie cinétique du rotor des générateurs par rapport au centre d'inertie dépendant des vitesses des générateurs. Le deuxième terme est une énergie de position, il s'agit de la variation de l'énergie potentielle du rotor par rapport au centre d'inertie. Le troisième terme est l'énergie magnétique emmagasinée dans la branche ij et le quatrième terme est l'énergie dissipée dans la même branche. La somme des trois derniers termes équivaut à l'énergie potentielle totale dépendant uniquement des angles des générateurs. Cette dernière équation montre bien que l'on obtient la fonction d'énergie transitoire en soustrayant la variation d'énergie associée au mouvement du centre d'inertie à l'énergie totale du rotor. Pour étudier la stabilité, on simule le réseau dans le domaine temporel afin de trouver les angles et les vitesses de tous les générateurs, puis on calcule l'énergie totale du réseau. Le but ultime est de trouver la frontière de la région de stabilité V_{cr} .

On retient que le centre d'inertie de l'angle δ_{COI} et de la vitesse ω_{COI} sont donnés par le système d'équations (6.6) :

$$\begin{aligned}\delta_{COI} &\triangleq \frac{\sum_{i=1}^n H_i \delta_i}{\sum_{i=1}^n H_i} \\ \omega_{COI} &\triangleq \frac{\sum_{i=1}^n H_i \omega_i}{\sum_{i=1}^n H_i}\end{aligned}\tag{6.6}$$

6.2 Concept du centre de puissance (*Center of Power, COP*)

Le concept du centre de puissance (COP) dont il est question dans cette thèse est inspiré du COI et de fait s'appuie sur les mêmes fondements théoriques. Dans cette section, on discute de ce qui différencie le COP du COI, de son application et des nouveaux indices proposés pour analyser la stabilité des réseaux électriques (Suprême et al., 2017).

6.2.1 Principes théoriques

Considérons un réseau divisé en n régions avec m générateurs comme montré à la Figure 6.1. On place des PMUs aux jeux de barres connectés aux générateurs et à des emplacements stratégiques du réseau de sorte que l'on a une connexion entre deux régions. Les différents PMUs fournissent les phaseurs de tension et de courant aux jeux de barres ainsi choisis. Un générateur dans une région est dénommé *Generator m* où m est le numéro dudit générateur. La région i renferme m_i générateurs et le nombre total de générateurs dans le réseau m est égal à la somme des m_i . P_{ej}^i , ω_j^i et δ_j^i désignent respectivement la puissance active instantanée délivrée, la vitesse et l'angle du rotor de la machine j dans la région i .

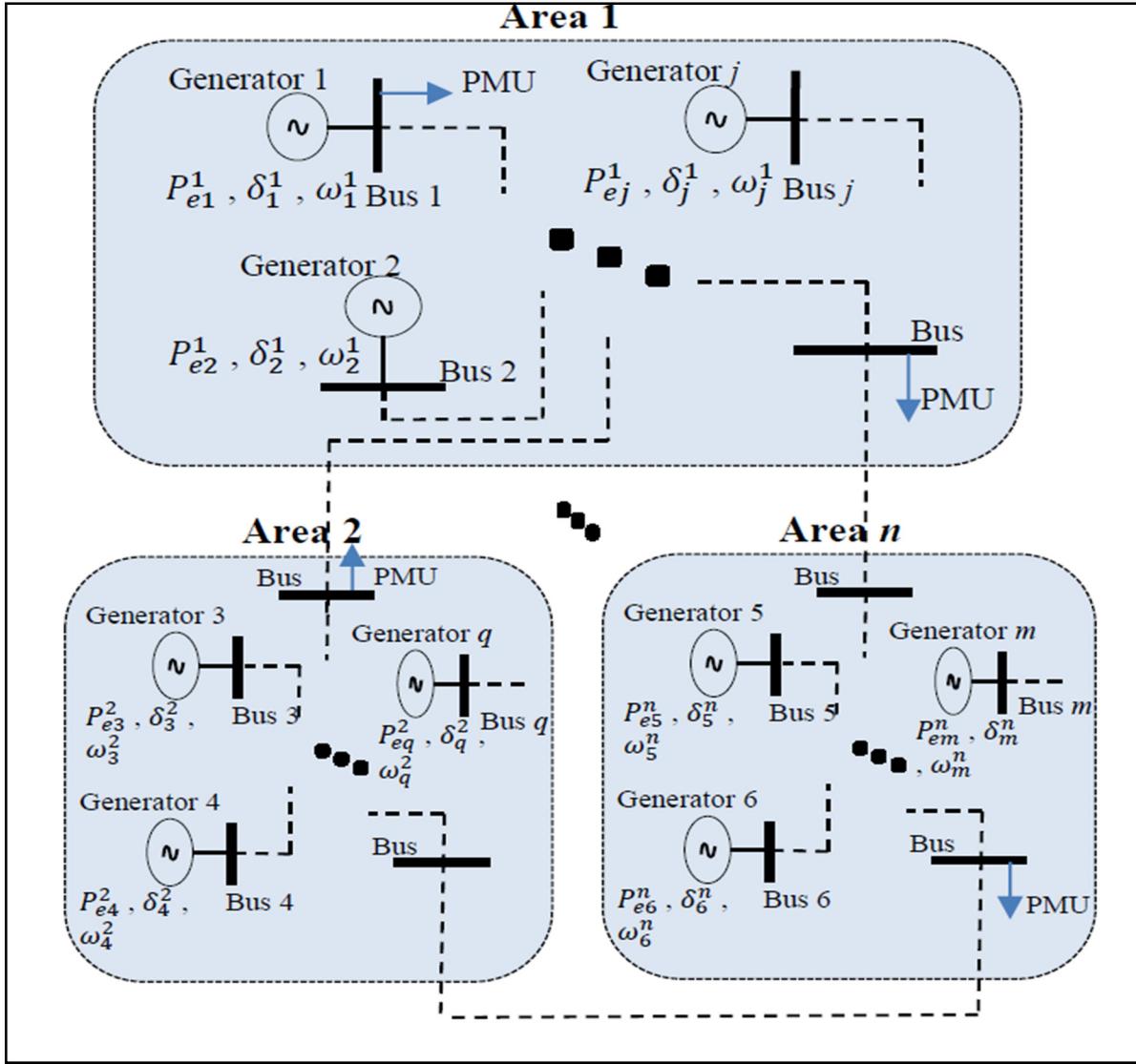


Figure 6.1 Représentation d'un réseau avec n régions et m générateurs

En supposant que chaque région soit cohérente suite à une contingence, son comportement peut être associé à celui d'un seul grand générateur (*single large machine*). Dans ce mémoire (El Aouni, 2014), on explore plusieurs méthodes pour trouver des groupes cohérents de machines. Une méthode faisant appel à des fonctions d'agrégation est aussi proposée. On retient aussi la méthode proposée dans (Chow, 2013) se servant du concept de cohérence et d'analyse sélective modale pour réduire la taille des modèles. Dans cet article (Kamwa et al., 2009b), on utilise la logique floue pour trouver un bon partitionnement du réseau en plusieurs

régions. On peut aussi se servir de cet outil *Power System Toolbox* (PST) implémenté dans Matlab © (Chow et Rogers, 2000) pour avoir les groupes cohérents.

Dans cette thèse, on introduit le concept du centre de puissance (*Center of Power*, COP) qui est une extension du COI et de l'approche de mesure de zone étendue et de système de contrôle de tension. Vu que la constante d'inertie des machines est difficile d'accès en temps réel en raison du nombre de générateurs qui peut varier d'un point d'opération à un autre, il est souhaitable de trouver des méthodes faisant fi de cette constante. En effet, le COP d'un réseau est un point de référence imaginaire situé à la position moyenne pondérée par les puissances des générateurs du réseau électrique. À supposer que les mesures de phaseurs de tension des générateurs soient disponibles dans la région i , il est alors possible de calculer une approximation de la vitesse du rotor des machines en résolvant les équations présentées dans l'équation (3.7). Le COP de vitesse de référence correspondant à la région i est défini comme suit :

$$\omega_{R_i}^{COP} = \frac{\sum_{j=1}^{m_i} \omega_j^i P_{e_j}^i}{\sum_{j=1}^{m_i} P_{e_j}^i} \quad (6.7)$$

Pour le réseau au complet, le COP de vitesse $\omega_{R_T}^{COP}$ peut être calculé avec la vitesse du rotor ω_k des machines en appliquant la formule suivante :

$$\omega_{R_T}^{COP} = \frac{\sum_{k=1}^m \omega_k P_{e_k}}{\sum_{k=1}^m P_{e_k}} \quad (6.8)$$

Ensuite, on propose une approche heuristique pour détecter l'instabilité à partir du COP en temps réel. La valeur absolue maximale de la différence de COP de vitesse entre deux régions quelconques du réseau indique un écart entre elles. Si cet écart a tendance à être constant ou à osciller autour d'une valeur constante, le réseau est stable. En revanche, s'il continue à augmenter dans le temps, le réseau se dirige vers un état instable. En d'autres termes, un cas stable correspond à des COP de vitesse des régions qui évoluent de manière cohérente. Pour

un cas instable, on assiste à des COP de vitesse des régions qui s'éloignent jusqu'à une séparation totale du reste du réseau. Dans un cas d'instabilité, des actions correctives peuvent être appliquées. En général, une action corrective consiste à redistribuer la production afin de réduire la puissance générée dans la zone qui tend vers la séparation du reste du système. Ces deux travaux (Zerigui, 2015) et (Zerigui et al., 2015) traitent de la question de redistribution et d'optimisation d'écoulement de puissance (*Transient Stability Constrained Optimal Power Flow, TSC-OPF*) plus en profondeur.

Pour illustrer le concept de COP, on considère le réseau test de Kundur décrit dans l'Annexe I. Il est divisé en deux régions identifiées : *Area 1* (contenant les générateurs *Generator 1* et *Generator 2* qu'on note ici G1 et G2) et *Area 2* (contenant les générateurs *Generator 3* et *Generator 4* qu'on note ici G3 et G4). On définit alors le COP de vitesse des régions 1 et 2 respectivement par $\omega_{R_1}^{COP}$ et $\omega_{R_2}^{COP}$ tel que suit dans les équations (6.9) et (6.10) :

$$\omega_{R_1}^{COP} = \frac{\omega_1^1 P_{e_1}^1 + \omega_2^1 P_{e_2}^1}{P_{e_1}^1 + P_{e_2}^1} \quad (6.9)$$

$$\omega_{R_2}^{COP} = \frac{\omega_3^2 P_{e_3}^2 + \omega_4^2 P_{e_4}^2}{P_{e_3}^2 + P_{e_4}^2} \quad (6.10)$$

Le COP de vitesse du réseau entier $\omega_{R_T}^{COP}$ est donné par l'équation (6.11) :

$$\omega_{R_T}^{COP} = \frac{\omega_1^1 P_{e_1}^1 + \omega_2^1 P_{e_2}^1 + \omega_3^2 P_{e_3}^2 + \omega_4^2 P_{e_4}^2}{P_{e_1}^1 + P_{e_2}^1 + P_{e_3}^2 + P_{e_4}^2} \quad (6.11)$$

On soumet à $t = 1$ s le réseau à un court-circuit triphasé à la terre d'une durée de 6 cycles. Dans un premier temps, suite à cette contingence, aucune action corrective n'est portée sur le réseau. Sur la Figure 6.2 b), on observe après 6 secondes, il y a une séparation qui se creuse au fur et à mesure entre les COP de vitesse des deux régions. Le réseau va donc perdre son synchronisme si aucune action n'est prise. Or, en appliquant une correction, on remarque sur

la Figure 6.2 a) que les COP de vitesse des régions tendent vers une même valeur constante. Ceci sous-entend que le réseau atteint son équilibre. À partir de cette même figure, on constate que l'écart entre les COP des régions peuvent effectivement servir à détecter l'instabilité après un phénomène transitoire.

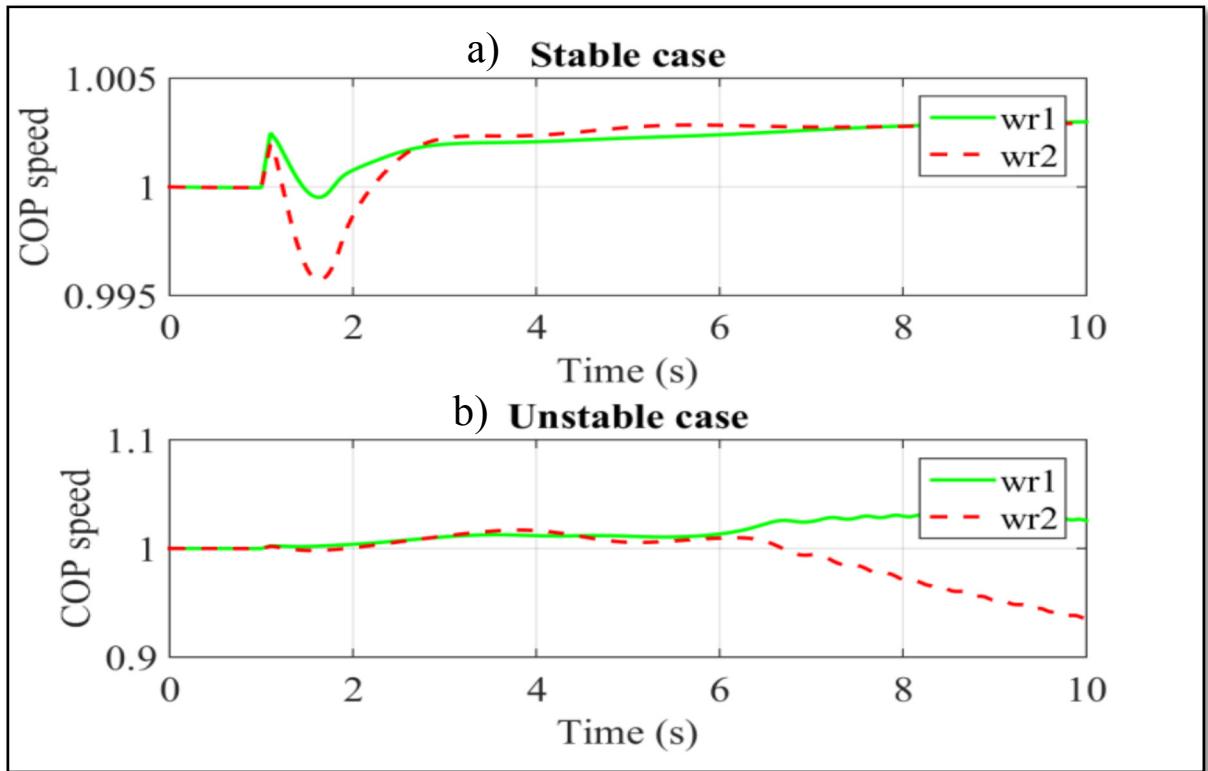


Figure 6.2 Variation du COP de vitesse pour chaque région du réseau test de Kundur (cas stable versus instable)

Partant de la même logique et en considérant les angles internes des générateurs, les angles équivalents des régions i et du réseau en entier sont donnés par les équations (6.12) et (6.13) :

$$\delta_{R_i}^{COP} = \frac{\sum_{j=1}^{m_i} \delta_j^i P_{e_j}^i}{\sum_{j=1}^{m_i} P_{e_j}^i} \quad (6.12)$$

$$\delta_{R_T}^{COP} = \frac{\sum_{k=1}^m \delta_k P_{e_k}}{\sum_{k=1}^m P_{e_k}} \quad (6.13)$$

6.2.2 Définition des nouveaux indices basés sur le COP et les mesures des synchrophaseurs

Pour détecter l'instabilité, on propose aussi dans cette thèse deux nouveaux indices basés sur le concept du COP et les mesures des synchrophaseurs. Ces indices sont obtenus à partir d'un calcul de produit scalaire modifié (*Dot Product*, DP). On s'est inspiré de la notion de DP défini dans (Fu et Bose, 1999) qui permet de trouver le point menant à l'instabilité avec les méthodes TEF. Par définition dans l'analyse de la stabilité avec le TEF, ce point est caractérisé par le premier maximum atteint par l'énergie potentielle transitoire du réseau dans sa configuration post-défaut. Le DP donne la mesure de la puissance d'accélération totale et décrit la réponse du réseau au défaut responsable de cette puissance d'accélération.

On calcule les indices γ_{COP} et κ_{COP} à partir des valeurs instantanées des COP de vitesse et des angles équivalents en appliquant les formules suivantes :

$$\gamma_{COP} = \sum_{l=1}^n \delta_{R_l}^{COP} (\omega_{R_l}^{COP} - \omega_{R_T}^{COP}) \quad (6.14)$$

$$\kappa_{COP} = \sum_{l=1}^n \omega_{R_l}^{COP} (\delta_{R_l}^{COP} - \delta_{R_T}^{COP}) \quad (6.15)$$

On fait remarquer que, contrairement aux autres indices de la littérature basés sur les méthodes TEF et les DP (Fu et Bose, 1999), γ_{COP} et κ_{COP} ne sont pas obtenus à partir du COI et de fait, ne dépendent pas de l'inertie des différentes machines du réseau. Pour prédire l'instabilité à partir de ces indices, on analyse l'allure de leur courbe représentative dans le temps. Pour des cas instables, on observe que les deux indices ont tendance à croître ou décroître rapidement pour ensuite diverger. La tendance est contraire pour des cas stables. Les indices convergent assez rapidement vers une valeur constante.

La Figure 6.3 illustre l'application de ces deux indices en considérant les mêmes défauts que dans le cas présenté précédemment. On trouve la courbe représentative de nos deux indices comme indiqué dans les équations (6.16) et (6.17).

$$\gamma_{COP} = \delta_{R_1}^{COP} (\omega_{R_1}^{COP} - \omega_{R_T}^{COP}) + \delta_{R_2}^{COP} (\omega_{R_2}^{COP} - \omega_{R_T}^{COP}) \quad (6.16)$$

$$\kappa_{COP} = \omega_{R_1}^{COP} (\delta_{R_1}^{COP} - \delta_{R_T}^{COP}) + \omega_{R_2}^{COP} (\delta_{R_2}^{COP} - \delta_{R_T}^{COP}) \quad (6.17)$$

On constate que les deux indices convergent dans un cas stable. Tandis que pour le cas instable, on remarque que légèrement avant 4.5 s, γ_{COP} commence à croître. Il en est de même pour κ_{COP} peu après 5 s.

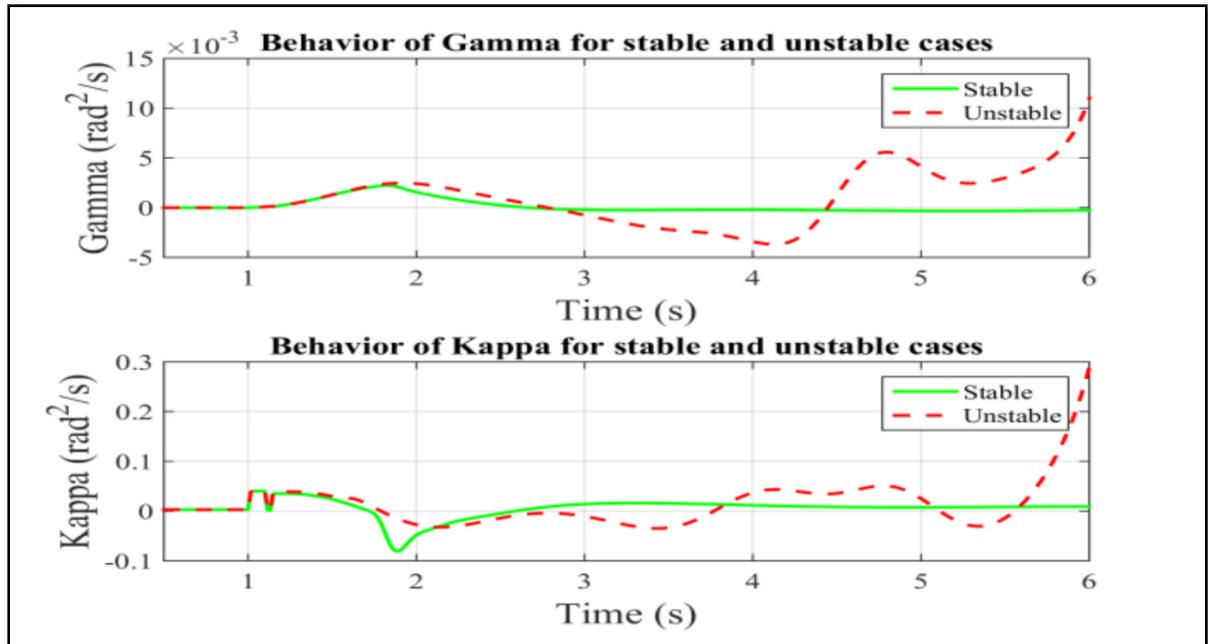


Figure 6.3 Comportement des indices γ_{COP} et κ_{COP} (cas stables versus instables)
Réseau Test Kundur

Notez que la contingence considérée pour montrer la performance du COP et des deux indices conduit à une instabilité lente. Dans le contexte de la stabilité transitoire, une instabilité lente peut être associée à une instabilité observée après plusieurs oscillations (*multi-swing instability*) tandis qu'une instabilité rapide apparaît à la première oscillation (*first swing instability*) (Susuki, Mezić et Hikihara, 2011). La Figure 6.4 montre que l'écart de 180^0 (critère d'instabilité largement adopté dans la littérature) entre les angles du rotor de deux générateurs

quelconques est atteint longtemps après l'effacement du défaut dans le cas instable. Par contre, on observe que les angles du rotor des générateurs évoluent dans le même sens pour le cas stable. Le COP de vitesse et les indices en découlant ont des comportements similaires aux angles du rotor des générateurs.

En fait, dans le cas de l'analyse de l'instabilité, le temps de préemption est le paramètre le plus important à surveiller. Pour être acceptable, il doit être suffisamment long pour permettre la réalisation des actions de contrôle (soit 10 cycles ou plus) en prenant en compte le temps de réponse du disjoncteur et tout retard de communication. Dans ce cas-ci, le temps écoulé entre la détection du COP et le temps théorique de l'instabilité est plus de 2 s.

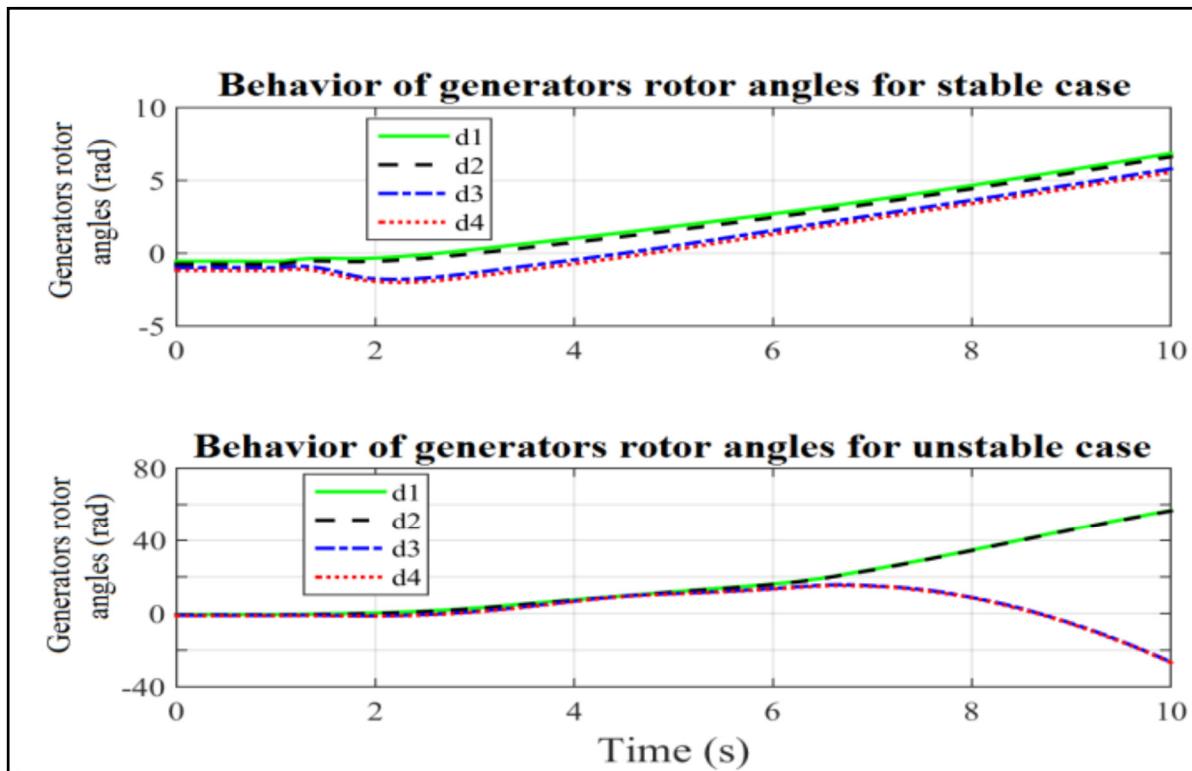


Figure 6.4 Comportement des angles du rotor des générateurs – Réseau Test Kundur

6.3 Pourquoi utiliser le centre de puissance ?

Dans cette sous-section on se propose de répondre à deux questions capitales pour cette thèse : 1) pourquoi utiliser le COP ? et 2) comment interpréter l'instabilité avec le COP proposé et les indices en découlant ?

L'émergence des différents types d'infrastructures d'énergies renouvelables sur les grands réseaux électriques vont rendre dans le futur la notion d'inertie floue et même caduque. Dans un certain sens, la notion d'inertie virtuelle initiée par certains auteurs est-elle réellement équivalente à l'inertie physique (Soni, Doolla et Chandorkar, 2013) – (Arani et El-Saadany, 2013) ? Si on considère un réseau sans générateur synchrone, on se demande quel est le sens du centre d'inertie (Zhu et al., 2011) – (Morren, Haan et Ferreira, 2005) – (Ekanayake et Jenkins, 2004) ? À date, ces questions restent légitimes et aucune réponse ni catégorique ni unanime n'est encore apportée.

En outre, les opérateurs de réseaux doivent effectuer de longs tests dispendieux pour valider les valeurs d'inertie fournies pour chacune des machines du réseau. Lors d'une redistribution de puissance pour des raisons économiques ou suite à une contingence, certains générateurs peuvent être désaffectés. À ce moment-là, les informations relatives à l'inertie ne sont pas facilement accessibles au PDC/WAMS où les données des PMU sont stockées et traitées en temps réel pour permettre un contrôle efficace et une protection rapide des équipements. Les valeurs exactes d'inertie associées à un générateur donné doivent être connues en tout temps. En effet, comme prouvé dans (Naik, Nair et Swain, 2016), de légères variations dans la valeur de l'inertie peuvent avoir des effets néfastes sur la détection de la stabilité transitoire.

En examinant l'équation (6.18), il est clair que le COP et le COI ne diffèrent que par le facteur de poids appliqué à ω_k , lesquels poids sont le moment d'inertie M_k pour le COI et la puissance active du générateur P_{ek} pour le COP. La comparaison entre les deux concepts se résument alors à savoir si ces deux facteurs de poids sont proportionnels entre eux.

$$\omega_{COI} = \frac{\sum_{k=1}^{N_g} \omega_k M_k}{\sum_{k=1}^{N_g} M_k} \quad \text{et} \quad \omega_{COP} = \frac{\sum_{k=1}^{N_g} \omega_k P_{e_k}}{\sum_{k=1}^{N_g} P_{e_k}} \quad (6.18)$$

$$H_k = \frac{\text{Énergie stockée à la vitesse nominale du générateur } j \text{ (MW.s)}}{MVA_{j,nominal}} \quad (6.19)$$

Pour adresser cette question de proportionnalité, on va réécrire l'expression du COI suivant les équations (6.20) et (6.21) :

$$\omega_{COI} = \frac{\sum_{k=1}^{N_g} \omega_k \frac{2H_k}{\omega_{j0m}^2} S_k^{VA-Base}}{\sum_{k=1}^{N_g} \frac{2H_k}{\omega_{j0m}^2} S_k^{VA-Base}} = \frac{\sum_{k=1}^{N_g} \omega_k \eta_k S_k^{VA-Base}}{\sum_{k=1}^{N_g} \eta_k S_k^{VA-Base}} \quad (6.20)$$

$$M_k = \frac{2H_k}{\omega_{j0m}^2} S_k^{VA-Base} \quad \text{et} \quad \eta_k = \frac{2H_k}{\omega_{j0m}^2} \quad (6.21)$$

Avec N_g le nombre de générateurs du réseau ayant comme moment d'inertie M_j et vitesse mécanique nominale ω_{j0m} (mech. rad/s), $S_j^{VA-Base}$ la puissance apparente nominale (VA) et η_j une constante dépendant du générateur (Kundur, Balu et Lauby, 1994). Selon les équations (6.18) à (6.21), les coefficients de poids du COI sont fonction de la vitesse apparente nominale des générateurs. De plus, on sait que la constante d'inertie H_j varie dans un intervalle bien défini de 2 à 10 (Pavella, Ernst et Ruiz-Vega, 2012) et qu'elle tend à augmenter quand la vitesse mécanique nominale augmente (i.e. des groupes thermiques aux petites unités hydro-électriques). Il s'ensuit que le terme le plus important à considérer dans l'équation (6.20) est la puissance nominale du générateur $S_j^{VA-Base}$. A fortiori, les machines ayant une plus grande puissance apparente nominale (Voltampères) contribuent davantage que celles ayant de plus faibles puissances dans la formule du COI. Partant de ce constat, le lien entre le COI et le COP se fait automatiquement : un générateur qui produit plus de puissance active (MW) contribue plus qu'un générateur produisant moins dans la formule du COP. Concrètement, si toutes les machines fonctionnent à leur puissance apparente nominale (VA), la différence entre le COP

et le COI est négligeable. En revanche, si un générateur de grande puissance produit une puissance relativement faible (exemple $P_{e_j} \leq 0.5 \text{ p.u.}$), le COP sous-estime son coefficient de poids. Cette problématique ne risque pas trop de se poser car, pour des raisons économiques, les générateurs opèrent généralement autour de leur puissance nominale (i.e. $P_{e_j} \approx 0.9 - 1 \text{ p.u.}$).

En résumé à la première question soulevée, le COP présente l'avantage de ne pas avoir à calculer explicitement le TEF. Sous les conditions normales d'opération du réseau dictées par un TSC-OPF et par une gestion efficace et efficiente des groupes de générateurs, le COI et le COP ont tendance à suivre des modèles de poids de pondération similaires comme indiqué dans l'équation (6.18). Cela est dû au fait que les paramètres pris en compte pour le COP et le COI sont étroitement liés ; en l'occurrence, pour le COI le poids associé à chaque centrale est basé sur sa taille indiquée sur la plaque signalétique des générateurs (inertie) et que pour le COP, il dépend de la puissance active fournie en temps réel par la centrale. Le COP se révèle donc comme une approximation raisonnable du COI, pour les situations où les inerties ne sont pas connues ou ne sont pas fiables, avec le grand avantage que son calcul peut être effectué uniquement à l'aide des données PMU, sans aucune information du SCADA ni aucune communication avec l'EMS.

Dans cette thèse, on a proposé le nouveau concept du COP et deux indices de détection d'instabilité qui en découlent. Avec les formules modifiées présentées dans les équations (6.7) – (6.8) et (6.12) à (6.15), on analyse la stabilité transitoire en deux étapes : 1) en scrutant l'écart entre le COP de vitesse de deux régions quelconques du réseau et 2) en statuant sur la divergence ou la convergence des deux indices proposés.

Pour ce qui est du COP de vitesse, on compare la courbe représentative de chacune des régions l'une par rapport à l'autre. L'écart maximal entre deux régions ne doit pas dépasser un certain seuil qui peut être une valeur prédéfinie. En cas d'instabilité lent, l'écart n'est pas brusquement grand mais prend du temps à se creuser, comme on le voit sur la Figure 6.4. Dans le cas d'instabilité rapide, la déviation devient rapidement très importante. En ce qui concerne les

indices γ_{COP} et κ_{COP} , on inspecte le comportement de chacun d'eux séparément dans le domaine temporel. La divergence reflète le fait que le réseau n'a pas pu maintenir le synchronisme après le défaut, i.e. que le réseau est devenu instable.

Notez qu'en augmentant le nombre de mesures de phaseurs sur le réseau, on peut améliorer la précision du calcul du COP de vitesse et du même coup améliorer l'effet de redondance.

6.4 Description de la nouvelle méthodologie de prédition en temps réel

Cette sous-section expose les différentes étapes de la nouvelle méthodologie de prédition de stabilité en temps réel proposée dans le cadre de cette thèse. On décrit tout d'abord les différents scénarios considérés et les séquences de calcul réalisées. Ensuite, on discute de la fenêtre glissante adoptée et de ses composantes. En dernier lieu, on fait état de la méthode de mise à jour proposée pour actualiser la base de données avec l'arrivée de nouvelles données.

6.4.1 Méthodologie globale de prédition

La méthodologie adoptée est articulée autour de quatre grandes phases : la collecte des données, la compilation et l'organisation des données, la modélisation et la prédition. La phase de collecte des données consiste à capter les données des PMUs. Se trouvant à des emplacements stratégiques du réseau, les PMUs fournissent les phaseurs (l'amplitude et la phase) de tension et de courant des jeux de barres où ils sont installés. On dispose alors de $|V_{bus_{ij}}|$, $\theta_{V_{bus_{ij}}}$, $|I_{bus_{ij}}|$ et $\theta_{I_{bus_{ij}}}$. Ensuite, on passe à la phase d'organisation des données.

Les données sont compilées dans deux blocs : Bloc I et Bloc II. Grâce aux PMUs placés aux bus des générateurs, on peut estimer la vitesse (ω_i) et l'angle du rotor (δ_i) des machines. Ces variables ainsi que les puissances électriques des machines (P_{e_i}) sont stockées dans Bloc I. Les autres mesures de phaseurs provenant des PMUs se trouvent dans le Bloc II. On y ajoute la puissance transitant sur le bus $P_{bus_{ij}}$ et une variable V_{cos} obtenue à partir de l'équation (6.22).

$$V_{cos} = |V_{bus_{ij}}| \cos \phi_{bus_{ij}} \quad (6.22)$$

Sachant toutes les régions cohérentes du réseau et les angles et vitesses du rotor des machines dans chacune d'elles, on peut déduire le COP de vitesse ω_{R_i} et l'angle équivalent δ_{R_i} . De ces valeurs, on déduit les indices temporels κ et γ . Toutes les dernières variables mentionnées se retrouvent dans le Bloc II.

La troisième phase se rapporte à la conception du modèle de décision. Pour cela, on envisage trois scénarios pour lesquels on construit une forêt aléatoire d'arbres de décision. Dans le Scénario 1, toutes les variables des Bloc I et II indistinctement jouent le rôle de prédicteurs. L'objectif, entre autres, est de cibler les variables les plus importantes pour l'analyse de la stabilité et de voir l'ordre d'importance des variables définies dans le cadre de cette thèse dans la prise de décision. Dans le Scénario 2, on réduit notre base de données soumise à l'apprentissage en n'utilisant que les variables du Bloc II comme prédicteur. L'objectif est de comparer les résultats obtenus avec ceux obtenus dans le Scénario 1 et de vérifier dans quelle mesure la réduction du nombre de prédicteurs affecte la prédiction. On étudie ainsi la valeur ajoutée de la notion de COP et des indices proposés. Pour généraliser la méthodologie indépendamment du réseau et du nombre de PMUs installés, on compose un Scénario 3 en intégrant toutes les données de simulation obtenues. Le minimum et le maximum des puissances actives des jeux de barres, le minimum et le maximum des V_{cos} des bus et les indices γ et κ jouent le rôle de prédicteurs. On ne tient compte ici que des valeurs maximales ou minimales car on sait quand le réseau devient instable un certain seuil de puissance ou de tension a été dépassé par au moins un générateur. Cette réduction aide également à alléger l'apprentissage hors-ligne et à réduire le temps de calcul. Les résultats de prédiction sont comparés à ceux des Scénarios 1 et 2.

En ayant un modèle de RF pour chacun des scénarios, on peut aisément procéder à la prédiction de l'état du réseau pour un événement donné. Suite à la prédiction, pour des cas stables, on effectue des analyses des contingences alors que pour les cas instables, il faut appliquer des

actions correctives. La Figure 6.5 reproduit les différentes étapes de la méthodologie décrite précédemment pour la prédition de l'instabilité en temps réel.

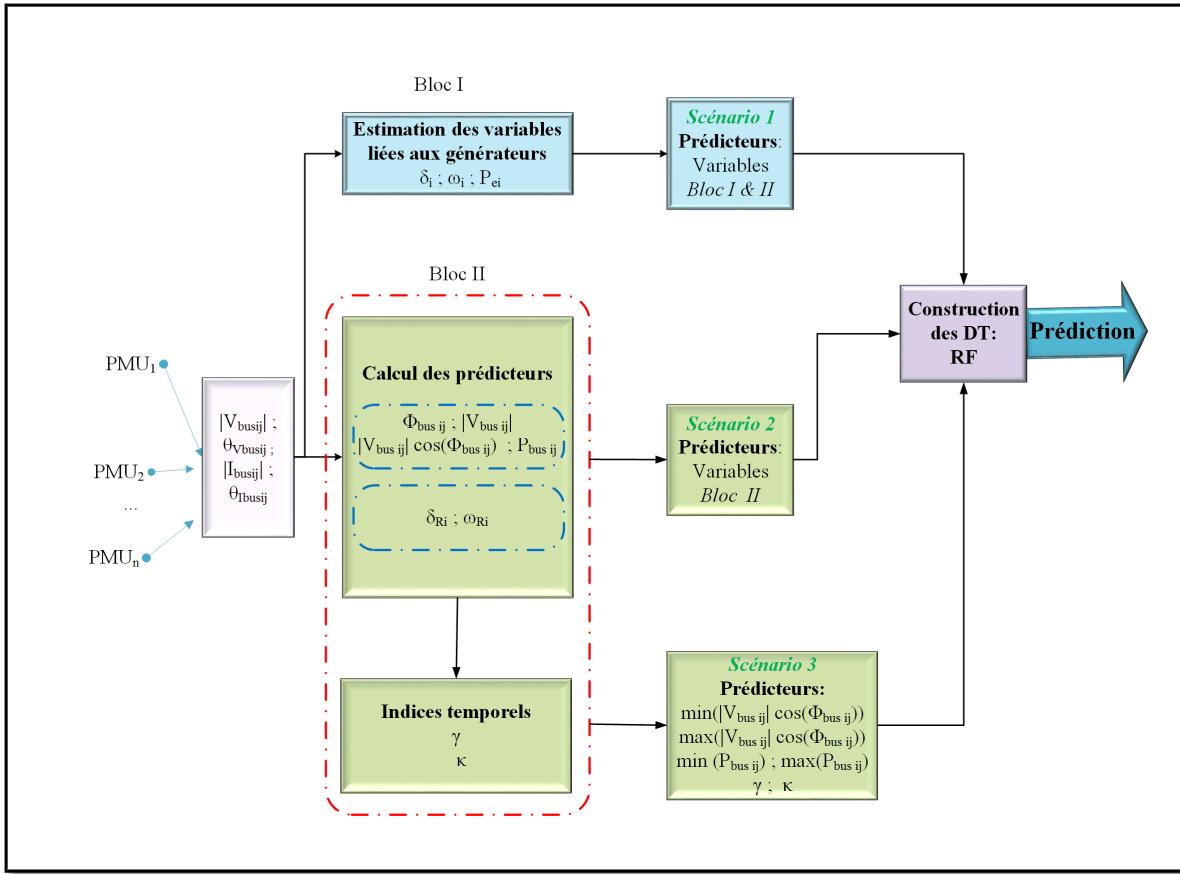


Figure 6.5 Description des séquences de calcul et des scénarios pour la prédition de l'instabilité

6.4.2 Approche de la fenêtre glissante (*sliding window*) versus *snapshot*

Généralement dans la littérature, on utilise un *snapshot* pour l'étude de la stabilité. Il s'agit d'une photo instantanée du réseau ou de la réponse du système pour un intervalle de temps donné après l'effacement du défaut (Zheng, Malbasa et Kezunovic, 2013) – (Dmitrova et al., 2015). Les temps couramment utilisés sont 150 ms et 300 ms (Kamwa, Grondin et Loud, 2001b) – (Kamwa, Samantaray et Joos, 2012). Contrairement à ces travaux, on propose de recourir à une approche de fenêtre glissante (*Sliding Window*, SW) dans le domaine du temps. Cette approche consiste à calculer la moyenne sur une période de temps donné au lieu de

considérer tous les éléments de l'intervalle un par un. Les principaux avantages d'une telle démarche sont la mise à jour automatique avec de nouvelles données entrantes et l'importance plus élevée qui leur est donnée sans perdre l'impact des précédentes valeurs. De plus, contrairement, aux *snapshots* pour lesquels on a besoin d'un détecteur de défaut pour repérer une contingence sur le réseau, le SW permet de décider en tout temps pour les cas d'instabilités lentes et rapides. L'Algorithme 6.1 résume les procédures suivies dans la formation de la fenêtre glissante.

Algorithme 6.1 Vue algorithmique de la formation de la fenêtre glissante

1. Déterminer la séquence S ;
2. Fixer la largeur de la fenêtre ($width_{sliding}$) ;
3. Définir le pas de glissement pour chaque itération ($incr_{sliding}$)
4. Calculer le nombre de fenêtre (nbr_{window})

$$nbr_{window} = \frac{(Longueur(S) - width_{sliding})}{incr_{sliding}}$$

5. Calculer la moyenne pour chaque fenêtre


```
for      i = 1 : nbrwindow
              meanwin(i) = mean(S(i: i + widthsliding))
      end
```

En considérant le même défaut introduit dans ce chapitre, la Figure 6.6 compare les valeurs réelles de l'indice κ_{COP} à celles calculées avec un SW. On constate que la différence entre les deux valeurs est très faible pour les cas stables et instables quand le signal n'est pas bruité.

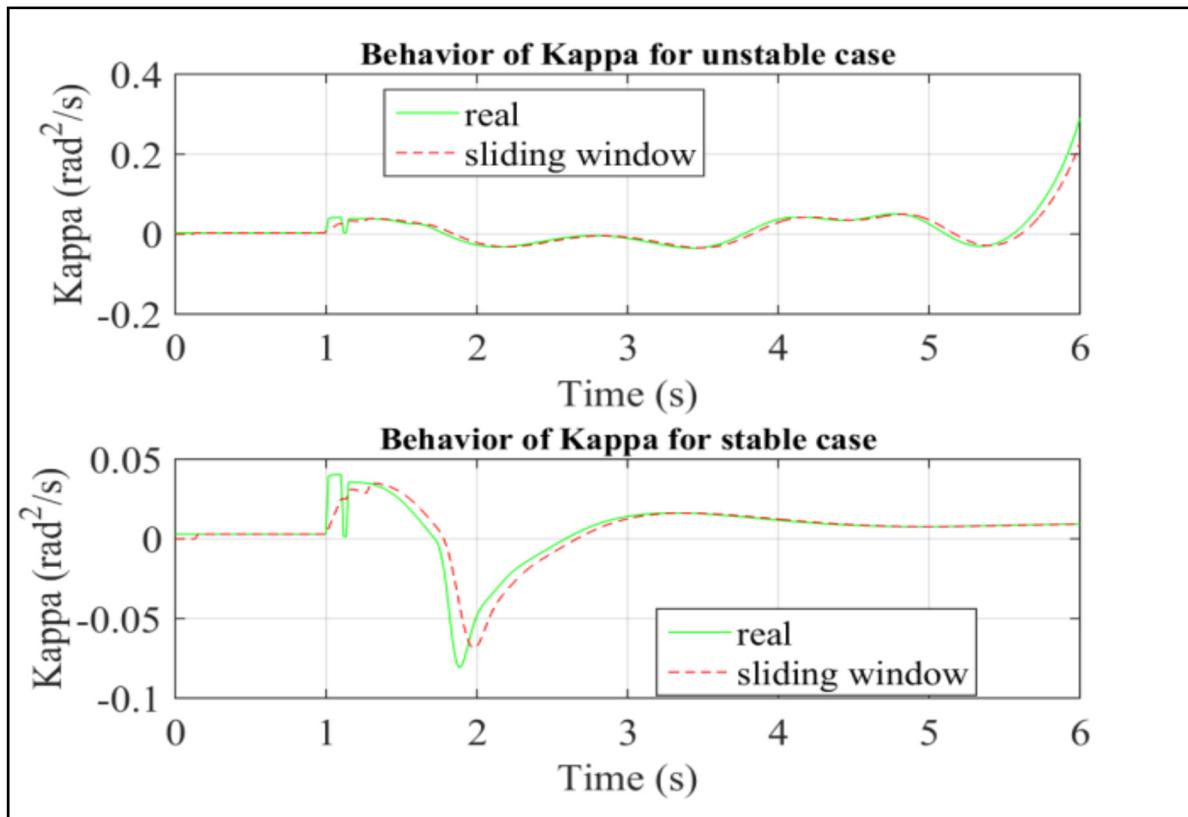


Figure 6.6 Comportement de κ_{COP} avec une fenêtre glissante pour des cas stables et instables – Réseau Test Kundur

6.4.3 Mise à jour automatique de la base de données et du modèle de prédiction

Une question courante qui revient avec l'exploitation du forage des données pour la prédiction est : une fois le modèle établi et qu'on observe des changements de routine dans les états du réseau, faut-il procéder ou non à un réapprentissage du RF ? Dans une certaine mesure, on peut dire qu'il est nécessaire de le faire. Quoique le RF soit robuste et qu'il soit conçu pour extraire sur un large éventail de conditions du réseau l'information « essentielle » quant à son état ; comme toute méthode inductive, les résultats de prédiction sont garantis pour des dynamiques assez proches de celles manipulées dans la phase d'apprentissage. Tandis que le nombre d'états possibles du réseau est infini, la dynamique du réseau est limitée à un nombre fini de groupes de caractéristiques tels que l'instabilité rapide, l'instabilité lente, l'instabilité de tension, etc...

Il est donc possible d'améliorer la qualité du RF en disposant de contrôles périodiques ou après des changements majeurs sur le réseau. Ces contrôles ne devraient pas être trop différents de ce qui est fait aujourd'hui pour les systèmes de protection spéciaux (*Special Protection Systems*, SPS) critiques dont les paramètres sont mis à jour sur une base de secondes ou de minutes selon les états du réseau (Trudel, Gingras et Pierre, 2005). Ces deux travaux (Sun et al., 2007) – (Alcheikh-Hamoud et al., 2009) donnent des techniques de mise à jour des données sur une base horaire pour la prédiction avec les DT. Compte tenu de la performance remarquable des prédicteurs proposés dans cette thèse, on soumet à la Figure 6.7 un nouveau dispositif pour mettre à jour le modèle de prédiction sur une base périodique en s'alimentant de nouvelles dynamiques du réseau. Ces dernières vont enrichir la base de données et rendre notre dispositif adaptatif.

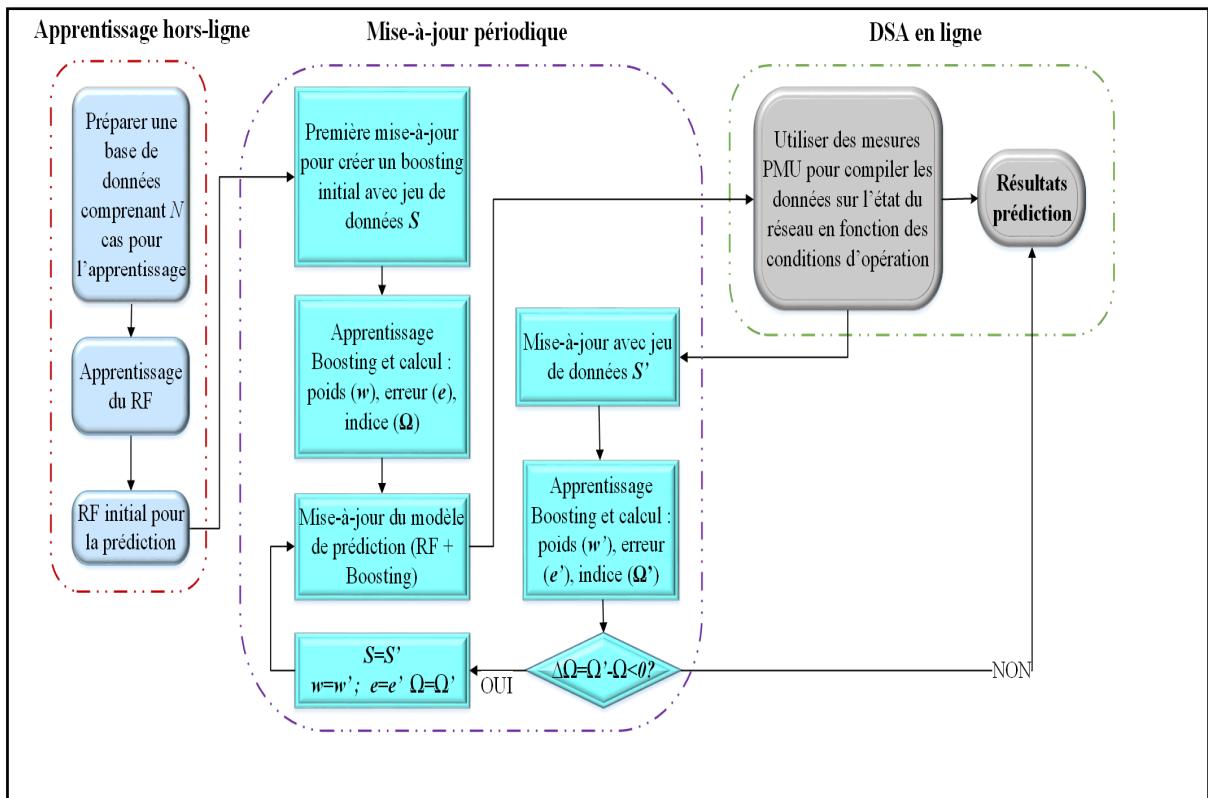


Figure 6.7 Dispositif de mise-à-jour périodique du modèle de prédiction de l'instabilité

Avec ce nouvel outil adaptatif de mise à jour du modèle de prédiction, l'état du réseau va être déterminé par une double vérification. La première étape revient à soumettre l'événement en question à un modèle découlant d'un RF initial. Cette base de données étant la plus dense, le RF est plus approprié pour produire un modèle de prédiction car les arbres élagués construits dans l'apprentissage sont décorrélés. Ainsi, on maximise la diminution de la variance, ce qui implique une réduction des erreurs. Le RF dans ce contexte-ci tient compte d'une configuration globale du réseau avec le maximum de cas possibles (par exemple, un apprentissage des données recueillies sur plusieurs années).

Avec l'acquisition de nouvelles données des PMUs, sur une base régulière (par exemple, semaine et/ou mois), on peut réaliser un deuxième modèle de prédiction à l'aide d'un boosting pour évaluer l'état du réseau. On privilégie un boosting à ce stade car on sait que la base de données à étudier est d'une dimension moindre et de plus, que le boosting ne se concentre que sur les apprenants faibles. Dans le contexte des DT, il s'agit d'arbres peu profonds ; à noter qu'on se limite à 100 arbres. Le boosting amoindrit donc l'erreur principalement en réduisant le biais. On se sert d'un premier jeu de données S pour entraîner ce second modèle noté B . Comme sortie de B , on stocke les poids w_i des arbres dans la prise de décision ainsi que les erreurs e_i associées à chacun de ces poids. On définit un indicateur Ω tel qu'énoncé dans l'équation (6.23) :

$$\Omega = \frac{\sum_{i=1}^{n_{tree}} w_i e_i}{\sum_{i=1}^{n_{tree}} w_i} \quad (6.23)$$

Cet indicateur a pour rôle essentiel de signaler si le modèle de boosting utilisé pour la prédiction doit être mis à jour ou non. Quand un nouveau jeu de données S' provenant des PMUs rentre, on déduit l'indicateur Ω' correspondant au modèle B' . Si Ω' est plus petit que Ω , on conserve le modèle B' comme deuxième modèle de prédiction ; sinon on garde le modèle B .

6.5 Conclusion

Dans ce chapitre, on présente le nouveau concept de centre de puissance pour l'étude de l'instabilité. Ce concept a les mêmes fondements théoriques que le COI mais présente l'avantage de prendre en compte les valeurs de puissance effectives du réseau déduites des mesures des PMUs. Pour chaque région cohérente du réseau, on définit un COP de vitesse et un angle équivalent, à partir desquels résultent deux indices γ et κ participant à la détection de l'instabilité. Au regard du COP de vitesse, le réseau est instable si pour deux régions quelconques on observe une déviation entre les COP correspondants. Quant aux indices, leur tendance à diverger confirme que le réseau est instable. On décrit également la méthodologie utilisée pour prouver que les prédicteurs proposés dans le cadre de cette thèse sont performants. Les étapes menant jusqu'à la formation du RF et de l'exécution du modèle de prédiction sont développées. Tous les scénarios pris en compte sont détaillés ainsi que la méthode de la fenêtre glissante exploitée en opposition au snapshot largement utilisé dans la littérature. Pour considérer les mises à jour des données, on propose aussi un nouveau dispositif couplant un RF à un boosting pour la prédiction de l'état du réseau en temps réel. Somme toute, l'originalité du travail est dans les variables utilisées comme prédicteurs. Le concept de COP utilisé dans notre approche permet avec un nombre limité de prédicteurs de dissocier les cas stables des cas instables en tout temps avec l'aide de la fenêtre glissante.

CHAPITRE 7

RÉSULTATS ET VALIDATION

7.1 Introduction

Ce chapitre se consacre à la présentation des résultats et à la validation de la performance de la méthodologie proposée tout au long de cette thèse. Pour les simulations, on considère trois réseau tests : celui de Kundur, le réseau Australien simplifié et le New England. Leurs caractéristiques sont données dans les Annexes I et II. Tout d'abord, on explique comment la base de données a été formée et les différents cas de figure étudiés. Ensuite, on montre le comportement du COP de vitesse et des indices proposés suite à des contingences menant ou non à une instabilité sur le réseau. On procède aussi à une évaluation de la performance de l'erreur out-of-bag. De plus, on analyse l'efficacité des prédicteurs proposés pour la construction du RF et du boosting et on livre une évaluation de la performance du modèle de prédiction. À cela, on ajoute les résultats obtenus lors de la mise à jour du modèle de prédiction. En guise de conclusion, on effectue une discussion des résultats dans laquelle certaines considérations sont abordées.

7.2 Présentation des cas étudiés et formation de la base des données

La Figure 7.1 affiche le cadre d'apprentissage pour l'évaluation de la sécurité du réseau. On y retrouve toutes les étapes partant de la génération de la base de données jusqu'à la prédiction de l'état du réseau suite à une contingence à partir des règles établies dans le modèle.

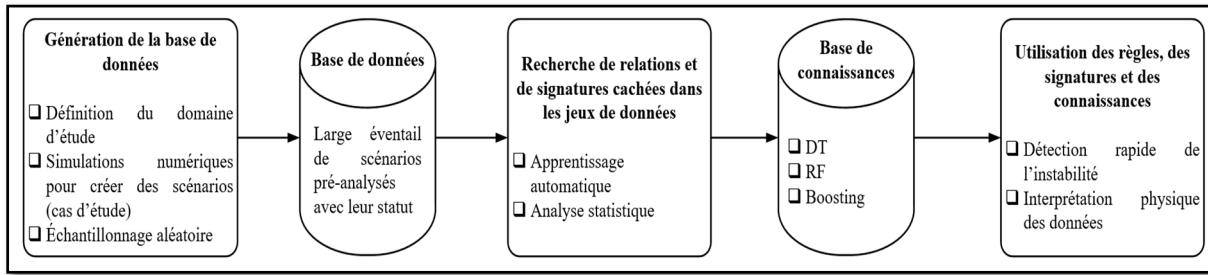


Figure 7.1 Cadre d'apprentissage pour l'évaluation de la sécurité du réseau

La base de données est bâtie en appliquant un défaut triphasé suivi de la perte d'une ligne à chaque jeu de barre, exception faite des bus liés aux générateurs. La durée des défauts provoqués (*clearing time*) varie de 1 à 12 cycles. On génère plusieurs points d'opération afin d'entraîner le RF et le boosting sur un réseau fonctionnant aux conditions limites et aussi pour étudier la robustesse de la méthodologie quant à leurs variations. Pour cela, on change les charges (active et réactive en même temps) de 90% à 120% de leurs valeurs nominales. Pour le réseau test New England, on déclenche 12 défauts sur chacun des 29 barres pour chaque point d'opération, ce qui fait un total de 1 044 cas d'observations. Suivant cette même logique, on obtient 1 620 cas pour le réseau Simplifié Australien et 144 cas pour le réseau de Kundur. Comme on adopte une fenêtre glissante dans l'analyse de la stabilité, il faut compter 600 points pour chacun des cas. Les PMU connectés sur le réseau fournissent des mesures pour estimer les variables des générateurs et pour calculer tous les prédicteurs mentionnés dans le diagramme de la Figure 6.5. En vue de compléter la base de données, tous les attributs sont stockés avec leur état « Stable » si le réseau est stable et « Instable » sinon. Il s'agit donc là d'un large éventail de prédicteurs pré-analysés avec leur statut $C_i = \{Stable, Instable\}$. On rappelle pour l'apprentissage que le réseau est instable si on constate un écart de 180° entre les angles du rotor de deux machines quelconques (Narasimhamurthi et Musavi, 1984) pendant les 5 secondes suivant l'effacement du défaut.

La prochaine étape est de rechercher des relations et des signatures cachées dans le jeu de données ainsi pré-analysé. On fait appel au logiciel libre (open-source) **R** (Liaw et Wiener, 2002) pour la mise en œuvre d'arbres de décision conventionnels, de forêts aléatoires et de

boosting. On examine les trois scénarios rapportés à la Figure 6.5 à partir d'une fenêtre glissante et de deux snapshots dont un de 150 ms et l'autre de 300 ms. Lors de la construction du modèle basé sur le boosting, on se limite à 100 arbres. En ce qui concerne le RF, pour la construction du modèle, on le configure pour qu'il choisisse aléatoirement un échantillon constitué 70% de la base de données pour la phase d'apprentissage et on se limite à un nombre maximal de 500 arbres. Contrairement au boosting, on s'assure que chacun des arbres soit élagué pour qu'ils atteignent leur taille maximale. Ainsi, on obtient un ensemble de règles et de signatures pour prédire l'état du réseau à partir de nouvelles données.

7.3 Comportement des indices proposés

Dans cette section, on se propose de montrer comment les COP de vitesse ainsi que les indices γ_{COP} et κ_{COP} définis dans le cadre de cette thèse permettent de détecter l'instabilité. Pour cela, on provoque des défauts sur le réseau New England et sur le réseau Simplifié Australien, puis on analyse l'évolution dans le temps des dernières variables citées.

On commence par inspecter l'évolution des valeurs réelles versus celles calculées à partir de la fenêtre glissante. On applique un défaut triphasé lié à la terre d'une durée de 8 cycles au jeu de barre numéroté 8 sur le réseau test New England (voir Annexe I) à $t = 1$ s. Dans un premier temps, aucune action n'est posée suite à cette contingence, on constate que le réseau tombe dans une instabilité lente. Dans un second cas, on rétablit la stabilité grâce à des compensateurs statiques-var (*Static-Var Compensator*, SVC) et des stabilisateurs de puissance conventionnels (*Power System Stabilizer*, PSS). La Figure 7.2 montre l'allure des courbes représentatives (valeurs réelles versus fenêtre glissante) de γ_{COP} dans les deux cas.

On constate que les allures sont similaires avec un léger décalage au niveau des valeurs prises par la fenêtre glissante (taux d'échantillonnage plus petit d'un facteur $incr_{sliding}$). On en déduit que l'approximation faite va donner des résultats du même ordre que les valeurs réelles avec l'avantage de prendre en compte des valeurs antérieures dans la décision.

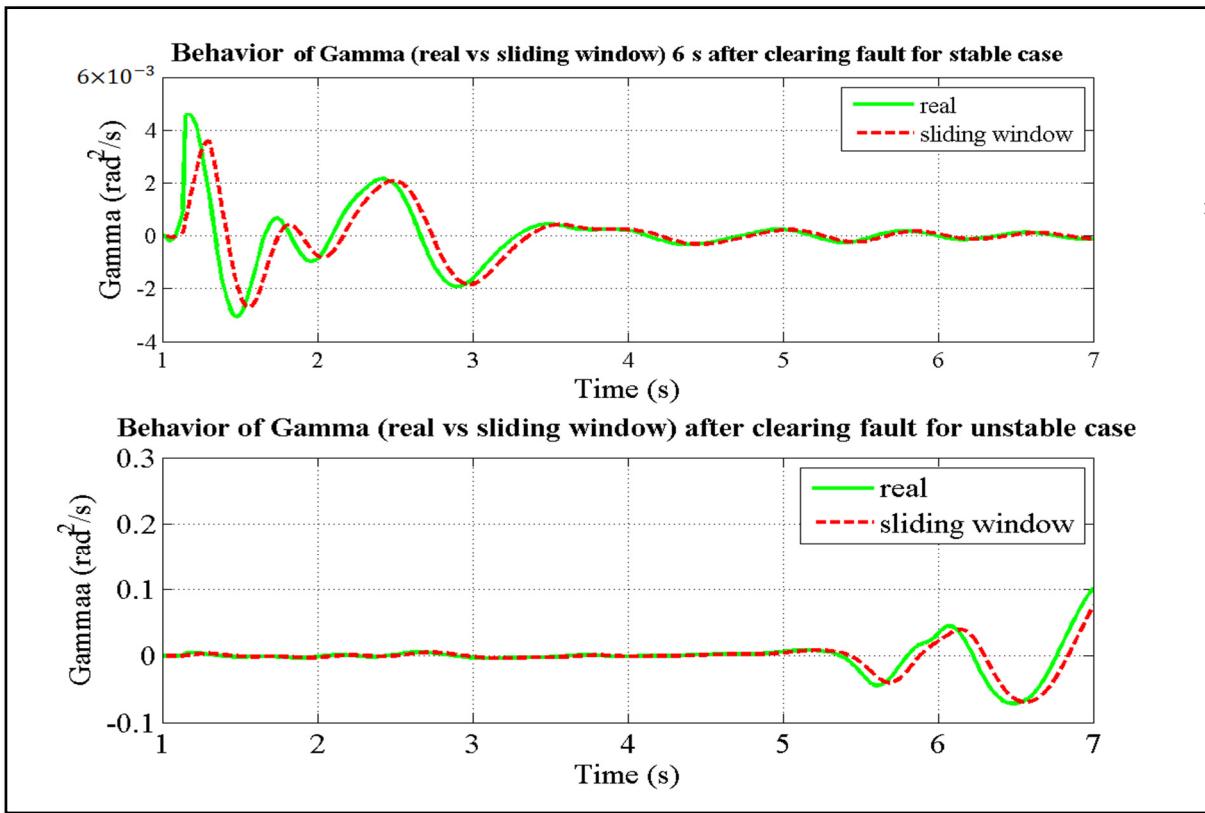


Figure 7.2 Comparaison entre les valeurs réelles et celles d'une fenêtre glissante pour des cas stables et instable – Réseau Test New England

Sur la Figure 7.3, on compare les variations de γ_{COP} et κ_{COP} 6 s après avoir effacé le défaut pour un cas stable et un cas instable. Remarquez que pour un cas stable, à partir de $t = 4s$, les deux indices commencent à se stabiliser autour d'une valeur constante. On peut se servir de la figure précédente pour avoir une meilleure vue sur l'indice γ_{COP} . Cette stabilisation s'explique par le PSS enclenchée pour ne pas perdre le synchronisme. Dans le cas contraire, il apparaît clairement que κ_{COP} commence à diverger au voisinage de $t = 4s$. La détection de l'instabilité se fait donc à partir de cet instant.

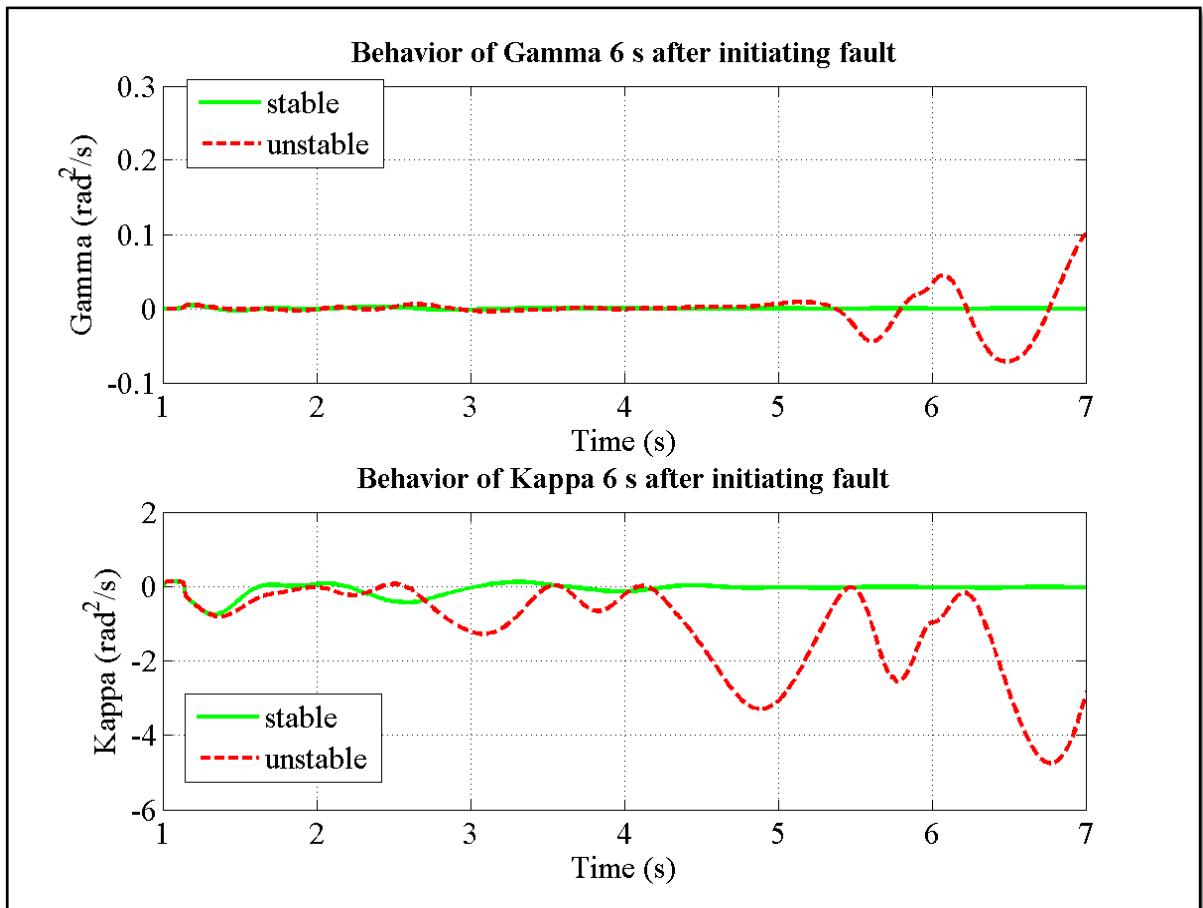


Figure 7.3 Comportement de γ_{COP} et κ_{COP} 6 s après l'effacement du défaut
Réseau Test New England

La Figure 7.4 montre que l'instabilité, du point de vue théorique (écart angulaire de 180 degrés entre deux machines quelconques), survient après $t = 9$ s. En examinant les courbes des COP de vitesse, on aperçoit qu'une des régions s'est détachée des autres bien avant $t = 9$ s. De plus, en se référant aux indices γ_{COP} et κ_{COP} , on déduit que le temps de préemption est supérieur à 5 s. Il s'ensuit que les indicateurs proposés permettent de détecter l'instabilité correctement dans des délais acceptables pour l'application d'actions adéquates.

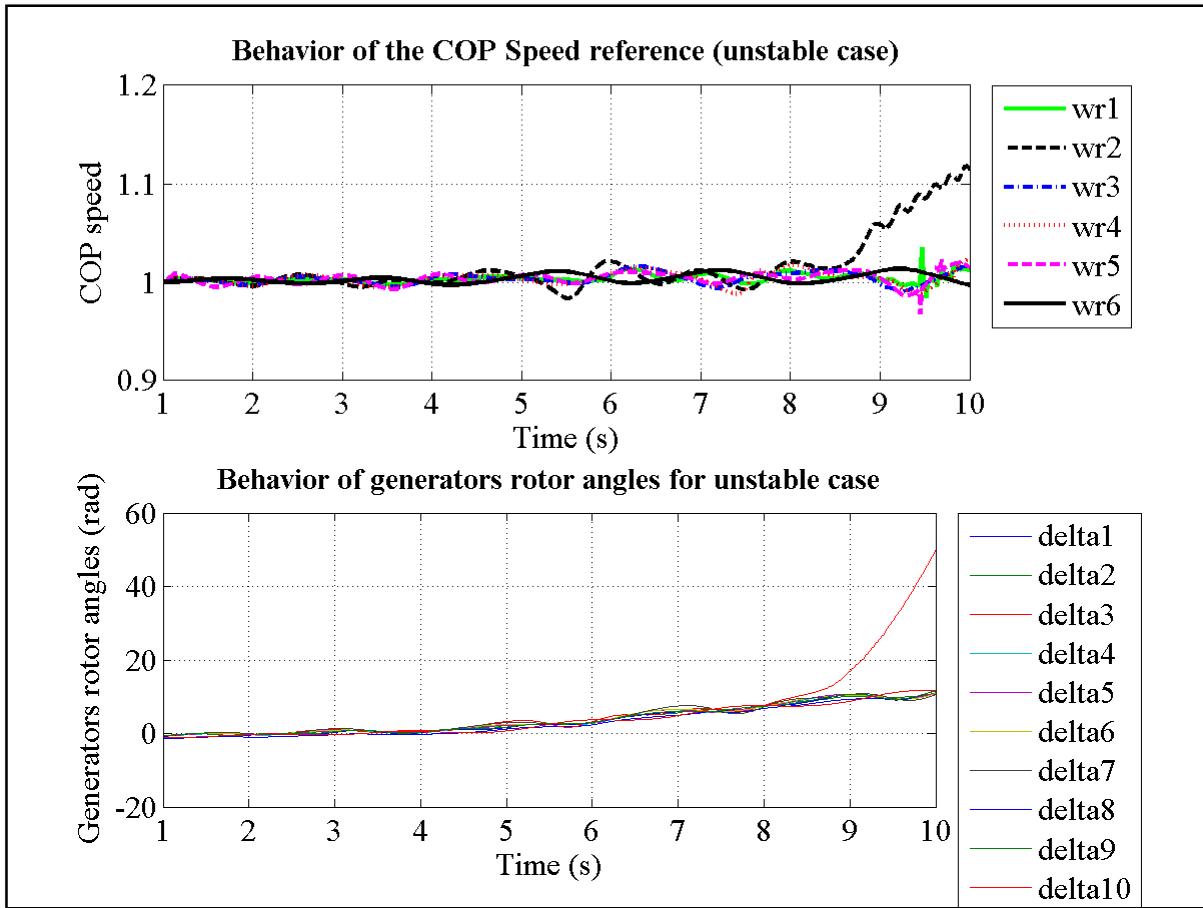


Figure 7.4 Comportement des COP de vitesse versus les angles du rotor des générateurs pour un cas instable – Réseau Test New England

La Figure 7.5 regroupe une comparaison générale entre plusieurs variables du réseau pour détecter l'instabilité. Dans la colonne de gauche, le réseau est instable suite à la contingence décrite ci-avant. On peut voir le détachement du COP de vitesse d'une des régions par rapport au groupe et la divergence totale de γ_{COP} et de κ_{COP} . On constate également que la tension des jeux de barre ne se stabilisent pas après l'événement. À l'opposé, en appliquant une compensation au moyen des SVC et des PSS, on remarque que tous les indicateurs proposés pour étudier l'instabilité convergent ; ce qui confirme que le réseau est stable.

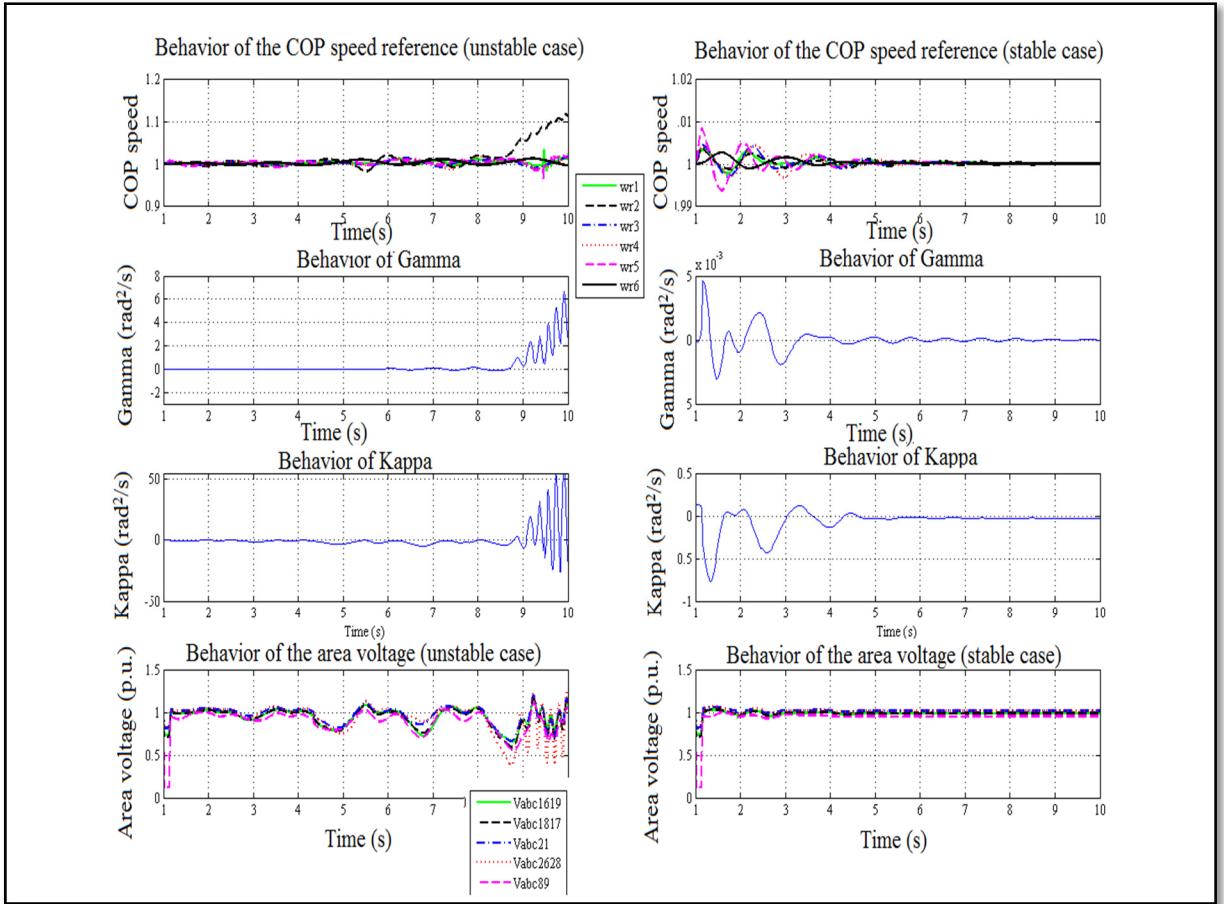


Figure 7.5 Comportement des variables pour un cas instable sans (gauche) et avec (droite) des mesures correctives – Réseau New England

À noter que dans le cas d'une instabilité rapide, les COP de vitesse des régions se détachent plus rapidement dépendamment de la sévérité de la contingence. La Figure 7.6 illustre bien ce cas. Il s'agit d'un défaut produit au bus 28 à $t = 1\text{ s}$ puis effacé 8 cycles plus tard. Si on n'applique aucune action, le réseau devient irrémédiablement instable à $t = 6.25\text{ s}$. Le COP de vitesse d'une des régions commence à s'éloigner à $t = 4.95\text{ s}$. Par contre, dans le cas stable obtenu grâce à des SVC et des PSS, on note que les indices γ_{COP} et κ_{COP} convergent et que les COP de vitesse évoluent de manière cohérente.

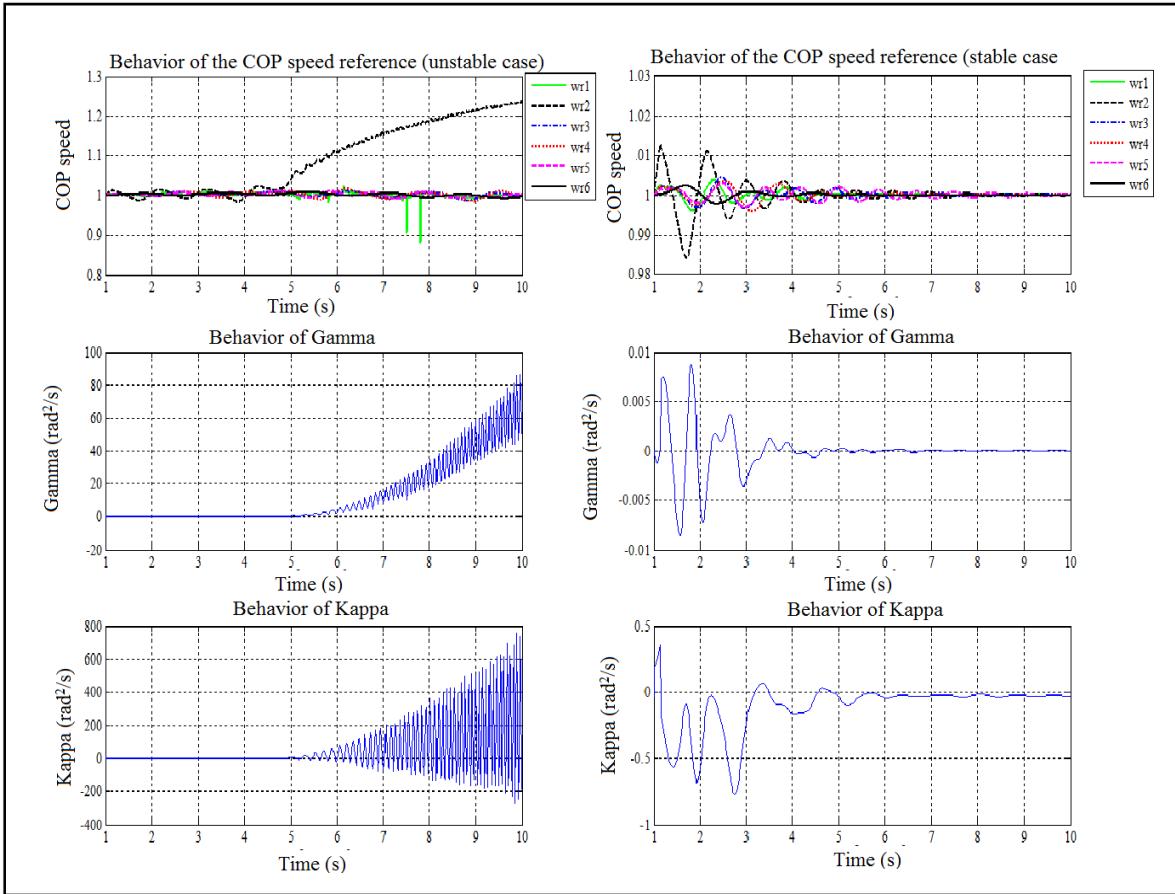


Figure 7.6 Courbe représentative des indicateurs d'instabilité pour une instabilité rapide sans (gauche) et avec (droite) des actions correctives – Réseau New England

Sur le réseau Simplifié Australien, amplement détaillé en Annexe II, on inflige un défaut triphasé lié à la terre d'une durée de 10 cycles au jeu de barre numéroté 310 à $t = 2\text{ s}$. Théoriquement, le réseau perd complètement son synchronisme à $t = 6.38\text{ s}$. On compare sur la Figure 7.7, les valeurs réelles et de la fenêtre glissante prises par les indices γ_{COP} et κ_{COP} . Dans les deux cas, on relève que les allures sont équivalentes avec une marge d'erreur très faible. En outre, on décèle que l'instant $t = 5.5\text{ s}$ marque le début de la divergence deux indices ; cela confirme que ces derniers ont bel et bien détecté l'instabilité du réseau à venir.

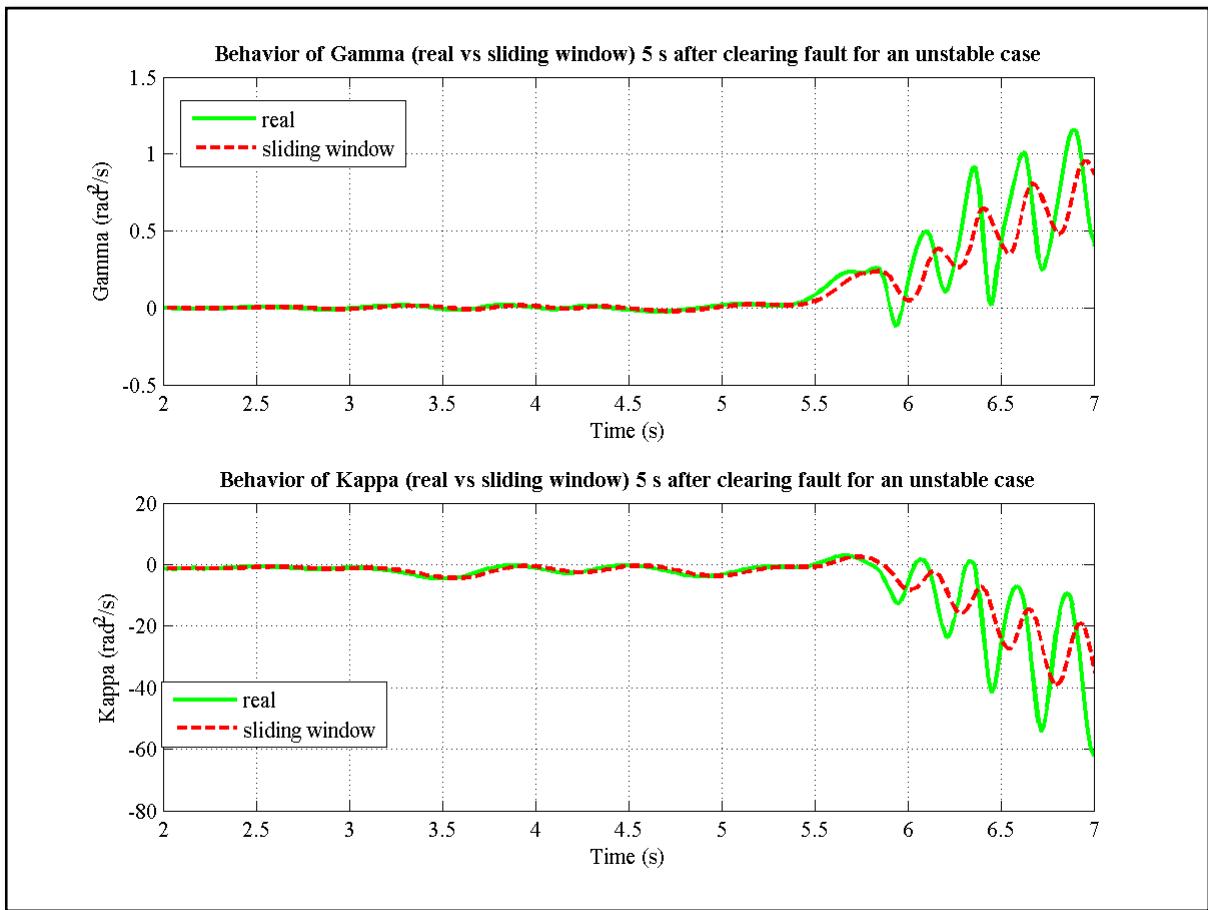


Figure 7.7 Comportement de γ_{COP} et κ_{COP} 5 s après l'effacement du défaut
Réseau Simplifié Australien

On reproduit à la Figure 7.8 une vue globale sur les différents indicateurs proposés pour prédire l'instabilité. Le graphe c) montre l'évolution de l'angle du rotor des 14 machines sur le réseau. On remarque que l'angle de certaines machines se séparent des autres, signe que le réseau va tomber en instabilité si aucune action corrective n'est apportée. Le graphe a) dénote que les COP de vitesse des régions suivent la même tendance que les angles du rotor. Les régions 1 et 2 se détachent totalement des trois autres pour rompre leur cohérence qui est une caractéristique de stabilité du réseau électrique. La notion de COP de vitesse constitue donc une alternative viable aux angles du rotor des générateurs pour la détection de l'instabilité. Autre que la déviation des COP de vitesse des régions, la divergence de γ_{COP} et de κ_{COP} sur les graphes b) et d) confirment l'instabilité du réseau.

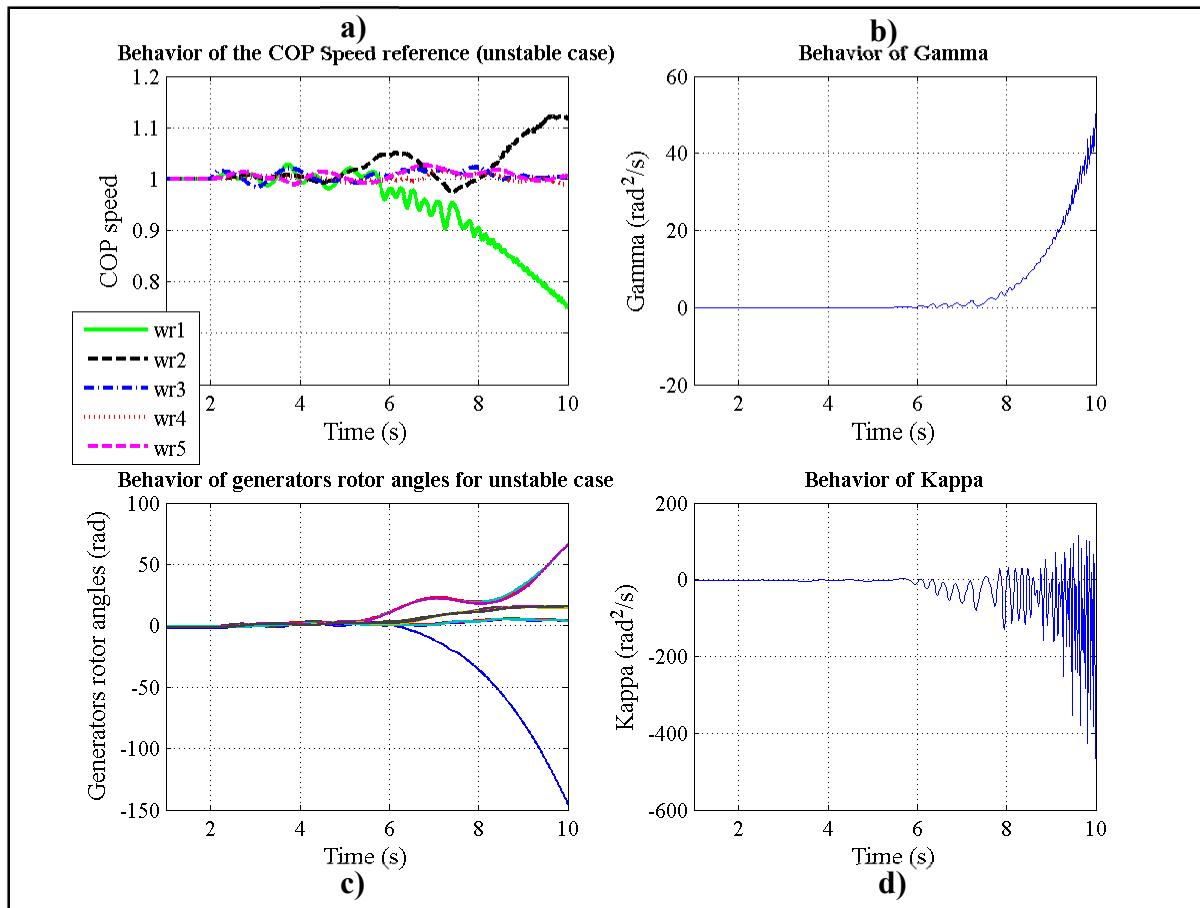


Figure 7.8 Vue globale sur les indicateurs de détection de l'instabilité
Réseau Simplifié Australien

7.4 Analyse de l'importance et de l'efficacité des prédicteurs

La précision de la classification dépend fortement de la qualité des attributs choisis comme prédicteurs. L'avantage principal du RF est qu'il couvre l'impact de chaque variable prédictive prise individuellement ainsi que les interactions multivariées avec les autres variables prédictives. On rappelle que l'importance d'une variable dans la classification correspond à la diminution moyenne de l'impureté qu'elle permet. Pour chaque arbre, la diminution totale de l'impureté liée à une variable correspond à la diminution de l'impureté cumulée sur l'ensemble des nœuds qu'elle régit. Cette diminution est ensuite moyennée sur l'ensemble des arbres.

En choisissant comme prédicteurs toutes les variables (Scénario 1), la Figure 7.9 affiche sur l'axe des ordonnées (y) les variables en ordre d'importance relative et sur l'axe des abscisses (x) leur importance. Par conséquent, les variables les plus importantes sont en haut et une estimation de leur importance est donnée par la position du point sur l'axe des x . Théoriquement, les deux critères MDA et MDG sont très proches et doivent être étudiés en parallèle car l'importance d'une variable dépend de sa fréquence d'apparition (variable globale, MDA) mais aussi des places qu'elle occupe dans chaque arbre (variable locale, MDG). Avec l'approche de la fenêtre glissante, on remarque pour les deux réseaux que la décision dépend largement des variables des régions notamment des COP de vitesse (i.e. $\omega_{R_2}, \omega_{R_4}, \omega_{R_T}, \dots$) et des variables dérivées directement des PMUs (i.e. $Vbus13, Vabc1619, \dots$). Les indices γ_{COP} et κ_{COP} apparaissent toujours au haut de la liste, en termes de diminution de la moyenne de Gini.

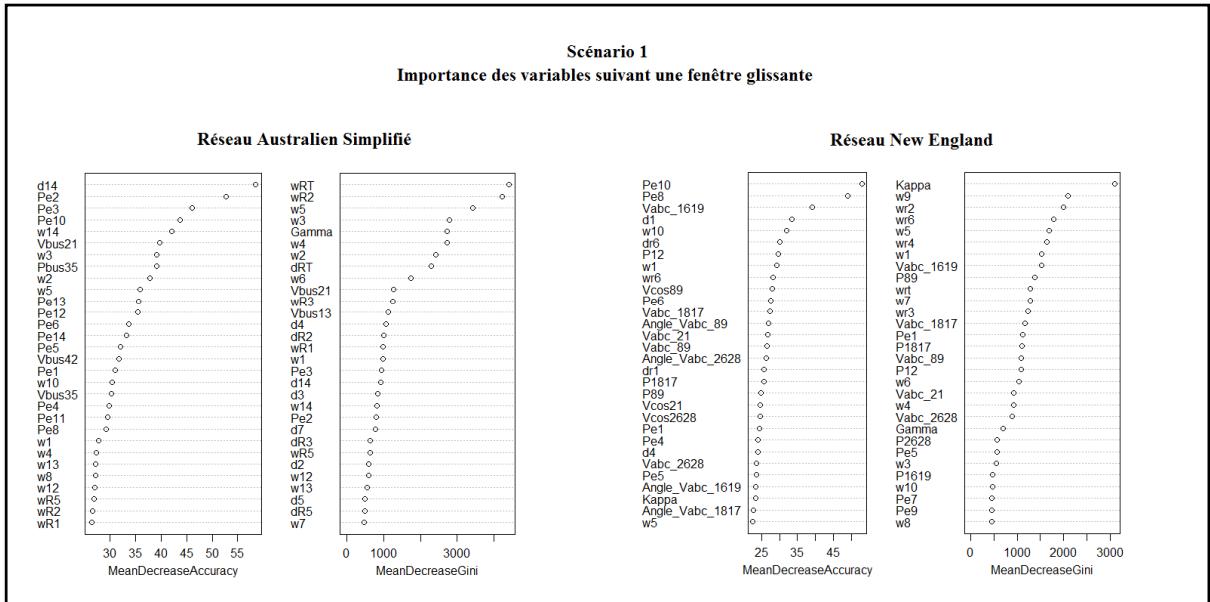


Figure 7.9 Importance des variables dans la construction de la forêt aléatoire pour le Scénario 1 – Réseau New England (droite) et Réseau Simplifié Australien (gauche)

Ces résultats prouvent que la réduction de la pureté du noeud est attribuable aux prédicteurs proposés lorsqu'on les utilise pour diviser les nœuds. À fortiori, les considérer uniquement comme des prédicteurs aux Scénarios 2 et 3 est une bonne stratégie pour minimiser la taille de la base de données dans le processus d'apprentissage. La Figure 7.10 confirme que les prédicteurs basés sur la notion de COP et les indices d'instabilité sont les variables les plus importantes dans la prise de décision dans le Scénario 2. Ils apparaissent toujours les premiers dans la liste de tous les prédicteurs.

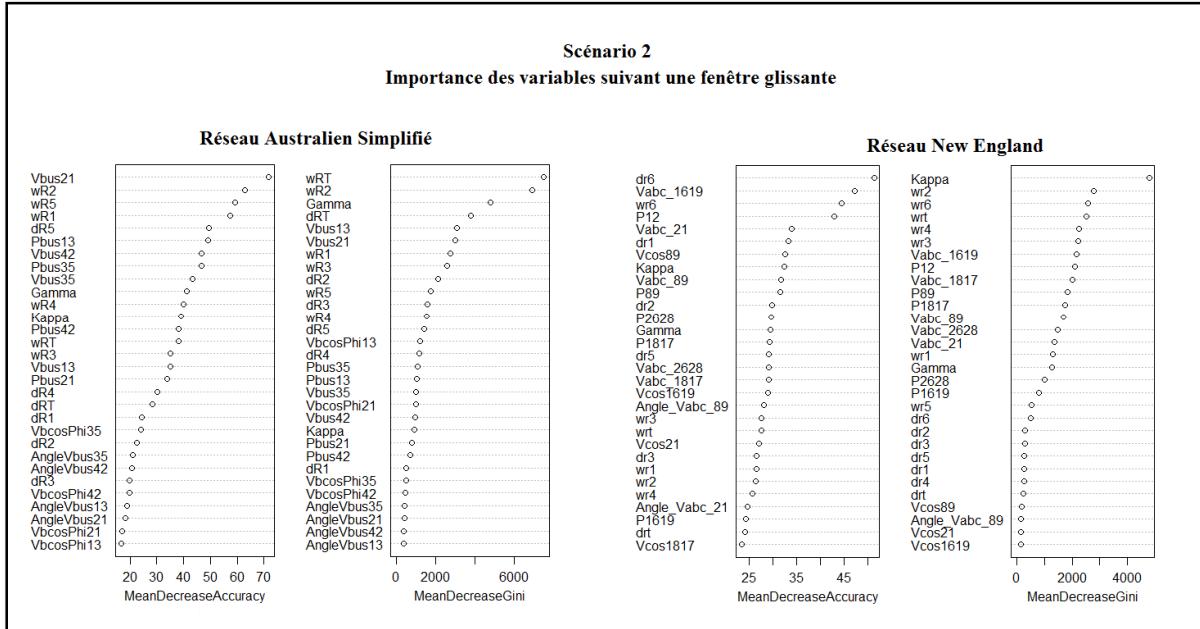


Figure 7.10 Importance des variables dans la construction de la forêt aléatoire pour le Scénario 2 – Réseau New England (droite) et Réseau Simplifié Australien (gauche)

Pour le Scénario 3, on fait encore le même constat tel que montré à la Figure 7.11. Les variables dérivant du COP ressortent toujours les premiers dans le classement.

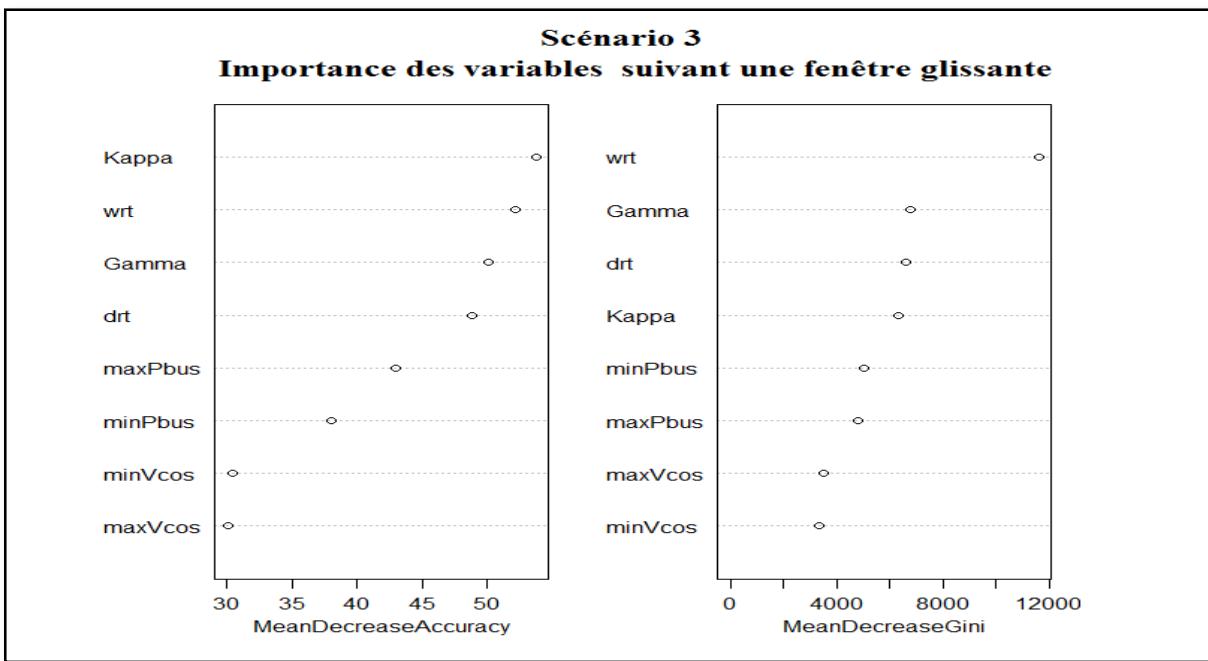


Figure 7.11 Importance des variables dans la construction de la forêt aléatoire pour le Scénario 3

7.5 Évaluation de la performance out-of-bag des forêts aléatoires

Idéalement, l'évaluation de la performance d'un algorithme de prédiction doit être effectuée à l'aide d'un ensemble de données indépendant des données utilisées dans la phase d'apprentissage. Dans la pratique, lorsque les données sont limitées, on procède à une forme de validation croisée (Hawkins, Basak et Mills, 2003) qui, malheureusement nécessite un long temps de calcul. Avec les échantillons OOB, le RF effectue l'équivalent d'une validation croisée parallèlement à l'étape d'apprentissage pour obtenir une estimation non biaisée de l'erreur de test. Sur la Figure 7.12, on évoque les changements dans le taux d'erreur associés aux prédictions du modèle RF pour les deux scénarios. L'axe des ordonnées (y) indique l'exactitude de la forêt des arbres (en termes de taux d'erreur) et l'axe des abscisses (x) représente le nombre d'arbres qui ont été inclus dans la forêt. On constate que les erreurs OOB des deux scénarios s'avoisinent pour les deux réseaux test. De plus, en tenant compte des deux scénarios pour le réseau Australien simplifié, l'erreur OOB des cas mal classés « Instable » est supérieure à ceux mal classés « Stable ». Pour le Réseau New England, on observe le contraire.

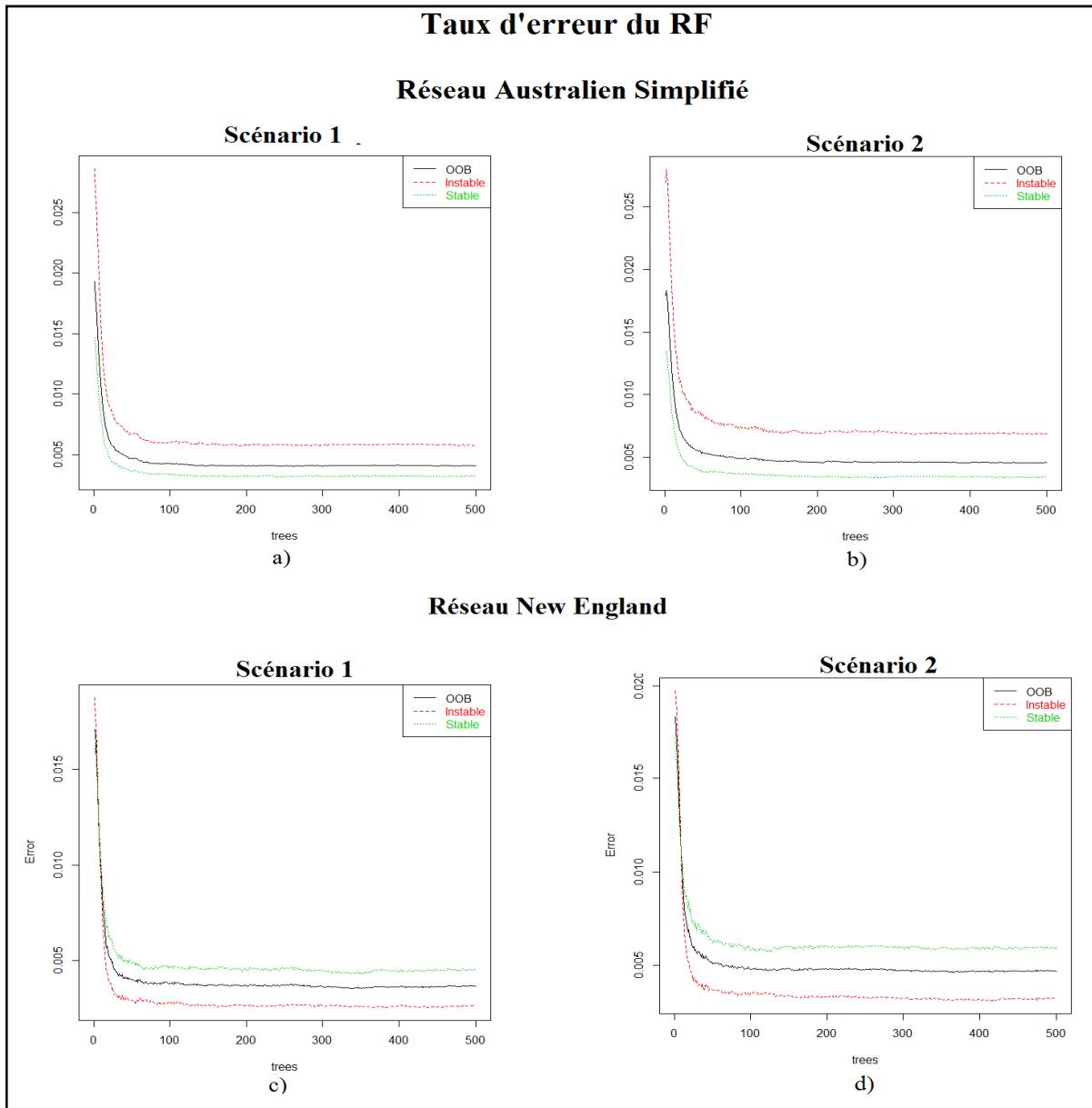


Figure 7.12 Caractéristiques de convergence de l'OOB pour l'apprentissage du RF avec la fenêtre glissante pour les Scénarios 1 et 2 – Réseau New England (bas) et Réseau Simplifié Australien (haut)

Pour le Scénario 3, comme illustré à la Figure 7.13, en combinant les deux bases de données, l'erreur OOB des cas mal classés « Stable » est supérieure à celle des cas mal classés « Instable ». Les Figures 7.12 et 7.13 montrent qu'aux alentours de 75 arbres, l'erreur OOB se stabilise. Elles suggèrent qu'il y a en fait très peu de changements dans la prédiction en

ajoutant des arbres supplémentaires à la forêt après en avoir atteint 75. Cette information est surtout utile pour des simulations futures dans un souci de gagner en temps de calcul et de simulation. Cela justifie en partie la raison pour laquelle le boosting produit plus loin est limité à 100 arbres.

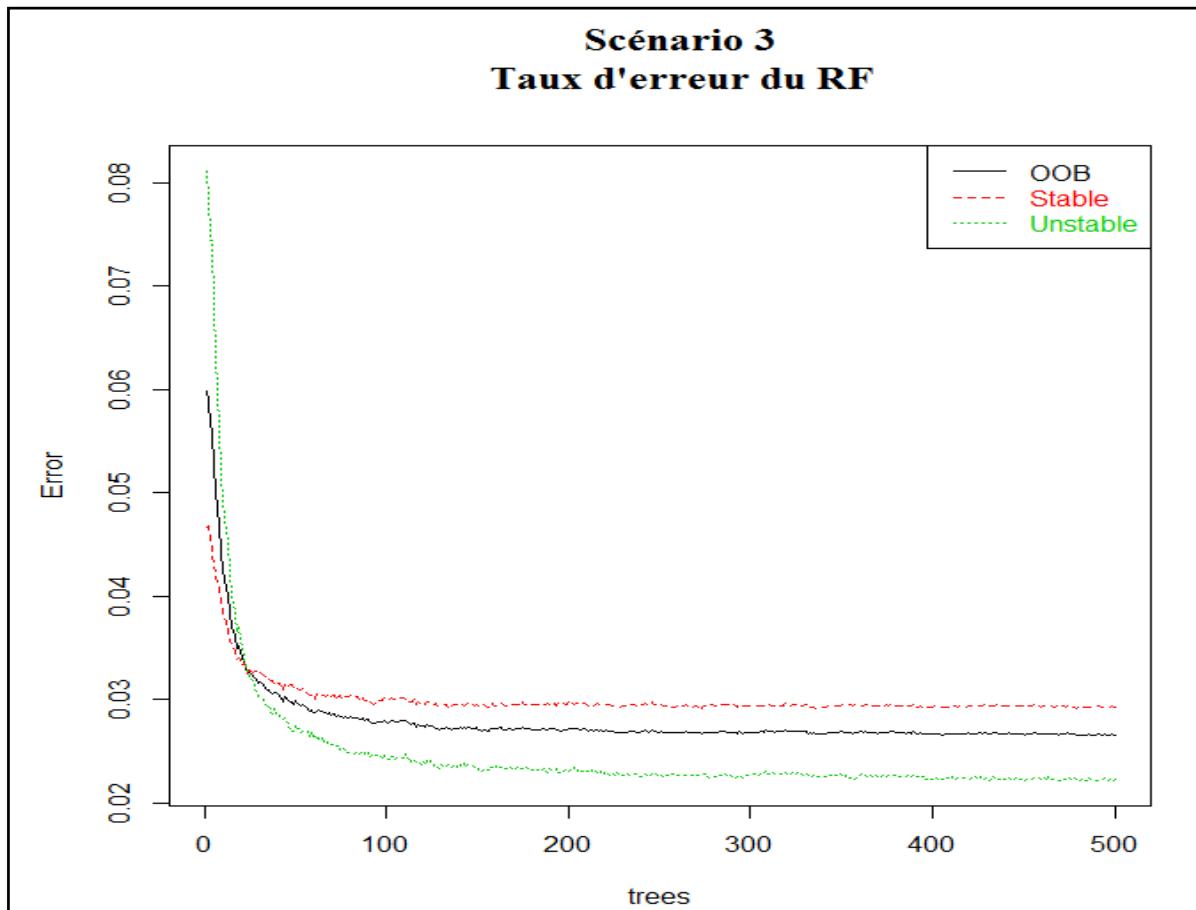


Figure 7.13 Caractéristiques de convergence de l'OOB pour l'apprentissage du RF avec la fenêtre glissante pour le Scénario 3

7.6 Évaluation de la performance du modèle de prédiction

Dans cette section, on se propose d'évaluer la performance du modèle de prédiction produit par le RF. En un premier temps, on compare les résultats obtenus en utilisant l'approche de la fenêtre glissante et celle des snapshots. Pour cela, on tient compte des Scénarios 1 et 2 pour les réseaux New England et Australien simplifié. En un deuxième temps, on compare les résultats

obtenus en utilisant les variables dérivées du COP et celles dérivées du COI comme prédicteurs dans la formation du RF. Finalement, on analyse les résultats du Scénario 3 en confrontant le COP au COI.

Le Tableau 7.1 réfère aux résultats de performance du RF pour la fenêtre glissante et les snapshots appliqués au réseau Australien simplifié. On peut faire deux constats principaux : 1) l'approche de la fenêtre glissante est celle avec de meilleurs résultats et 2) dans le cas du Scénario 2, on perd un peu dans la qualité du modèle de prédiction. En effet, en comparant le Scénario 1 au Scénario 2, dans tous les cas, la différence est de moins de 1% en termes d'exactitude, de sécurité et de fiabilité. Cependant, on observe une augmentation de l'ordre de 0.05% environ en termes de mauvaise détection et de fausse alarme. Sachant que la base de données est plus dense dans le Scénario 1, on juge que cette légère perte dans la qualité du modèle de prédiction constitue un bon compromis par rapport au temps de calcul et de simulation pour l'adoption du Scénario 2. Les prédicteurs proposés permettent donc de prédire correctement l'état d'instabilité du réseau en temps réel suite à une contingence sans tenir compte de toutes les mesures internes des générateurs dans la phase d'apprentissage.

Tableau 7.1 Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau Australien simplifié

	Snapshot 150 ms		Snapshot 300 ms		Fenêtre glissante	
	Scénario 1	Scénario 2	Scénario 1	Scénario 2	Scénario 1	Scénario 2
α_s	19384	19383	18752	18571	27395	27304
α_u	912	910	1183	1180	14096	14154
β_{ds}	147	149	142	145	80	94
β_{du}	80	81	91	92	89	113
Accuracy (%)	98.89	98.88	98.84	98.81	99.59	99.5
Security (%)	99.446	99.419	99.374	99.341	99.675	99.586
Reliability (%)	90.644	90.918	92.364	91.801	99.432	99.336
Misdetection (%)	0.716	0.73	0.7	0.73	0.19	0.23
False alarm (%)	0.39	0.39	0.45	0.46	0.21	0.27

Sur la Figure 7.14, on fait ressortir une comparaison plus nette entre les 3 approches pour le Scénario 2. Il apparaît clairement que la fenêtre glissante proposée dans le cadre de cette thèse jouit d'une meilleure exactitude et sécurité dans la prédiction du modèle. En termes de fiabilité, l'apport du SW est encore plus remarquable, on a une amélioration de plus de 5%. On remarque également qu'on diminue la mauvaise détection et la fausse alarme de près de la moitié.

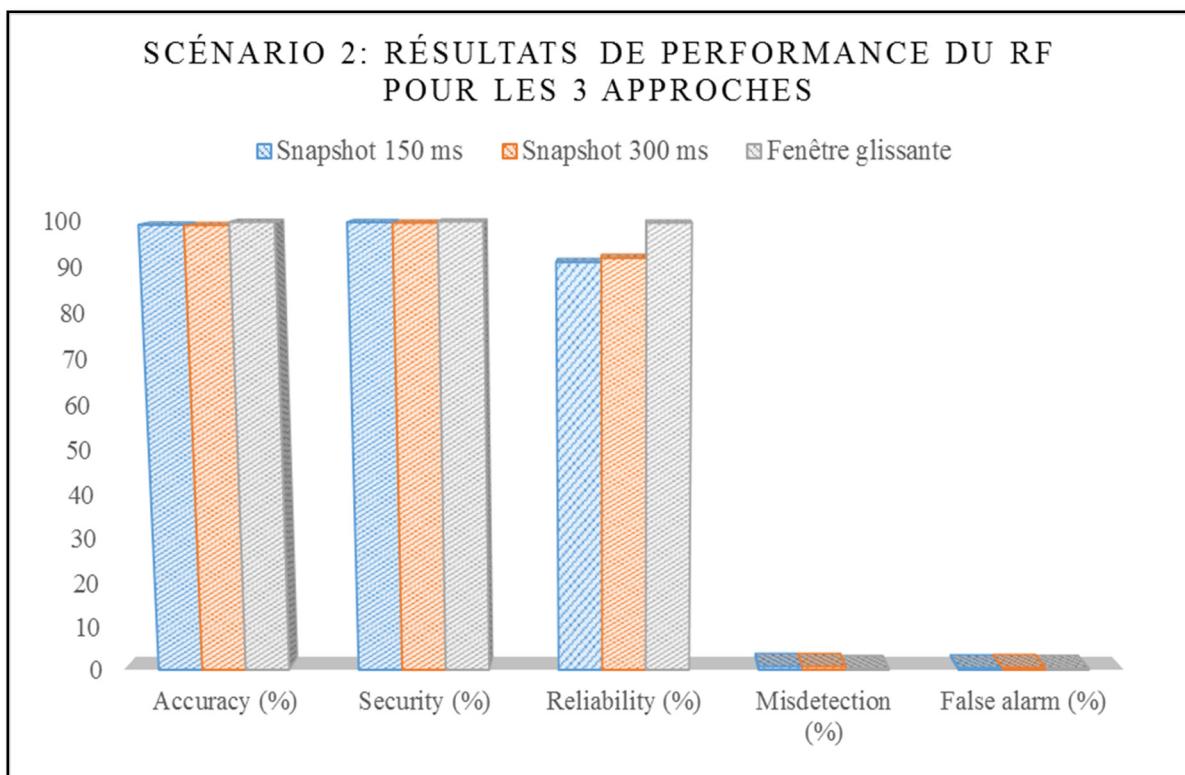


Figure 7.14 Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau Australien simplifié

Le Tableau 7.2 compare la performance du RF pour les 3 approches en considérant les Scénarios 1 et 2 appliqués au réseau New England. À comparer le Scénario 1 au Scénario 2 pour les snapshots, on constate une perte dans l'exactitude, la sécurité et la fiabilité du modèle de l'ordre du dixième de pourcentage. Par contre pour le SW, la différence est de l'ordre du centième de pourcentage quand la valeur n'est pas la même. Pour les snapshots, on remarque aussi une augmentation de la mauvaise détection et de la fausse alarme en passant du Scénario 1 au Scénario. Là encore, il s'agit de variation de l'ordre du dixième de pourcentage. Pour le

SW, la mauvaise détection est inchangée alors que la fausse alarme a augmenté de 0.04%. Fort de cela, on peut affirmer qu'en maintenant uniquement les prédicteurs du Scénario 2 dans la phase d'apprentissage, on obtient un modèle de prédiction presqu'équivalent en utilisant toutes les variables.

Tableau 7.2 Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau New England

	Snapshot 150 ms		Snapshot 300 ms		Fenêtre glissante	
	Scénario 1	Scénario 2	Scénario 1	Scénario 2	Scénario 1	Scénario 2
α_s	3024	3020	3199	3187	14980	14969
α_u	235	228	420	418	12332	12332
β_{ds}	25	32	37	39	33	33
β_{du}	40	44	39	51	62	73
α	3259	3248	3619	3605	27312	27301
β	65	76	76	90	95	106
Accuracy	98.045%	97.714%	97.943%	97.564%	99.653%	99.613%
Security	98.677%	98.543%	98.781%	98.400%	99.586%	99.512%
Reliability	89.362%	85.965%	91.190%	90.670%	99.732%	99.732%
Misdection	0.752%	0.963%	1.001%	1.055%	0.120%	0.120%
False alarm	1.203%	1.324%	1.055%	1.380%	0.226%	0.266%

Sur les deux réseaux, avec de telle performance du RF avec les prédicteurs du Scénario 2, on conclut que les prédicteurs basés sur le concept de COP, les mesures des PMUs et les indices proposés sont très performants et permettent de réduire considérablement la base de données servant à construire le modèle de prédiction. Pour analyser l'apport du SW, on recourt à la Figure 7.15 qui met en parallèle la performance des snapshots à la sienne. On remarque encore une fois que globalement la fenêtre glissante est l'approche qui produit le modèle de prédiction avec les meilleurs résultats. On assiste à une augmentation de l'exactitude, de la sécurité et de la fiabilité du modèle dans les mêmes proportions que précédemment (Figure 7.14). Par contre, en termes de mauvaise détection et de fausse alarme, on peut voir de nettes diminutions et on peut noter qu'elles sont plus significatives que celles rapportées pour le réseau Australien simplifié.

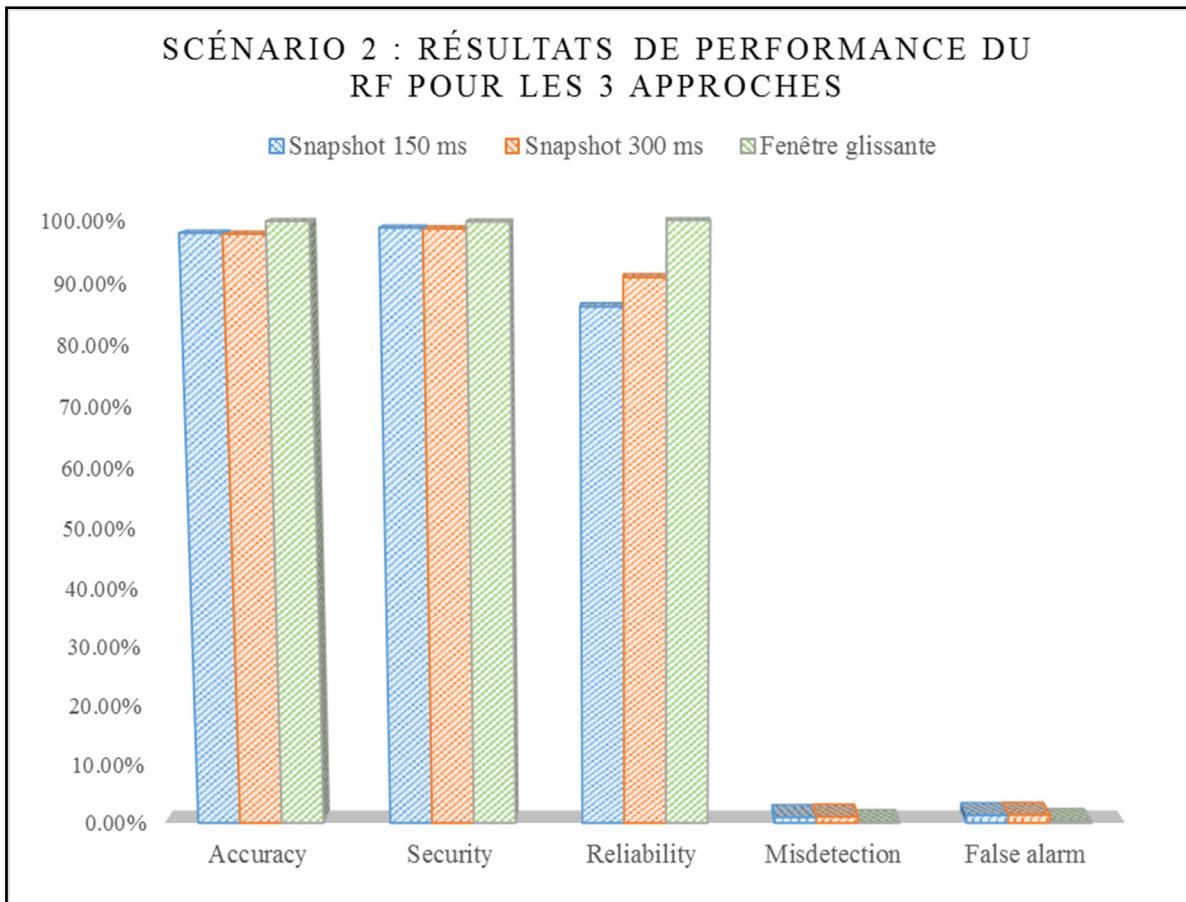


Figure 7.15 Comparaison des résultats de performance des modèles issus de RF pour les 3 approches – Réseau New England

On se propose à présent de comparer la performance de modèles de prédiction formés à partir de prédicteurs provenant de variables COP et COI. Pour ce faire, on priorise l'approche de la fenêtre glissante appliquée aux réseaux New England et Australien simplifié. De surcroît, on s'assure de prendre en compte les mêmes attributs. On définit les indices COI ainsi :

$$\gamma_{COI} = \sum_{l=1}^n \delta_{R_l}^{COI} (\omega_{R_l}^{COI} - \omega_{R_T}^{COI}) \quad (7.1)$$

$$\kappa_{COI} = \sum_{l=1}^n \omega_{R_l}^{COI} (\delta_{R_l}^{COI} - \delta_{R_T}^{COI}) \quad (7.2)$$

L'objectif ici n'est pas de mettre en compétition le COI et le COP afin de statuer définitivement sur lequel des deux convient le mieux pour la prédiction de l'instabilité. On s'intéresse tout simplement à faire une comparaison entre les résultats pour voir si les indices proposés fonctionnent aussi bien en exploitant le concept de COI. Toutefois, on recommande à ce que des études supplémentaires soient faites pour déduire lequel des deux concepts présente le plus d'avantages dans l'étude de l'instabilité. Le Tableau 7.3 résume les résultats de performance des modèles issus des RF. Suivant ces derniers, avec le COP on améliore légèrement l'exactitude, la sécurité et la fiabilité du modèle pour les deux réseaux. Le gros avantage se retrouve au niveau de la mauvaise détection que l'on diminue près de la moitié pour le réseau New England et de 33% environ pour le réseau Australien simplifié. La réduction de la fausse alarme avoisine les 40% pour le premier réseau alors qu'elle décroît aux alentours de 37% pour le second. Comme mentionné récemment, ces résultats ne suffisent pas pour affirmer péremptoirement que le COP est supérieur au COI.

Tableau 7.3 Comparaison des résultats de performance des modèles issus de RF
sur une fenêtre glissante avec des variables COP et COI comme prédicteurs
Réseaux Australien simplifié et New-England

	New England		Australian 14 generators	
	COI	COP	COI	COP
α_s	14931	14969	27276	27304
α_u	12245	12332	14070	14154
β_{ds}	111	33	141	94
β_{du}	120	73	178	113
α	27176	27301	41346	41458
β	231	106	319	207
Accuracy	99.157%	99.613%	99.234%	99.503%
Security	99.196%	99.512%	99.347%	99.586%
Reliability	99.094%	99.732%	98.998%	99.336%
Misdetection	0.405%	0.120%	0.338%	0.226%
False alarm	0.438%	0.266%	0.427%	0.271%

Finalement, on conçoit la base de données du Scénario 3 en concaténant la moitié de chacune des deux bases de données obtenues pour les deux réseaux en appliquant une fenêtre glissante. D'un côté, on prend comme prédicteurs les variables COP et celles qui en découlent ; et de l'autre côté, les attributs COI. On constate sur la Figure 7.16 demeure très performant avec un fort pourcentage exactitude et ne fiabilité. Les fausses alertes restent très faibles. La cohérence des résultats démontre donc la généralité de la méthodologie proposée et son indépendance par rapport réseau considéré. La Figure 7.16 montre aussi que le modèle de prédiction formé avec le COP jouit d'une plus grande exactitude, d'une plus grande sécurité et fiabilité. De plus, on décèle que le pourcentage de fausse alarme et de mauvaise détection sont moindres pour le COP.

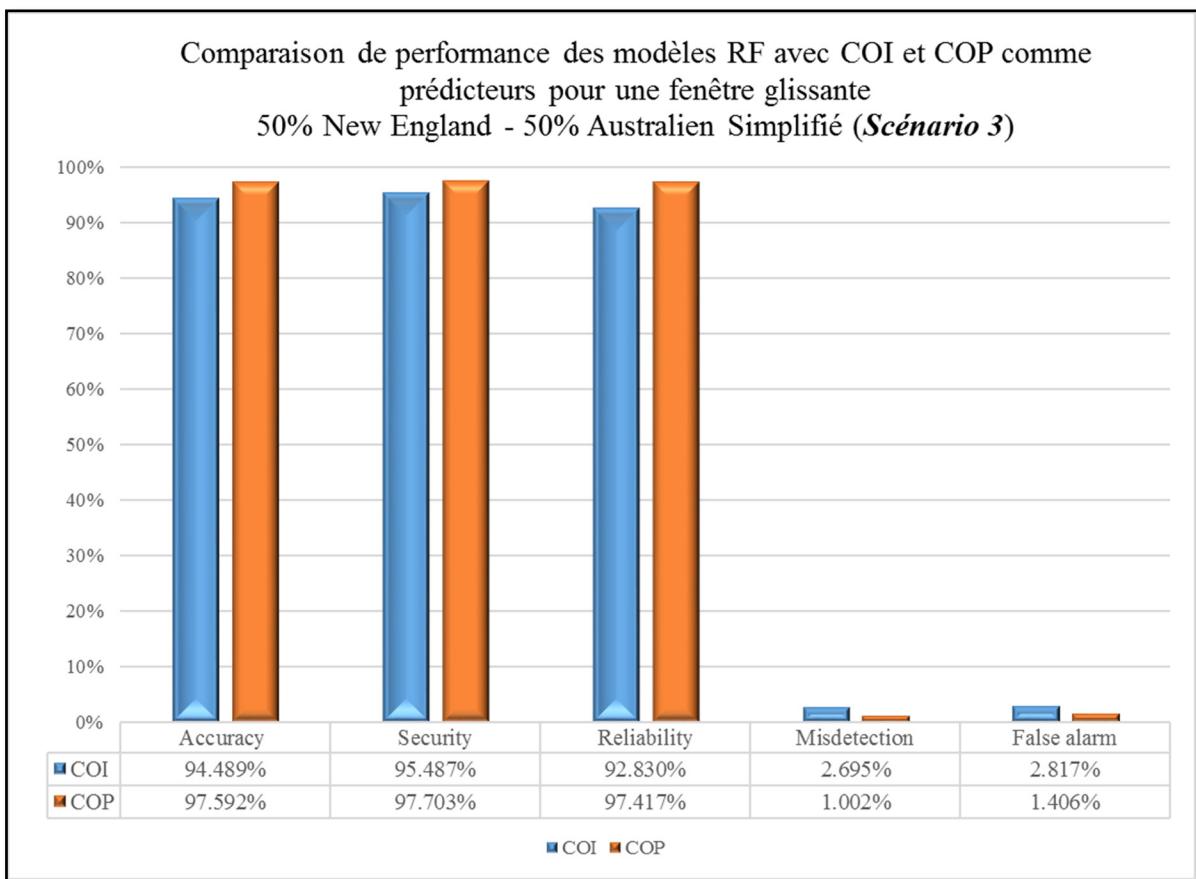


Figure 7.16 Comparaison des résultats de performance des modèles issus de RF pour le Scénario 3

7.7 Mise à jour automatique du modèle de prédiction

Afin d'illustrer la procédure de mise à jour proposée, on fait appel à 5 nouveaux jeux de données que l'on obtient en provoquant des défauts sur les deux réseaux : New England et Australien simplifié. Toutes les étapes menant à la création de ces nouveaux jeux de données sont identiques à celles suivies dans la création de la base de données initiale.

Avec la méthode proposée, au lieu de reconstruire un modèle complet de RF basé sur la base de données initiale et l'entrée de nouvelles données de manière périodique, on préfère construire un boosting avec ce nouveau jeu de données moins dense. Le RF initial et le boosting ainsi construit fonctionnent en cascade. On arrive donc à faire une double vérification sur l'état en temps réel du réseau avec deux modèles de prédiction indépendants : le premier tenant compte d'une base de données plus dense et le second prenant en compte l'apparition d'éventuelles nouvelles signatures dans les données. Dans ce cas, on s'assure d'intégrer dans la prédiction les connaissances passées sur le réseau et les connaissances actuelles. La Figure 7.17 présente les résultats de performance des modèles de prédiction dans les deux cas.

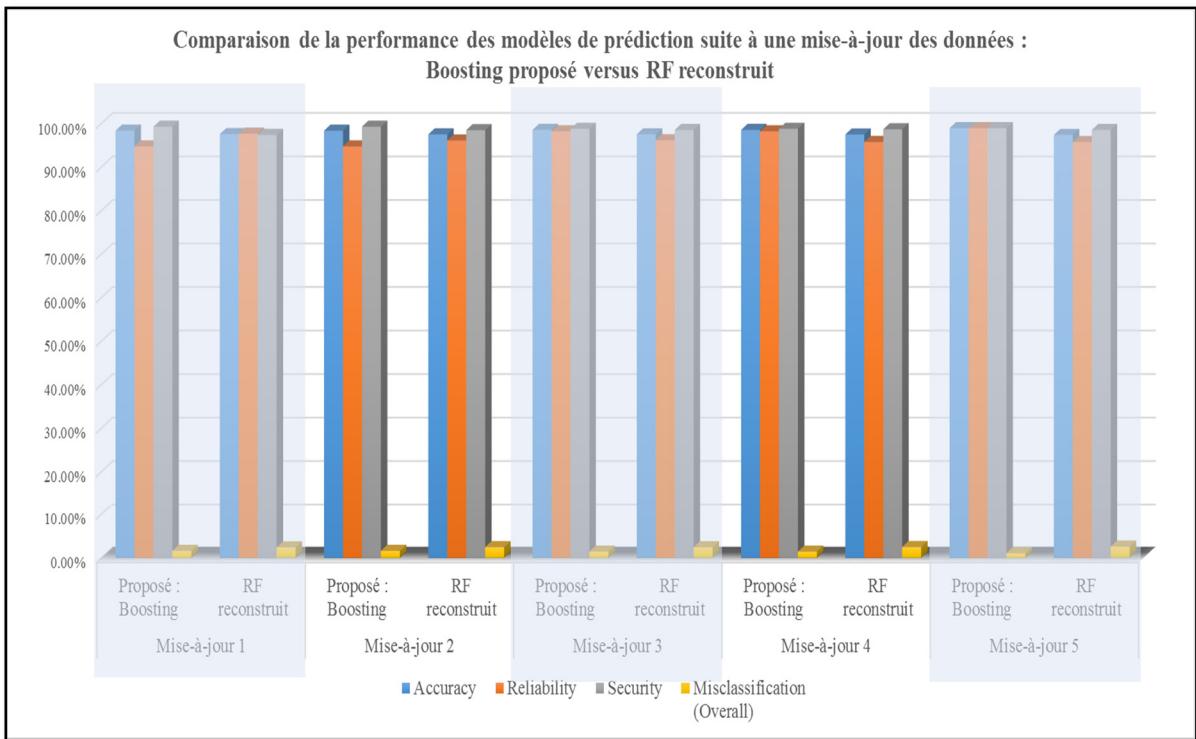


Figure 7.17 Comparaison des résultats de performance des modèles issus du Boosting proposé et du RF reconstruit suite à une mise à jour de la base de données – Scénario 3

Dans ce cas, la mauvaise classification (*Misclassification overall*) est la somme des pourcentages de mauvaise détection et de fausse alarme. On fait remarquer que pour les 5 mises à jour, la probabilité de mal classer les événements est plus élevée en reconstruisant un RF complet avec une base de données étendue. Parmi tous les paramètres de mesure de performance du modèle, c'est dans ce cas que l'écart est le plus élevé entre les deux modèles. Globalement, la sécurité et l'exactitude des deux modèles sont les mêmes à chaque mise à jour. En revanche, au niveau de la fiabilité, on note une certaine alternance dans la tendance. Par exemple, pour la première mise à jour, le RF reconstruit présente une meilleure fiabilité alors que l'on relève l'inverse lors de la troisième.

Un autre avantage que l'on retrouve dans le boosting réalisé pour la mise à jour est le temps de calcul du modèle qui est nettement inférieur à celui du RF reconstruit. Ce résultat n'est pas étonnant vu que la quantité de données traitées est moindre pour le boosting. La Figure 7.18 montre les résultats y relatifs.

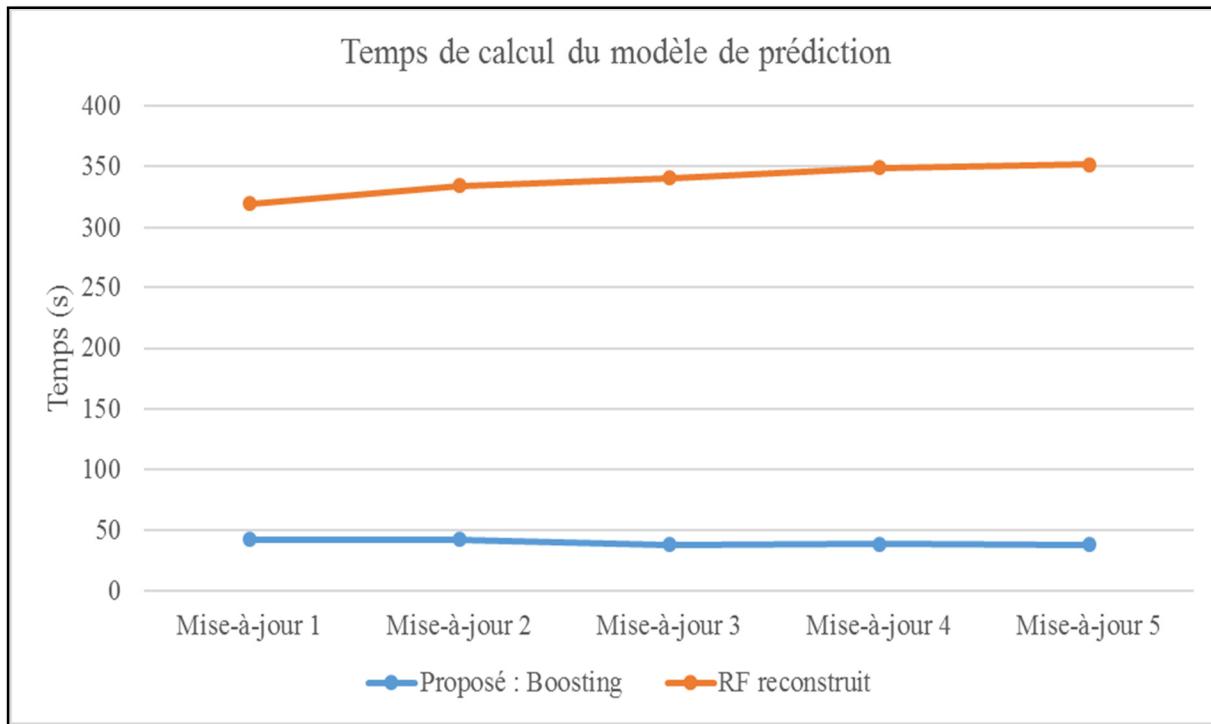


Figure 7.18 Temps de calcul du modèle de prédiction lors de la mise à jour – Boosting proposé vs RF reconstruit

Le Tableau 7.4 montre l'évolution de la base de données retenue avec les mises à jour périodiques. On rappelle que suivant la méthodologie proposée dans la Figure 6.7, quand la différence $\Delta\Omega$ est positive on garde la base de données courante de référence. Par exemple, en passant de S_1 à S_2 , comme $\Delta\Omega$ est positif, on garde le modèle de prédiction fourni par S_1 jusqu'à une prochaine mise à jour. Pour le modèle obtenu avec S_3 , on voit que $\Delta\Omega$ est négatif, alors ce nouveau modèle est conservé pour les prédictions futures.

Tableau 7.4 Suivi de la mise à jour de la base de données en fonction de l'indicateur Ω

Mise à jour	Test	Ω	$\Delta\Omega$	Données sauvegardées
$S1$		0.0318393	$\times \times$	$\times \times$
$S2$		0.0342118	-0.0023725	$S1$
$S3$		0.0267303	0.005109	$S3$
$S4$		0.0917415	-0.0650112	$S3$
$S5$		0.0235954	0.0031349	$S5$

7.8 Discussion des résultats

L'exécution des modèles par le logiciel **R** utilisé est réalisée à l'aide d'un ordinateur portable de mémoire 4.00Gb et doté d'un processeur Intel® Core™ i5-4300U CPU @1.90GHz 2.50GH. Le Tableau 7.5 résume le temps d'exécution pour l'apprentissage du RF initial pour les Scénarios 1 et 2 et pour chacun des réseaux. Ce dernier relate que le temps d'exécution de l'apprentissage du modèle est réduit de plus de la moitié pour les deux réseaux en considérant uniquement comme prédicteurs les variables dérivées de la notion de COP proposée. Le temps économisé en exécutant l'apprentissage dans le Scénario 2 constitue un bon compromis comparé à la précision de l'ordre du centième perdu par rapport au Scénario 1.

Tableau 7.5 Temps d'exécution en minutes pour l'apprentissage du modèle

	Scénario 1	Scénario 2
Réseau Australien simplifié	58.75	20.62
Réseau New England	14.32	5.99

En généralisant dans le Scénario 3, l'apprentissage du modèle RF initial a pris 3.53 minutes avec les variables dérivées du COP et 5.10 minutes avec celles dérivées du COI.

Pour les réseaux pris en compte dans le cadre cette étude, le début de l'instabilité est généralement détecté à partir de 0.05 s après l'effacement du défaut. Ce temps correspond au

début de la divergence des indices γ et κ considérés. Dans le cas d'une instabilité rapide, les opérateurs sont donc informés dans un temps raisonnable pour ne pas perdre le synchronisme du réseau (temps de préemption de 10 cycles généralement requis). Le pire temps de détection est de 0.833 s après le défaut. Il s'agit là des cas d'instabilité lente, comme traité à la Figure 7.4. Puisque le réseau tombe en instabilité un long temps après l'effacement du défaut (plus de 5 s dans le cas susmentionné), le temps de préemption est suffisant pour apporter les actions correctives nécessaires.

Les deux algorithmes privilégiés (RF et boosting) pour la prédiction de l'instabilité transitoire sont indéniablement plus efficaces qu'un arbre de décision pris seul. En effet, dans les algorithmes les plus connus de construction d'arbres de décision, tous les descripteurs sont testés pour leur performance de fractionnement à chaque nœud. Pour un RF, on ne teste qu'un nombre limité de descripteurs choisis aléatoirement. Le boosting quant à lui améliore les performances des classificateurs faibles en pondérant l'importance des observations mal-estimées lors de l'itération précédente. De plus, il a été prouvé empiriquement que le RF et le boosting surclassent le DT et le bagging (Meyer, Leisch et Hornik, 2003).

La base de données initiale étant plus dense, on utilise le RF pour établir le modèle de prédiction de base. Cela s'explique par le fait que le RF n'a fondamentalement qu'un seul paramètre à régler pour améliorer la qualité du modèle de prédiction : le nombre d'attributs à sélectionner au hasard à chaque nœud, noté n_{try} dans cette thèse. Pour le boosting, il en existe plusieurs dont le nombre d'arbres et la profondeur (ou nombre de feuilles). De plus les RF sont plus robustes au surapprentissage et nécessitent moins de consensus dans son algorithme pour l'éviter. Il va de soi qu'il est préférable d'exploiter les RF pour les grandes bases de données et de se servir du boosting lors des mises à jour fréquentes.

De nos jours, dans les travaux, le concept de COI est très utilisé (Kamwa, Samantaray et Joos, 2009) – (Cepeda et al., 2014). En pratique, dans les centres d'EMS/SCADA, il n'est pas facile d'obtenir une valeur précise du moment d'inertie des générateurs par les opérateurs du réseau électrique suite à une redistribution de la puissance par exemple. En définissant des prédicteurs

d'instabilité indépendants de ceux-ci comme le COP de vitesse des régions et les indices basés sur des valeurs globales, il est possible de s'affranchir de ces données manière concrète et pratique. En se concentrant sur les valeurs associées aux zones, la base de données à analyser est réduite de manière significative. Le Scénario 1 confirme l'efficacité des variables proposées comme prédicteurs pour évaluer l'instabilité du réseau. En effet, ces dernières apparaissent en premier dans la liste d'importance relative des variables selon le MDG et le MDA.

En outre, les prédicteurs proposés sont plus fiables que ceux généralement utilisés pour la construction de DT (Kamwa, Samantaray et Joos, 2009) – (Kai et al., 2007). Pour les trois scénarios, la perte de précision, de sécurité et de fiabilité sont très faibles et l'augmentation de la probabilité de mauvaise détection et de fausse alarme sont négligeables. De plus, ce travail montre que l'approche proposée basée sur l'approche de la fenêtre glissante est très efficace par rapport aux approches snapshots très répandues dans la littérature. En effet, le SW peut augmenter significativement la fiabilité tout en réduisant de près de la moitié la fausse alarme et la mauvaise détection. Finalement, notez également que les indices proposés permettent de distinguer sans équivoque un cas stable d'un cas instable dans le domaine temporel.

CONCLUSION

Les réseaux électriques modernes sont en expansion permanente, les interconnexions se multiplient et des sources d'énergie renouvelables diverses s'y rajoutent. Pour satisfaire à la demande croissante d'énergie électrique, les réseaux sont exploités à leurs limites de sécurité. Cela va sans dire que tous ces facteurs ajoutent un degré de difficulté dans le maintien de la stabilité et de la sécurité des réseaux. Vu l'importance de la conservation de la continuité du service, suite à des contingences, il est souhaitable de pouvoir prédire en tout temps le comportement des réseaux en fonction des différents paramètres et variables qui les définissent.

Avec l'émergence des technologies de synchrophaseurs, on assiste à une explosion quantitative des données numériques fournissant des informations en temps réel sur l'état du réseau. Les chercheurs doivent trouver de nouvelles manières d'analyser ces flux de données en découvrant les signatures cachées qui régissent ces données multidimensionnelles. La présente thèse s'inscrit dans un tel contexte. Elle se veut de proposer une nouvelle approche de détection en temps réel de l'instabilité transitoire et dynamique des réseaux électriques grâce aux mesures captées par les PMUs.

En effet, on a introduit la nouvelle notion de centre de puissance pour la prédiction de l'instabilité. Grâce aux mesures des PMUs, on peut aisément estimer en temps réel la vitesse et l'angle du rotor des générateurs du réseau. Au lieu d'analyser individuellement les vitesses de tous les générateurs, on a considéré une représentation équivalente de chacune des régions cohérentes du réseau : le centre de puissance de vitesse des régions. Ce nouveau concept a le principal avantage d'être indépendant de l'inertie des générateurs. On s'intéresse aux puissances déduites des mesures de synchrophaseurs captées sur le réseau. Ces dernières pondèrent les vitesses et tiennent compte de l'apport réel et actuel de chacun des générateurs dans le fonctionnement du réseau. Dans le cas où les COP de vitesse de deux régions quelconques du réseau continuent à s'éloigner, le réseau est considéré instable. Toutefois, si

tous les COP de vitesse évoluent de manière cohérente, le synchronisme du réseau est réputé maintenu.

En faisant appel à la notion de COP et au produit scalaire modifié, lequel permet de trouver le point menant à l'instabilité avec les méthodes de fonction d'énergie, on a défini deux nouveaux indices de détection de l'instabilité sur le réseau électrique : γ_{COP} et κ_{COP} . L'instabilité est caractérisée par la divergence de ces deux variables dans le temps.

Généralement, dans l'étude de la stabilité, on utilise des snapshots. Dans cette thèse, on a proposé une technique de fenêtre glissante. Elle a comme forces principales de mettre à jour la base de données sur l'état du réseau avec les nouvelles données entrantes sans perdre les informations contenues dans les données précédentes. De plus, le SW permet de décider en tout temps contrairement au snapshot qui privilégie une photo du réseau à un instant donné.

À partir des mesures PMUs et des variables estimées en fonction d'elles, on a créé une riche base de données en se référant aux réseaux test : Kundur, New England et Australien simplifié. Pour extraire des informations cachées et invisibles à l'œil nu de cette base de données et pour pouvoir prédire l'état du réseau (instable ou stable) par la donnée de certains paramètres, on a eu recourt à des techniques de forage de données. Dans la classification supervisée, les arbres de décision sont reconnus pour la simplicité de leur modèle, leur lecture facile et la facilité de les expliquer. Mais pris seuls, ils sont sensibles au bruit et ont une forte tendance à surapprendre les données, i.e. à apprendre à la fois les relations entre les données et le bruit présent dans l'ensemble d'apprentissage. Pour pallier à cette difficulté, la forêt aléatoire adoptée dans cette thèse a permis de réaliser l'apprentissage sur de multiples arbres de décision non-élagués entraînés sur des sous-ensembles de données légèrement différents. Pour évaluer la pertinence de la notion de COP introduite et de l'importance des indices de détection de l'instabilité proposés, on a considéré trois scénarios qui diffèrent principalement par les attributs considérés comme prédicteurs dans la phase d'apprentissage du modèle de prédiction pour chacun des réseaux. Dans le Scénario 1, on intègre toutes les variables mesurées et estimées comme prédicteurs dans la phase d'apprentissage. On a remarqué que celles dépendant directement du COP suggéré, des indices de détection de l'instabilité proposés et des PMUs ressortent toujours

parmi les variables les plus importantes. Il est donc raisonnable d'envisager au Scénario 2 de prendre uniquement comme prédicteurs le dernier groupe d'attributs cités. Le modèle de prédiction obtenu au Scénario 2 est aussi performant que celui obtenu au Scénario 1. Les écarts dans l'exactitude, la fiabilité, la sécurité et la mauvaise classification des deux modèles RF sont très minimes. Considérant que la base de données a été réduite et que les temps de calcul des modèles ont considérablement diminué, on a jugé que cette faible perte de précision du modèle de prédiction constitue un très bon compromis. La comparaison des snapshots à la fenêtre glissante a prouvé que cette approche jouit d'une meilleure efficacité dans la prédiction de l'instabilité en temps réel.

En généralisant au Scénario 3, on a prouvé que les prédicteurs proposés sont indépendants du réseau. En comparant l'approche du centre d'inertie au COP, il est ressorti des résultats que les prédicteurs dérivés du COP sont plus performants dans la construction du modèle de prédiction avec des probabilités d'exactitude, de fiabilité et de sécurité élevées et un faible taux de mauvaise détection et de fausse alarme.

En dernier lieu, on a présenté une technique de mise à jour du modèle de prédiction en temps réel avec l'arrivée de nouveaux jeux de données. Essentiellement, on a mis en cascade le RF initialement construit et un boosting. On s'est référé à ce boosting pour savoir si de nouvelles signatures cachées s'ajoutent à la base de données. À chaque modèle créé, on a attribué un indicateur le caractérisant. Pour savoir si le modèle doit être mis à jour, on consulte la différence de ces deux indicateurs : si elle est négative, le modèle est mis à jour, et conservé sinon. Les résultats obtenus ont montré l'efficacité d'une telle démarche.

Somme toute, dans cette thèse, on a développé une nouvelle approche pour la prédiction en temps réel de la stabilité transitoire et dynamique des réseaux électriques en exploitant les mesures des synchrophaseurs. Cette approche prend également en compte la mise à jour périodique des données aux centre de collecte et de traitement des données sur les réseaux électriques.

RECOMMANDATIONS

Dans cette partie, on formule les recommandations pour les travaux futurs :

1. On a présenté des résultats comparatifs en utilisant le centre d'inertie et le centre de puissance proposé. Toutefois, ils n'ont pas été mis en compétition. Il serait intéressant de poursuivre les études comparatives sur ces deux approches et de déterminer laquelle est la plus adaptée pour l'analyse et la prédition de l'instabilité. Les résultats obtenus laissent présager que l'utilisation du COP pourrait être plus pertinente, mais des travaux plus pointus doivent être faits en ce sens ;
2. Généralement, dans l'analyse de la stabilité transitoire, les générateurs sont divisés en deux groupes. Pour faire cette division, on utilise la notion de COI. On pourrait envisager de diviser les machines suivant le concept de COP présenté et voir laquelle des deux méthodes permet d'avoir un moindre temps critique de défaut et un angle critique d'instabilité plus adéquat ;
3. L'analyse élaborée dans le cadre de cette thèse concerne typiquement trois réseaux tests, il serait important d'étendre l'approche proposée à de plus grands réseaux dans le but d'agrandir la base de données et de valider qu'effectivement les prédicteurs proposés sont indépendants des réseaux. On pourrait aussi appliquer la méthodologie proposée à des réseaux existants ;
4. On pourrait aussi développer suivant la méthodologie proposée des outils et des composantes pour enrichir la librairie de Simscape Power Systems™ (SimPowerSystems™) dans Matlab.

ANNEXE I

CARACTÉRISTIQUES DES RÉSEAUX ÉLECTRIQUES TEST KUNDUR TWO-AREA ET NEW ENGLAND

Cette annexe présente les caractéristiques de deux réseaux électriques test utilisés lors des simulations. Il s'agit du réseau *Two-area system* (aussi connu sous le nom de *Kundur Two-area test System*) et du réseau *New England*. On mentionne le nombre de machines, le nombre de jeux de barres et la quantité de lignes. On décrit les modèles des machines, les modèles des charges et des éléments de transmission des réseaux électriques. On montre également les différentes régions cohérentes considérées dans notre étude. On rappelle qu'une base de 100 MVA est choisie pour toutes les conversions en per unit.

Réseau test *Two-area System*

Le réseau test *Two-area System*, aussi appelé *Kundur two-area test system*, est un réseau proposé dans cet ouvrage (Kundur, Balu et Lauby, 1994). Il est utilisé pour les études de stabilité transitoire, de stabilité dynamique, d'amortissement d'oscillations, de transit de puissances etc... Comme montré à la Figure-A- I.1, ce réseau est divisé en deux régions (*Area 1* et *Area 2*) ayant chacune deux générateurs, ce qui fait un total de 4 générateurs. Il renferme également 11 bus, deux charges et deux *shunt capacitors*.

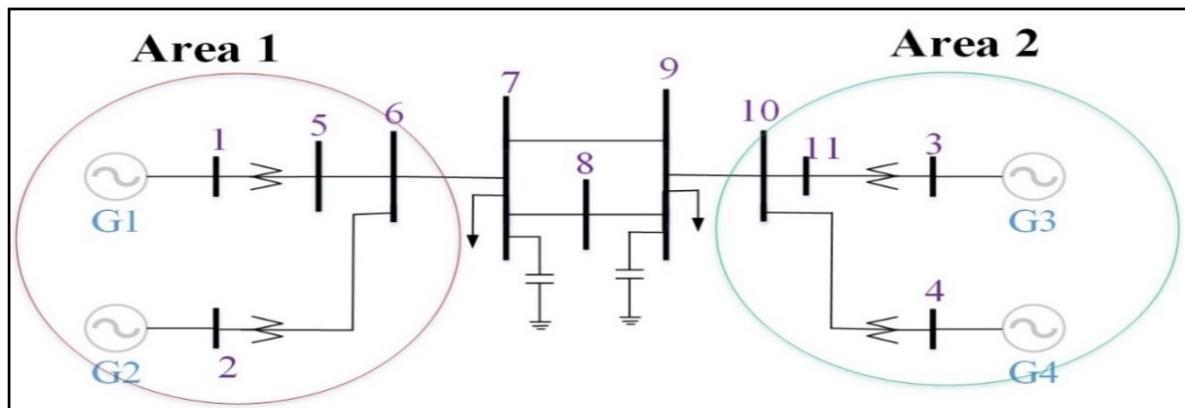


Figure-A I-1 Réseau test *Two-area System*

Les caractéristiques des machines sont regroupées dans le Tableau-A I-1 et les données de l'excitatrice des machines dans le Tableau-A I-2. Tous les générateurs utilisent un régulateur de tension de machine synchrone de type IEEE – I couplé à une excitatrice.

Tableau-A I-1 Caractéristiques des générateurs du réseau test *Two-area System*

	Générateurs	G1	G2	G3	G4
Paramètres					
S (MVA)	900	900	900	900	900
V (kV)	20	20	20	20	20
H (s)	6.5	6.5	6.5	6.5	6.5
T'_{d_0} (s)	8	8	8	8	8
T'_{q_0} (s)	0.4	0.4	0.4	0.4	0.4
x_d (p.u.)	1.8	1.8	1.8	1.8	1.8
x'_d (p.u.)	0.3	0.3	0.3	0.3	0.3
x_q (p.u.)	1.7	1.7	1.7	1.7	1.7
x'_q (p.u.)	0.55	0.55	0.55	0.55	0.55
R_s (p.u.)	0.0025	0.0025	0.0025	0.0025	0.0025

Tableau-A I-2 Paramètres de l'excitatrice des générateurs

	Paramètres	Tous les générateurs
Gain du régulateur	K_a	200
	K_e	1
	K_f	0
Constante de temps (s)	T_a	0.001
	T_e	0
	T_f	0

Les caractéristiques des lignes, des charges et des *shunt capacitors* sont fournies dans les tableaux suivants.

Tableau-A I-3 Caractéristiques des lignes

Barre <i>i</i> vers <i>j</i>		Résistance (p. u.)	Réactance (p. u.)	Susceptance (p. u.)
1	5	0.0000000	0.15	0.0000000
3	11	0.0000000	0.15	0.0000000
4	9	0.0000000	0.15	0.0000000
2	7	0.0000000	0.15	0.0000000
5	6	0.0001	0.001	0.00175
6	7	0.0001	0.001	0.00175
7	8	0.0001	0.001	0.00175
8	9	0.0001	0.001	0.00175
9	10	0.0001	0.001	0.00175
10	11	0.0001	0.001	0.00175

Tableau-A I-4 Caractéristiques des charges et des *shunt capacitors*

	Puissance active (MW)	Puissance réactive inductive (MVAR)	Puissance réactive capacitive (MVAR)
Les 2 charges	967 (1767)	100	187
Les 2 shunt capacitors	0	0	200 (350)

Réseau test *New England*

Le réseau *New England* est divisé en 6 régions cohérentes (Yusof, Rogers et Alden, 1993) comme le montre la représentation unifilaire de la Figure-A I-2. Il renferme 10 générateurs, 39 jeux de barres, 19 charges, 12 transformateurs et 46 lignes (IEEE).

Notez que les PMUs sont branchés aux générateurs et aux barres reliant les régions, en l'occurrence les barres 26, 18, 16, 8 et 2.

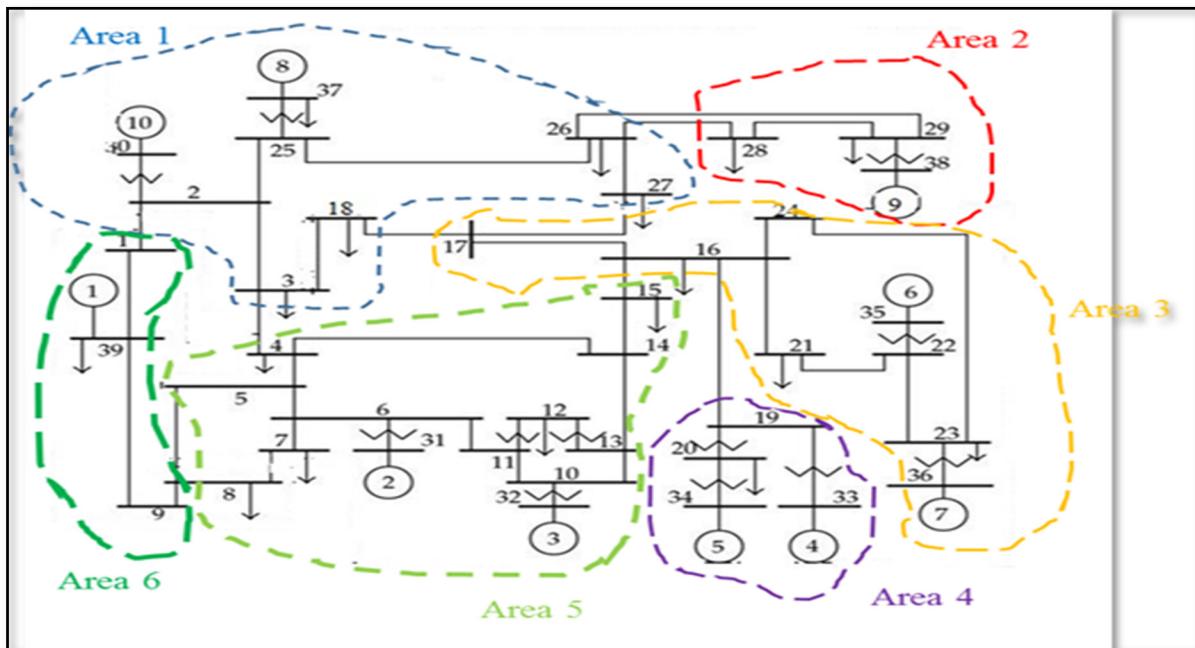


Figure-A I-2 Réseau test *New England*

Tous les générateurs utilisent un régulateur de tension de machine synchrone de type IEEE – I couplé à une excitatrice. La Figure-A I-3 montre le modèle d'une telle excitatrice (Sauer et Pai, 1998).

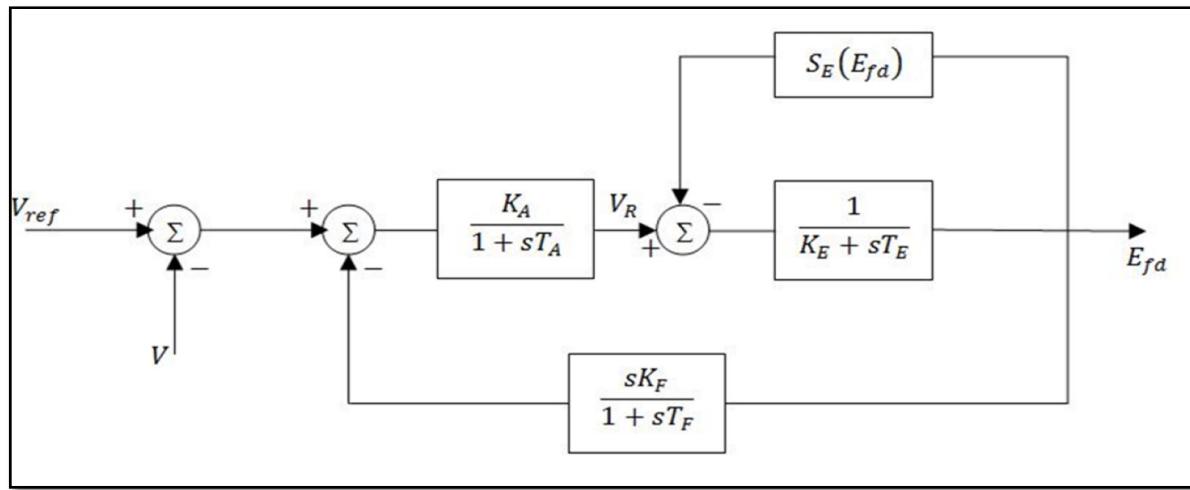


Figure-A I-3 Modèle excitatrice IEEE-Type I

Les données de l'excitatrice des machines sont données dans le Tableau-A I-5.

Tableau-A I-5 Paramètres de l'excitatrice des générateurs

	Paramètres	Tous les générateurs
Gain du régulateur	K_A	200
	K_E	1
	K_F	0
Constante de temps (s)	T_A	0.015
	T_B	10
	T_R	0.01
	T_C	1

On présente les caractéristiques des générateurs dans le Tableau-A I-6 et le Tableau-A I-7.

Tableau-A I-6 Caractéristiques des 5 premiers générateurs du réseau test *New England*

	Générateurs	G1	G2	G3	G4	G5
Paramètres						
x_l (p.u.)	0.03	0.35	0.304	0.295	0.54	
r_a (p.u.)	0.001	0.027	0.00386	0.00222	0.0014	
x_d (p.u.)	0.2	2.95	2.495	2.62	6.7	
x'_d (p.u.)	0.06	0.697	0.531	0.436	1.32	
x''_d (p.u.)	0.01	0.01	0.01	0.01	0.01	
T'_{d_o} (s)	7	6.56	5.7	5.69	5.4	
T''_{d_o} (s)	0.003	0.003	0.003	0.003	0.003	
x_q (p.u.)	0.19	2.82	2.37	2.58	6.2	
x'_q (p.u.)	0.08	1.7	0.876	1.66	1.66	
x''_q (p.u.)	0.03	0.03	0.03	0.03	0.03	
T'_{q_0} (s)	0.7	1.5	1.5	1.5	0.44	
T''_{q_0} (s)	0.005	0.005	0.005	0.005	0.005	
Puissance active (MW)	1000	520.81	650	632	508	
V (kV)	345	22	22	22	22	
H (s)	50	3.03	3.58	2.86	2.6	

Tableau-A I-7 Caractéristiques des 5 derniers générateurs du réseau test *New England*

	Générateurs	G6	G7	G8	G9	G10
Paramètres						
x_l (p.u.)	0.224	0.322	0.28	0.298	0.125	
r_a (p.u.)	0.0615	0.00268	0.00686	0.003	0.0014	
x_d (p.u.)	2.54	2.95	2.9	2.106	1	
x'_d (p.u.)	0.5	0.49	0.57	0.57	0.31	
x''_d (p.u.)	0.01	0.01	0.01	0.01	0.01	
T'_{d_o} (s)	7.3	5.66	6.7	4.79	10.2	

	Générateurs	G6	G7	G8	G9	G10
Paramètres						
$T_{d_o}''' (s)$	0.003	0.003	0.003	0.003	0.003	0.003
$x_q (p.u.)$	2.41	2.92	2.8	2.05	0.69	
$x'_q (p.u.)$	0.814	1.86	0.911	0.587	0.08	
$x''_q (p.u.)$	0.03	0.03	0.03	0.03	0.03	
$T'_{q_0} (s)$	0.4	1.5	0.41	1.96	1.5	
$T''_{q_0} (s)$	0.005	0.005	0.005	0.005	0.005	
Puissance active (MW)	650	560	540	830	250	
$V (kV)$	22	22	22	22	22	
$H (s)$	3.48	2.64	2.43	3.45	4.2	

Le Tableau-A I-8 donne les caractéristiques des lignes allant d'un bus à un autre. On fournit la résistance, la réactance, la charge, la valeur des changeurs de prise et la tension nominale transittée.

Tableau-A I-8 Caractéristiques des lignes, *New England*

Barre <i>i</i> vers <i>j</i>		Résistance (p. u.)	Réactance (p. u.)	Charge (p. u.)	Changeur de prise	Tension (kV)
1	2	0.0035	0.0411	0.6987	0	345
1	39	0.001	0.025	0.75	0	345
2	3	0.0013	0.0151	0.2572	0	345
2	25	0.007	0.0086	0.146	0	345
2	30	0	0.0181	0	1.025	22
3	4	0.0013	0.0213	0.2214	0	345
3	18	0.0011	0.0133	0.2138	0	345
4	5	0.0008	0.0128	0.1342	0	345
4	14	0.0008	0.0129	0.1382	0	345
5	8	0.0008	0.0112	0.1476	0	345

Barre <i>i</i> vers <i>j</i>		Résistance (<i>p.u.</i>)	Réactance (<i>p.u.</i>)	Charge (<i>p.u.</i>)	Changeur de prise	Tension (<i>kV</i>)
6	5	0.0002	0.0026	0.0434	0	345
6	7	0.0006	0.0092	0.113	0	345
6	11	0.0007	0.0082	0.1389	0	345
7	8	0.0004	0.0046	0.078	0	345
8	9	0.0023	0.0363	0.3804	0	345
9	39	0.001	0.025	1.2	0	345
10	11	0.0004	0.0043	0.0729	0	345
10	13	0.0004	0.0043	0.0729	0	345
10	32	0	0.02	0	1.07	22
12	11	0.0016	0.0435	0	1.006	345
12	13	0.0016	0.0435	0	1.006	345
13	14	0.0009	0.0101	0.1723	0	345
14	15	0.0018	0.0217	0.366	0	345
15	16	0.0009	0.0094	0.171	0	345
16	17	0.0007	0.0089	0.1342	0	345
16	19	0.0016	0.0195	0.304	0	345
16	21	0.0008	0.0135	0.2548	0	345
16	24	0.0003	0.0059	0.068	0	345
17	18	0.0007	0.0082	0.1319	0	345
17	27	0.0013	0.0173	0.3216	0	345
19	33	0.0007	0.0142	0	1.07	22
19	20	0.0007	0.0138	0	1.06	345
20	34	0.0009	0.018	0	1.009	22
21	22	0.0008	0.014	0.2565	0	345
22	23	0.0006	0.0096	0.1846	0	345
22	35	0	0.0143	0	1.025	22
23	24	0.0022	0.035	0.361	0	345

Barre <i>i</i> vers <i>j</i>		Résistance (<i>p. u.</i>)	Réactance (<i>p. u.</i>)	Charge (<i>p. u.</i>)	Changeur de prise	Tension (<i>kV</i>)
23	36	0.0005	0.0272	0	1	22
25	26	0.0032	0.0323	0.513	0	345
25	37	0.0006	0.0232	0	1.025	22
26	27	0.0014	0.0147	0.2396	0	345
26	28	0.0043	0.0474	0.7802	0	345
26	29	0.0057	0.0625	1.029	0	345
28	29	0.0014	0.0151	0.249	0	345
29	38	0.0008	0.0156	0	1.025	22
31	6	0	0.025	0	1	22

Le Tableau-A I-9 fournit les caractéristiques des charges du réseau.

Tableau-A I-9 Caractéristiques des charges du réseau test *New England*

Barres	Puissance active (MW)	Puissance réactive capacitive (MVAR)	Puissance réactive inductive (MVAR)
25	224	0	47.2
26	139	0	17
27	281	0	75.5
28	206	0	27.6
29	283.5	0	29.6
18	158	0	30
16	329	0	32.3
24	308.6	92	0
33	322	0	2.4
39	1104	0	250
8	522	0	176
12	233	0	84

Barres	Puissance active (MW)	Puissance réactive capacitive (MVAR)	Puissance réactive inductive (MVAR)
4	500	0	184
15	320	0	153
20	628	0	103
23	247.5	0	84.6
21	274.6	0	115.6
3	322	0	2.4
7	233	0	84

Les caractéristiques des transformateurs se trouvent dans le Tableau-A I-10

Transformateurs	Barre <i>i</i> vers <i>j</i>		Résistance (p. u.)	Réactance (p. u.)	Changeur de prise
1	2	30	0	0.0181	1.025
2	6	31	0	0.025	1.07
3	10	32	0	0.02	1.07
4	12	11	0.0016	0.0435	1.006
5	12	13	0.0016	0.0435	1.006
6	19	20	0.0007	0.0138	1.06
7	19	33	0.0007	0.0142	1.07
8	20	34	0.0009	0.018	1.009
9	22	35	0	0.0143	1.025
10	25	37	0.0006	0.0232	1.025
11	29	38	0.0008	0.0156	1.025
12	23	36	0.0005	0.0272	1.0

ANNEXE II

CARACTÉRISTIQUES DU RÉSEAU ÉLECTRIQUE *SIMPLIFIED 14-GENERATORS AUSTRALIAN TEST SYSTEM*

Cette annexe présente les caractéristiques du modèle simplifié 14 générateurs du réseau australien. Dans cette thèse, il est dénommé réseau Australien simplifié pour alléger l'écriture mais le nom retrouvé dans la littérature est *Simplified 14-Generators Australian Test System*. Comme montré à la Figure-A II-1, il renferme 14 générateurs, 59 bus et est divisé en 5 régions. Les générateurs sont identifiés par des lettres suivie d'un nombre indiquant la région dans laquelle ils se trouvent (par exemple, VPS_2 correspond à un générateur se trouvant dans la région 2). On applique au modèle 29 charges et 9 compensateurs statiques d'énergie réactive (*Static VAR compensator, SVC*) identifiés suivant la même terminologie utilisée pour les générateurs (par exemple, PSVC_5 est un SVC dans la 5^{ème} région). Les PMUs sont placés sur les barres 102, 309, 315 et 205. Notez qu'une base de 100 MVA est choisie pour toutes les conversions en per unit.

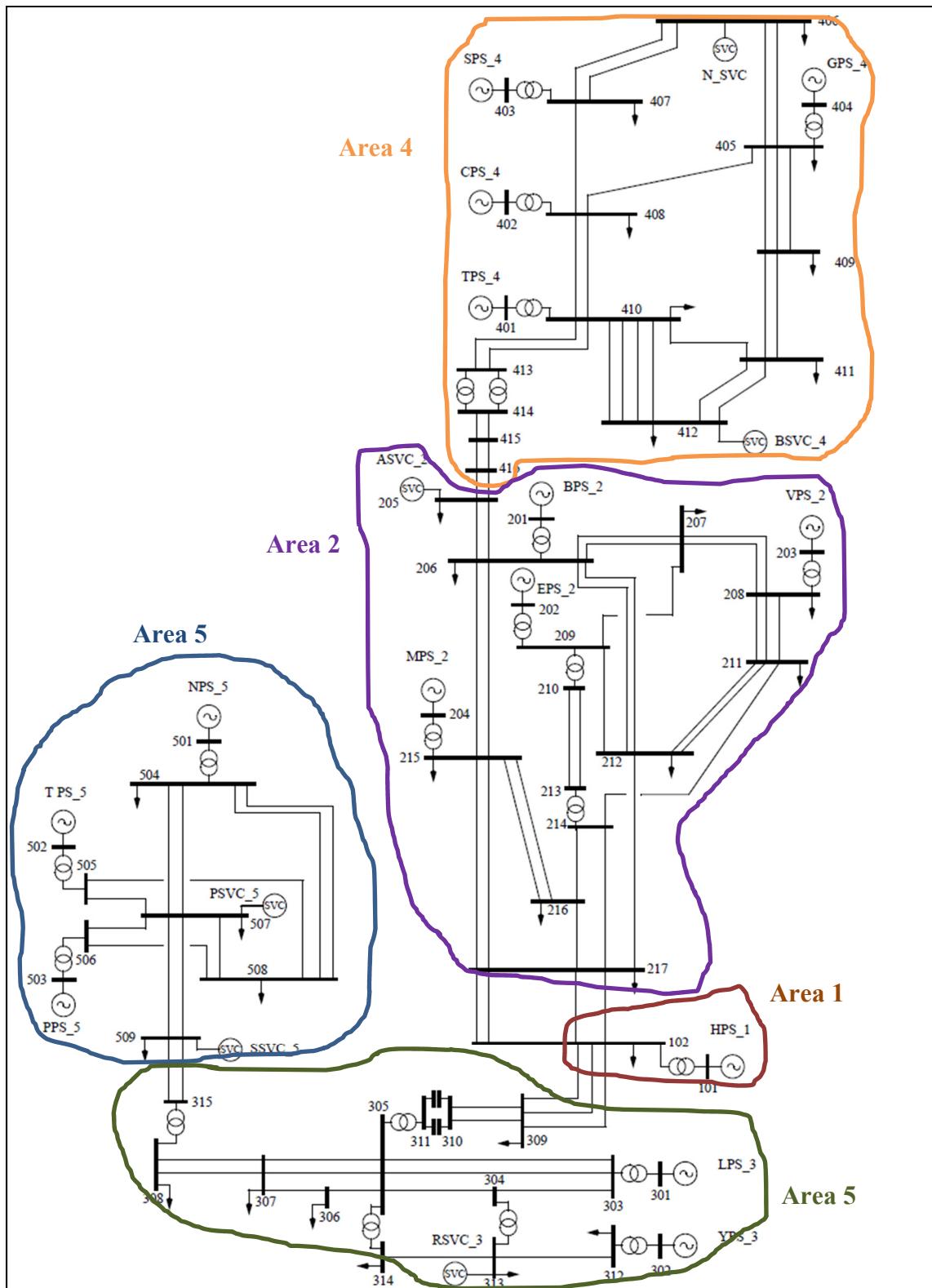


Figure-A II-1 Simplified 14-Generators Australian Test System

Les paramètres des générateurs dans les régions 1, 2 et 3 se retrouvent dans le Tableau-A II-1.

Tableau-A II-1 Paramètres des générateurs dans les régions 1, 2 et 3
Simplified 14-Generators Australian Test System

	Générateurs	HPS_1	BPS_2	EPS_2	MPS_2	VPS_2	LPS_3	YPS_3
Paramètres								
S (MVA)	333.3	666.7	555.6	666.7	555.6	666.7	444.4	
Nombre d'unités	12	6	5	6	4	8	4	
H (s)	3.6	3.2	2.8	3.2	2.6	2.8	3.5	
x_a (p.u.)	0.14	0.2	0.17	0.2	0.2	0.2	0.15	
x_d (p.u.)	1.1	1.8	2.2	1.8	2.3	2.7	2	
x_q (p.u.)	0.65	1.75	2.1	1.75	1.7	1.5	1.8	
x'_d (p.u.)	0.25	0.3	0.3	0.3	0.3	0.3	0.25	
T'_{d_0} (p.u.)	8.5	8.5	4.5	8.5	5	7.5	7.5	
x''_d (p.u.)	0.25	0.21	0.2	0.21	0.25	0.25	0.2	
T''_{d_0} (s)	0.05	0.04	0.04	0.04	0.03	0.04	0.04	
x'_q (p.u.)	-	0.7	0.5	0.7	0.4	0.85	-	
T'_{q_0} (s)	-	0.3	1.5	0.3	2	0.85	-	
x''_q (p.u.)	0.25	0.21	0.21	0.21	0.25	0.25	0.2	
T''_{q_0} (s)	0.2	0.08	0.06	0.08	0.25	0.12	0.25	

Les paramètres des générateurs dans les régions 4 et 5 sont présentés dans le Tableau-A II-2

Tableau-A II-2 Paramètres des générateurs dans les régions 4 et 5
Simplified 14-Generators Australian Test System

	Générateurs	CPS_4	GPS_4	SPS_4	TPS_4	NPS_5	TPS_5	PPS_5
Paramètres								
S (MVA)	333.3	333.3	444.4	444.4	333.3	250	166.7	
Nombre d'unités	3	6	4	4	2	4	6	
H (s)	3	4	2.6	2.6	3.5	4	7.5	
x_a (p.u.)	0.2	0.18	0.2	0.2	0.15	0.2	0.15	
x_d (p.u.)	1.9	2.2	2.3	2.3	2.2	2	2.3	
x_q (p.u.)	1.8	1.4	1.7	1.7	1.7	1.5	2	
x'_d (p.u.)	0.3	0.32	0.3	0.3	0.3	0.3	0.25	
T'_{d_0} (p.u.)	6.5	9	5	5	7.5	7.5	5	
x''_d (p.u.)	0.26	0.24	0.25	0.25	0.24	0.22	0.17	
T''_{d_0} (s)	0.035	0.04	0.03	0.03	0.025	0.04	0.022	
x'_q (p.u.)	0.55	0.75	0.4	0.4	0.8	0.8	0.35	
T'_{q_0} (s)	1.4	1.4	2	2	1.5	3	1	
x''_q (p.u.)	0.26	0.24	0.25	0.25	0.24	0.22	0.17	
T''_{q_0} (s)	0.04	0.13	0.25	0.25	0.1	0.2	0.035	

Trois types d'excitatrice sont utilisés le ST5B, AC4A et AC1A (IEEE et Standard, 2006). La Figure-A II-2 représente le modèle petit signal de l'excitatrice de types STB et AC4A quand ($T_{B_1} = T_{C_1} = 0$).

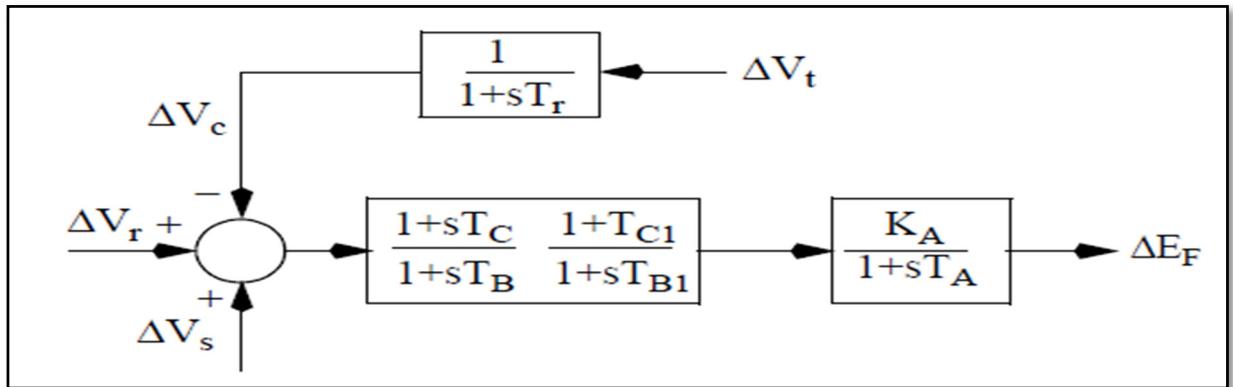


Figure-A II-2 Modèle petit signal de l'excitatrice de types ST5B et AC4A avec $T_{B_1} = T_{C_1} = 0$

La Figure-A II-3 représente le système d'excitation de type AC1A en mode petit signal avec les effets de démagnétisation du courant de champ négligés.

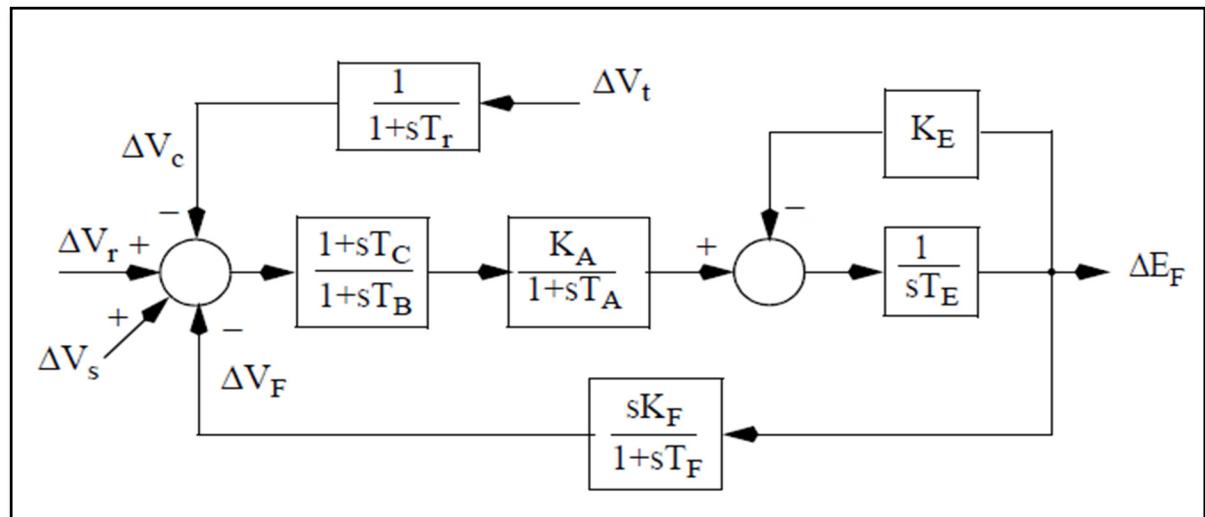


Figure-A II-3 Modèle petit signal de l'excitatrice de types AC1A

Les paramètres des excitatrices des générateurs dans les régions 1, 2 et 3 se retrouvent dans le Tableau-A II-3.

Tableau-A II-3 Caractéristiques des excitatrices des générateurs des régions 1, 2 et 3,
Simplified 14-Generators Australian Test System

	Générateur s	HPS_1	BPS_2	EPS_2	MPS_2	VPS_2	LPS_3	YPS_3
Paramètres		Types		AC4A	AC4A	AC1A	AC4A	AC4A
Gain du régulateu r	K_A	200	400	400	400	300	400	200
	K_E	-	-	1	-	-	-	1
	K_F	-	-	0.029	-	-	-	0.02
Constante de temps (s)	T_A	0.1	0.02	0.02	0.02	0.01	0.05	0.05
	T_B	13.25	1.12	0	1.12	0.7	6.42	0
	T_F	-	-	1	-	-	-	0.8
	T_E	-	-	1	-	-	-	1.333
	T_R	0	0	0	0	0	0	0
	T_C	2.5	0.5	0	0.5	0.35	1.14	0

Les paramètres des excitatrices des générateurs dans les régions 4 et 5 sont dans le Tableau-A II-4.

Tableau-A II-4 Caractéristiques des excitatrices des générateurs des régions 1, 2 et 3,
Simplified 14-Generators Australian Test System

	Générateur s	CPS_4	GPS_4	SPS_4	TPS_4	NPS_5	TPS_5	PPS_5
Paramètres		Types		AC4A	AC4A	AC4A	AC4A	AC1A
Gain du régulateu r	K_A	300	250	300	300	1000	400	300
	K_E	-	-	-	-	1	-	-
	K_F	-	-	-	-	0.004	-	-
	T_A	0.05	0.2	0.01	0.1	0.04	0.5	0.01

	Générateur s	CPS_4	GPS_4	SPS_4	TPS_4	NPS_5	TPS_5	PPS_5
Paramètres								
Constante de temps (s)	T_B	9.8	0.0232	0.7	40	0	16	0.8
	T_F	-	-	-	-	0.27	-	-
	T_E	-	-	-	-	0.87	-	-
	T_{B_1}	0	0	0	0	0	0.05	0
	T_{C_1}	0	0	0	0	0	0.6	0
	T_R	0.02	0	0	0	0	0	0
	T_c	1.52	0.136	0.35	4	0	1.4	0.2

Le modèle petit signal du système d'excitation des SVC utilisés est représenté à la Figure-A II-4 avec les gains utilisés ; la base nominale étant de 100 MVA.

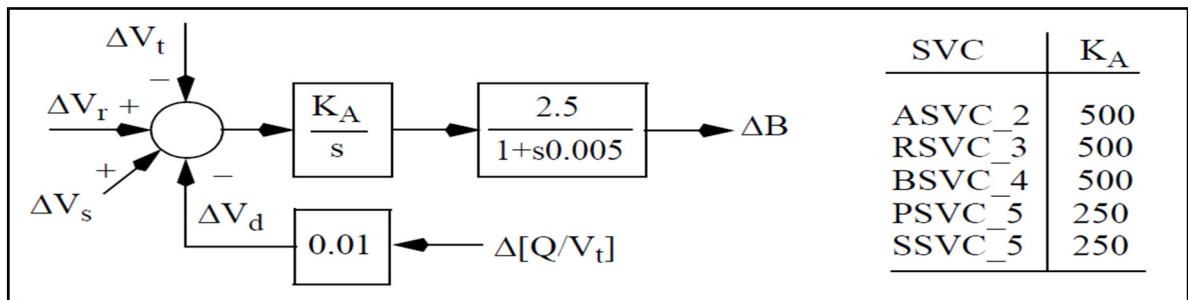


Figure-A II-4 Modèle petit signal du système d'excitation des SVC,
Simplified 14-Generators Australian Test System

Les caractéristiques des lignes sont fournies dans le Tableau-A II-5

Tableau-A II-5 Caractéristiques des lignes, *Simplified 14-Generators Australian Test System*

Barre <i>i</i> vers <i>j</i>		Numéro de ligne	Résistance (p. u.)	Réactance (p. u.)	Susceptance (p. u.)
102	217	1, 2	0.0084	0.0667	0.817
102	217	3, 4	0.0078	0.062	0.76
102	309	1, 2	0.0045	0.0356	0.437
102	309	3	0.0109	0.0868	0.76
205	206	1, 2	0.0096	0.076	0.931
205	416	1, 2	0.0037	0.046	0.73
206	207	1, 2	0.0045	0.0356	0.437
206	212	1, 2	0.0066	0.0527	0.646
206	215	1, 2	0.0066	0.0527	0.646
207	208	1, 2	0.0018	0.014	0.171
207	209	1	0.0008	0.0062	0.076
208	211	1, 2, 3	0.0031	0.0248	0.304
209	212	1	0.0045	0.0356	0.437
210	213	1, 2	0.001	0.0145	1.54
211	212	1, 2	0.0014	0.0108	0.133
211	214	1	0.0019	0.0155	0.19
212	217	1	0.007	0.0558	0.684
214	216	1	0.001	0.0077	0.095
214	217	1	0.0049	0.0388	0.475
215	216	1, 2	0.0051	0.0403	0.494
215	217	1, 2	0.0072	0.0574	0.703
216	217	1	0.0051	0.0403	0.494
303	304	1	0.001	0.014	1.48
303	305	1, 2	0.0011	0.016	1.7
304	305	1	0.0003	0.004	0.424
305	306	1	0.0002	0.003	0.32

Barre <i>i</i> vers <i>j</i>		Numéro de ligne	Résistance (p. u.)	Réactance (p. u.)	Susceptance (p. u.)
305	307	1,2	0.0003	0.0045	0.447
306	307	1	0.0001	0.0012	0.127
307	308	1, 2	0.0023	0.0325	3.445
309	310	1, 2	0.009	0.0713	0.874
310	311	1, 2	0	-0.0337	0
312	313	1	0.002	0.015	0.9
313	314	1	0.0005	0.005	0.52
315	509	1, 2	0.007	0.05	0.19
405	406	1, 2	0.0039	0.0475	0.381
405	408	1	0.0054	0.05	0.189
405	409	1, 2, 3	0.018	0.122	0.79
406	407	1, 2	0.0006	0.0076	0.062
407	408	1	0.0042	0.0513	0.412
408	410	1,2	0.011	0.128	1.01
409	411	1,2	0.0103	0.0709	0.46
410	411	1	0.0043	0.0532	0.427
410	412	1, 2, 3, 4	0.0043	0.0532	0.427
410	413	1, 2	0.004	0.0494	0.4
411	412	1, 2	0.0012	0.0152	0.122
414	415	1, 2	0.002	0.025	0.39
415	416	1, 2	0.0037	0.046	0.73
504	507	1, 2	0.023	0.15	0.56
504	508	1, 2	0.026	0.019	0.87
505	507	1	0.0008	0.0085	0.06
505	508	1	0.0025	0.028	0.17
506	507	1	0.0008	0.0085	0.06
506	508	1	0.003	0.028	0.14

Barre <i>i</i> vers <i>j</i>		Numéro de ligne	Résistance (p. u.)	Réactance (p. u.)	Susceptance (p. u.)
507	508	1	0.002	0.019	0.09
507	509	1, 2	0.03	0.22	0.9

Pour les transformateurs utilisés, la Figure A II-4 montre la convention employée pour le rapport de transformation des transformateurs.

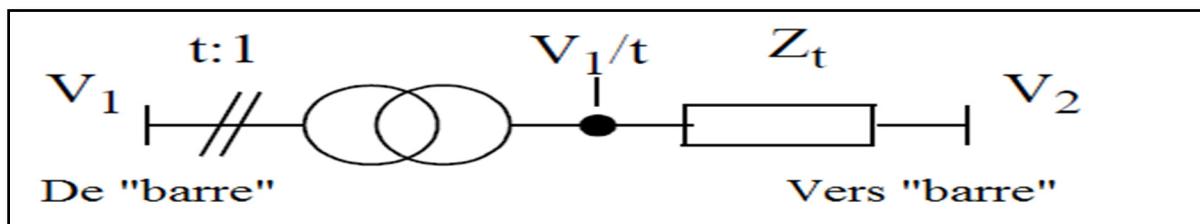


Figure-A II-5 Convention de rapport de transformation

On indique dans le Tableau-A II-6 le rapport de transformation de chacun des transformateurs, leur réactance ainsi que la part de la puissance nominale du générateur qui y est associé.

Tableau-A II-6 Caractéristiques des transformateurs
Simplified 14-Generators Australian Test System

Barre <i>i</i> vers <i>j</i>		Résistance par transformateur		Rapport de transformation
		% <i>S</i>	(p. u.)	
101	102	12	0.036	0.948
201	206	16	0.024	0.939
202	209	16	0.0288	0.939
203	208	17	0.0306	0.939
204	215	16	0.024	0.939
209	210	17	0.0272	0.976
213	214	17	0.0272	1

Barre <i>i</i> vers <i>j</i>		Résistance par transformateur		Rapport de transformation
		% <i>S</i>	(p.u.)	
301	303	16	0.024	0.93
302	312	15	0.0338	0.952
304	313	16	0.032	0.948
305	311	12	0.024	1
305	314	17	0.0243	1
308	315	10	0.027	1
401	410	15	0.0338	0.939
402	408	17	0.051	0.952
403	407	15	0.0338	0.952
404	405	17	0.051	0.952
413	414	6	0.008	1
501	504	17	0.051	0.952
502	505	16	0.064	0.93
503	506	16.7	0.1	0.93

Les caractéristiques des charges sont données dans le Tableau-A II-7.

Tableau-A II-7 Caractéristiques des charges
Simplified 14-Generators Australian Test System

Barres	Puissance active (MW)	Puissance réactive inductive (MVAR)
102	475	50
205	410	40
206	140	15
207	1975	200
208	220	25

Barres	Puissance active (MW)	Puissance réactive inductive (MVAR)
211	1785	180
212	1740	180
215	505	50
216	1930	200
217	1320	140
306	1450	150
307	770	80
308	770	80
309	230	25
312	140	15
313	2840	290
314	300	30
405	1215	120
406	905	90
407	0	0
408	185	20
409	310	30
410	650	65
411	700	70
412	1535	155
504	300	60
507	1100	220
508	800	160
509	100	20

Les puissances des bancs de condensateurs (*shunt capacitor*) et de réacteurs (*reactor banks*) sont fournies dans le Tableau-A II-8

Tableau-A II-8 Caractéristiques des shunt capacitors et des reactor banks
Simplified 14-Generators Australian Test System

	Barres	Q (MVAR)
<i>Shunt capacitor</i>	211	100
	212	150
	216	150
	409	60
	411	30
<i>Reactor banks</i>	414	30
	415	60
	416	60
	504	90

ANNEXE III

EXEMPLES D'APPLICATIONS DES ARBRES DE DÉCISION ET DE LA FORÊT ALÉATOIRE D'ARBRE DE DÉCISION

Cette annexe propose un aperçu sommaire de l'utilisation de **R** dans l'analyse statistique des données et la construction de modèles de prédiction en se basant sur des techniques de forage de données. On rappelle qu'avec l'interface de base de **R** montrée à la Figure-A III-1, les tâches se réalisent grâce à des lignes de code.

```
RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
[Icons]

> library(rattle)
> rattle()
> building <- TRUE
> scoring <- ! building
>
> # A pre-defined value is used to reset the random seed so that results are repeatable.
>
> crv$seed <- 42
>
> # =====
> # Rattle timestamp: 2016-09-09 11:23:02 i386-w64-mingw32
>
> # Load the data.
>
> crs$dataset <- read.csv(system.file("csv", "weather.csv", package="rattle"), encoding="UTF-8")
>
> # =====
> # Rattle timestamp: 2016-09-09 11:23:02 i386-w64-mingw32
>
> # Note the user selections.
>
> # Build the training/validate/test datasets.
>
> set.seed(crv$seed)
> crs$obs <- nrow(crs$dataset) # 366 observations
> crs$sample <- crs$train <- sample(nrow(crs$dataset), 0.7*crs$obs) # 256 observations
> crs$validate <- sample(setdiff(seq_len(nrow(crs$dataset)), crs$train), 0.15*crs$obs) # 54 observations
> crs$test <- setdiff(setdiff(seq_len(nrow(crs$dataset)), crs$train), crs$validate) # 56 observations
>
> # The following variable selections have been noted.
>
> crs$input <- c("MinTemp", "MaxTemp", "Rainfall", "Evaporation", "Sunshine", "WindGustDir", "WindGustSpeed", "WindDir9am", "WindDir3pm",
+ "WindSpeed9am", "WindSpeed3pm", "Humidity9am", "Humidity3pm", "Pressure9am", "Pressure3pm", "Cloud9am", "Cloud3pm", "Temp9am", "Temp3pm",
+ "RainToday")
>
> crs$numeric <- c("MinTemp", "MaxTemp", "Rainfall", "Evaporation", "Sunshine", "WindGustSpeed", "WindSpeed9am", "WindSpeed3pm",
+ "Humidity9am", "Humidity3pm", "Pressure9am", "Pressure3pm", "Cloud9am", "Cloud3pm", "Temp9am", "Temp3pm")
>
> crs$categorical <- c("WindGustDir", "WindDir9am", "WindDir3pm", "RainToday")
>
> crs$target <- "RainTomorrow"
```

Figure-A III-1 Interface de R sous Windows

Sur la Figure-A III-1, on observe qu'on a appelé la librairie *Rattle* pour avoir une interface plus conviviale pour traiter nos données. Par la suite, on a appelé un fichier de données *crs\$dataset* renfermant 366 observations. Comme mentionné dans la section 5.6, on l'a divisé en trois échantillons *crs\$sample*, *crs\$validate*, *crs\$test* suivant une proportion 70%, 15%, 15% de la base de données initiale. On peut voir également que les prédicteurs ainsi que la variable cible

ont été définies. L'objectif est de prédire la probabilité de pluie demain (*RainTomorrow*) étant donné les conditions d'aujourd'hui. Pour poursuivre, on utilise l'interface Rattle pour exécuter les mêmes lignes de code que précédemment en faisant des clicks avec la souris. On s'entend pour dire que l'interface Rattle est plus abordable pour manipuler les données. Il apparaît clairement les différents algorithmes de data mining, les différentes manipulations que l'on peut faire sur les données et une zone de texte pour afficher les résultats de simulation. Cependant, si l'on veut faire des études poussées et apportées des changements spécifiques dans des lignes de codes, on doit se référer obligatoirement à l'interface de base de **R**.

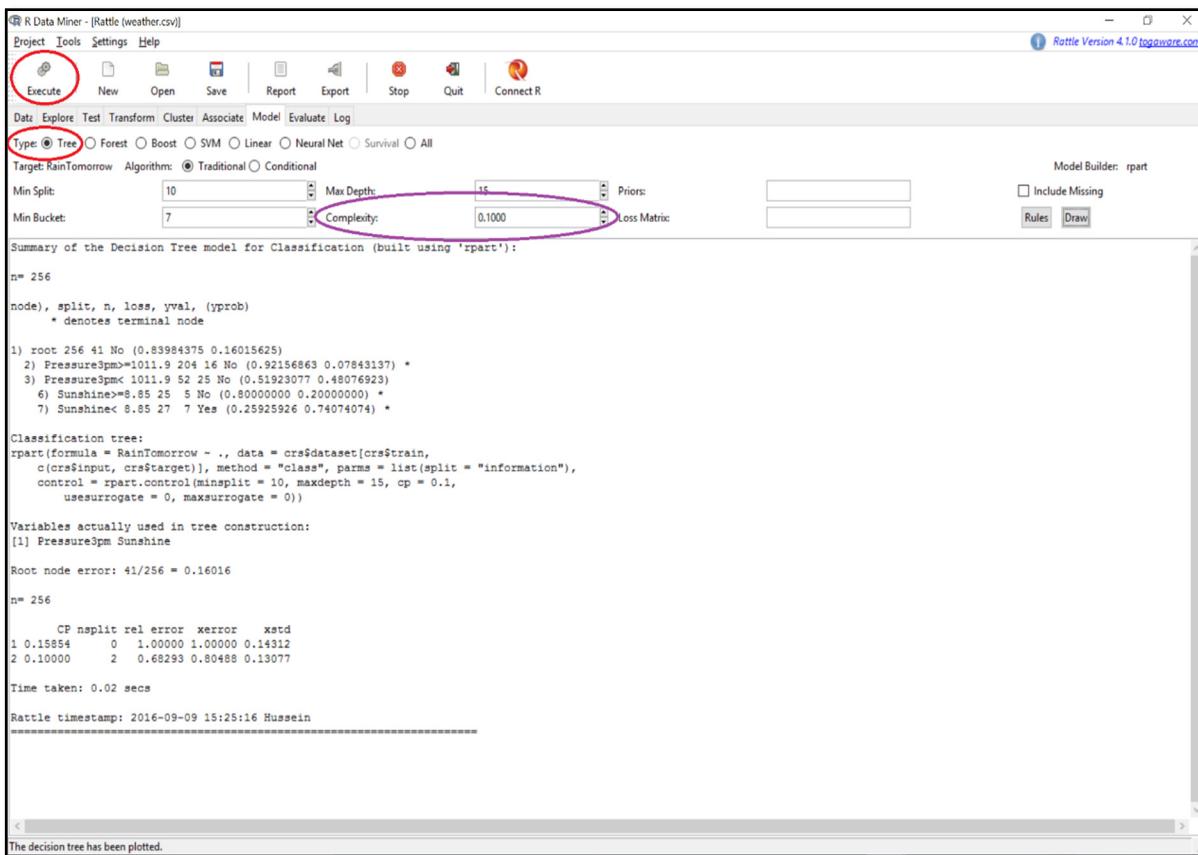


Figure-A III-2 Interface de Rattle sous Windows lors de la construction d'un arbre de décision

Sur la Figure-A III-2, on montre la construction d'un arbre de décision où *Complexity* est la variable qui permet de contrôler la profondeur et l'élagage de l'arbre construit. Une valeur de

0 produit un arbre non élagué tel il est le cas dans une forêt aléatoire (RF). La variable MinSplit désigne le nombre minimal d'observations qui doivent exister dans un nœud pour tenter la division des nœuds en nœuds filles. La variable MinBucket indique le nombre d'observations à considérer dans les feuilles (ou nœuds terminaux) et la variable MaxDepth fixe la profondeur maximale d'un nœud de l'arbre final en comptant le nœud racine ayant pour profondeur 0. On dispose également d'un résumé de la construction de l'arbre avec les règles, le nombre d'observations considérées et les variables prises en compte dans l'apprentissage. La Figure-A III-3 et la Figure-A III-4 montrent deux arbres de décision différents pour le même jeu de donnée en modifiant la profondeur des arbres : le premier étant profondeur maximale (non élagué) et le second élagué. Dans les deux cas, on remarque la facilité de lire les règles rapidement et de pouvoir classer une nouvelle observation sachant les valeurs prises par chacun des prédicteurs.

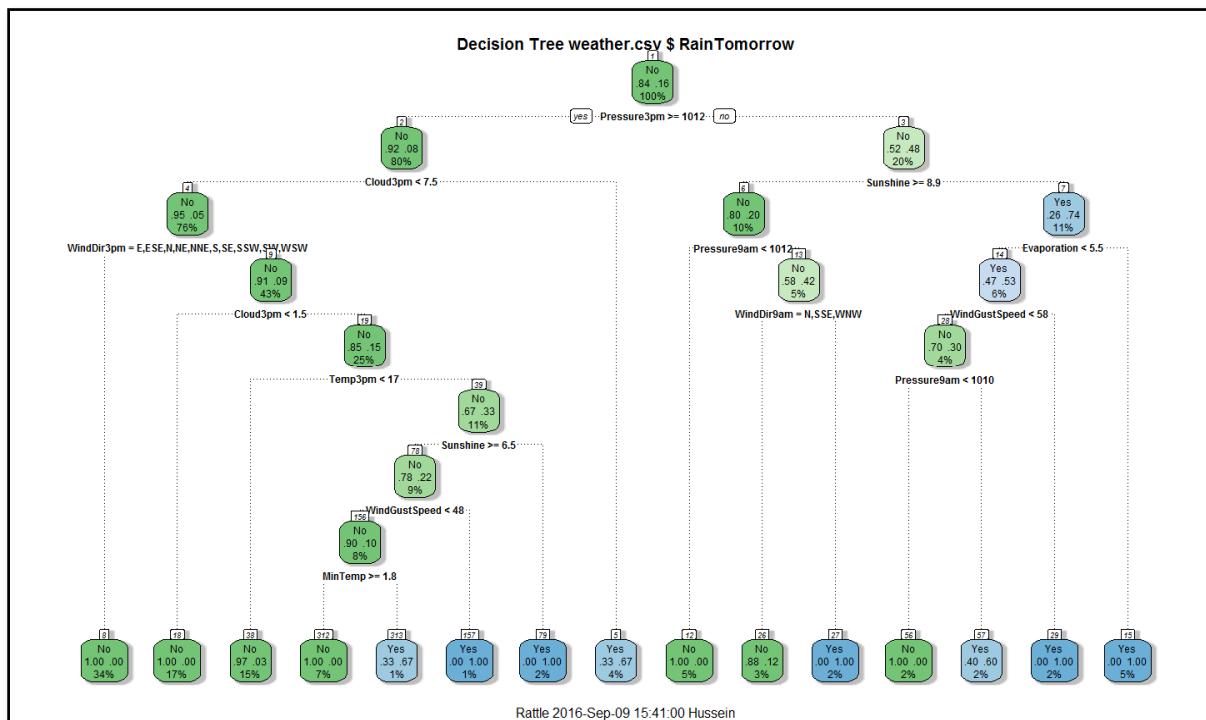


Figure-A III-3 Arbre de décision non élagué avec RainTomorrow comme variable cible

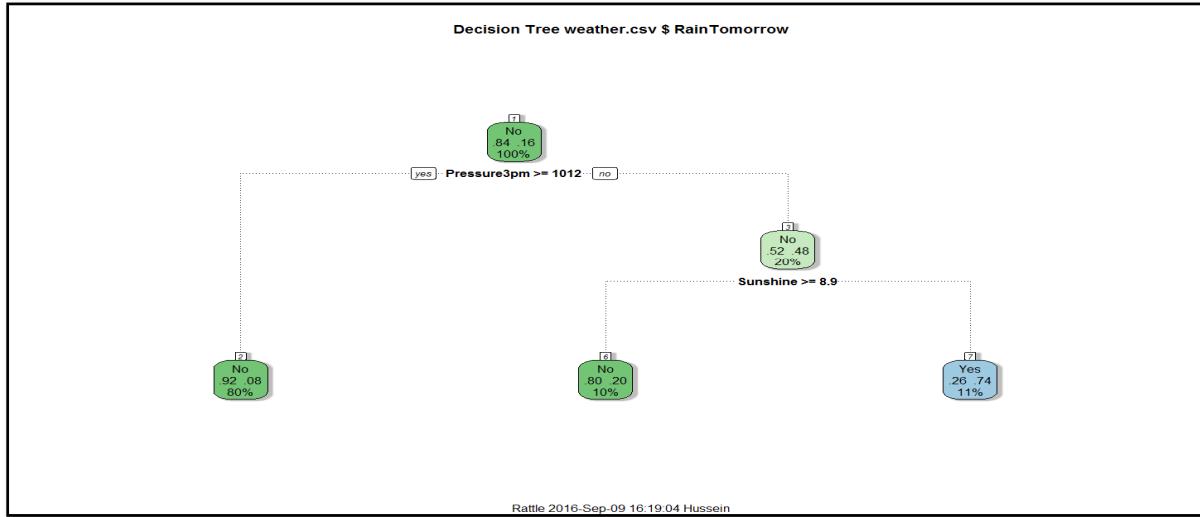


Figure-A III-4 Arbre de décision élagué avec RainTomorrow comme variable cible

Compte tenu du fait que l'arbre de décision final varie selon la profondeur souhaitée pour l'arbre de décision, on utilise une forêt aléatoire d'arbres de décision pour avoir un plus grand nombre d'arbres pris en compte en sélectionnant au hasard les prédicteurs dans la construction des arbres. La Figure-A III-5 montre l'interface Rattle correspondant à cette démarche. Dans la zone de texte, on peut voir le nombre d'arbres qui a été utilisé pour construire le modèle soit 500, le nombre d'observations utilisées soit 256 et le nombre de variables considérées lors de la scission de chaque nœud soit 4 variables ($\sqrt{23} \cong 4$). On fait remarquer que la base de données a été divisée en trois échantillons comme précédemment et qu'il s'agit des mêmes échantillons. Notez également que la matrice de confusion est fournie. Cette dernière enregistre le désaccord entre les prédictions du modèle final et les résultats réels des observations d'apprentissage. Les observations réelles sont les lignes de cette table, tandis que les colonnes correspondent à ce que le modèle prédit pour une observation et les cellules comptent le nombre d'observations dans chaque catégorie. Autrement dit, le modèle et l'échantillon d'apprentissage conviennent qu'il ne pleuvra pas pour 205 des observations. Ils conviennent qu'il va pleuvoir pour 14 des observations. Cependant, il y a 27 jours pour lesquels le modèle prédit qu'il ne pleut pas le jour suivant, et pourtant il pleut. De même, le modèle prédit qu'il va pleuvoir le lendemain pour dix des observations alors qu'en fait, il ne pleut pas.

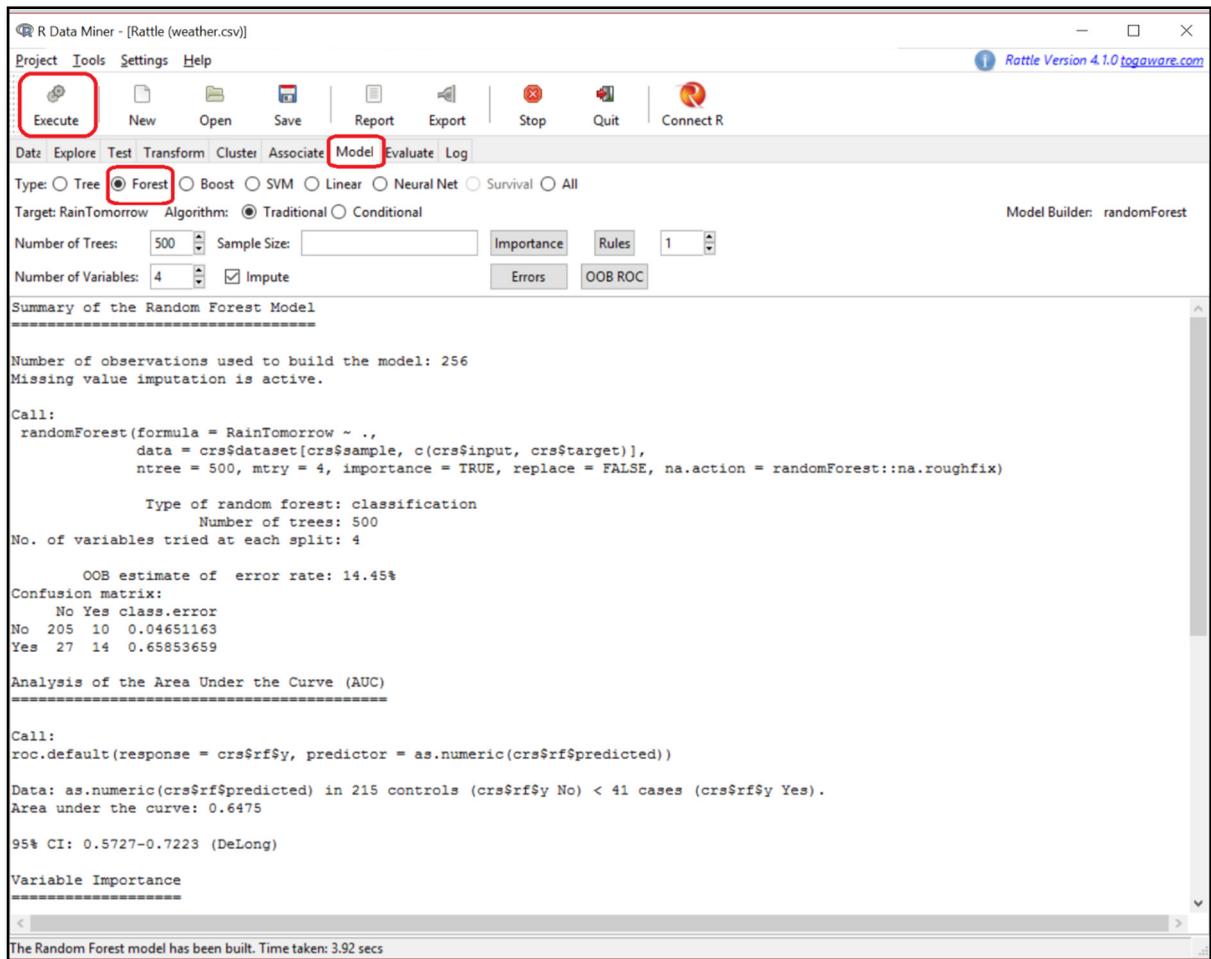


Figure-A III-5 Construction d'une forêt aléatoire avec Rattle

La Figure-A III-6 représente l'importance des variables suivant le MDA et le MDG. On rappelle que les variables se trouvant en tête de liste indiquent qu'elles sont « relativement plus importantes » dans la construction de la forêt aléatoire. Pressure3pm et Sunshine sont les variables les plus importantes suivant le MDG. Selon le MDA, on retrouve Cloud3pm comme deuxième variable plus importante. Remarquez la cohérence de cette donnée par rapport à la construction d'un seul arbre non élagué et de l'arbre élagué de la Figure-A III-4 et la Figure-A III-5. Ces mêmes variables ont servi pour la construction de l'arbre de décision final précédemment présenté.

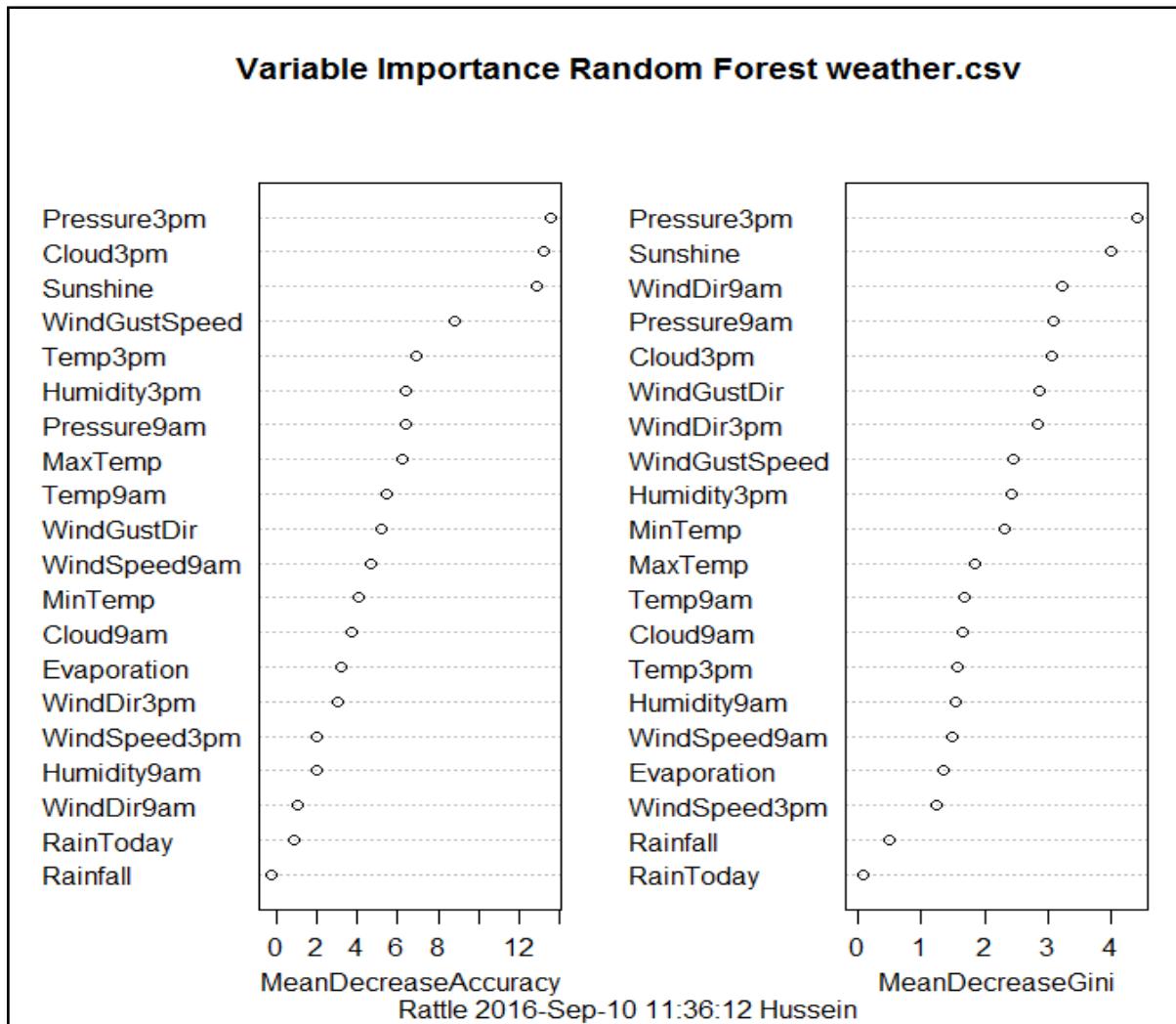


Figure-A III-6 Importance des variables (MDA, MDG) mesurées par la forêt aléatoire

On présente l'erreur OOB sur la Figure-A III-7. Le graphique rapporte la précision de la forêt en termes de taux d'erreur sur l'axe des ordonnées versus le nombre d'arbres qui ont été inclus dans la forêt (axe des abscisses). L'information essentielle à retenir est qu'à partir d'un certain nombre d'arbres, l'ajout de nouveaux arbres n'apporte que très peu de changements dans le modèle. Selon le graphique, le modèle n'a besoin que de 25 arbres pour avoir un très bon rendement.

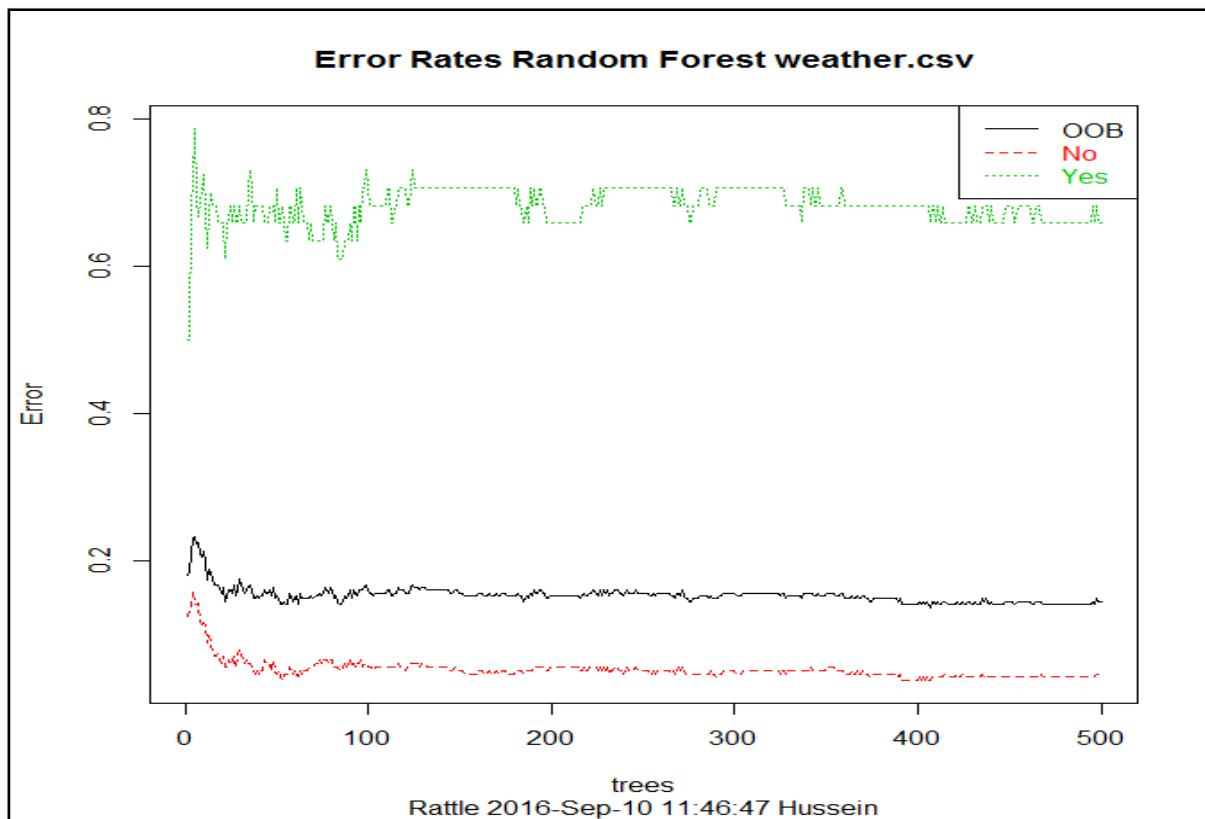


Figure-A III-7 Erreur OOB du modèle en fonction du nombre d'arbres de la forêt aléatoire

Comme il apparaît sur la Figure-A III-8, l'option Boost de l'onglet des Modèles construit un ensemble de 50 arbres en utilisant l'approche de renforcement des observations mal classées. Pour la construction de chacun des arbres, la méthode développée plus haut est utilisée. Dans la zone de texte, on peut voir la matrice de confusion qui indique la performance du modèle. Il est aussi mentionné les variables utilisées dans la construction des arbres et à quelle fréquence ces variables ont été appelées. Par exemple, comme prédicteur, la variable WindDir9am a été utilisée à 36 reprises et WindSSpeed9am à une seule reprise.

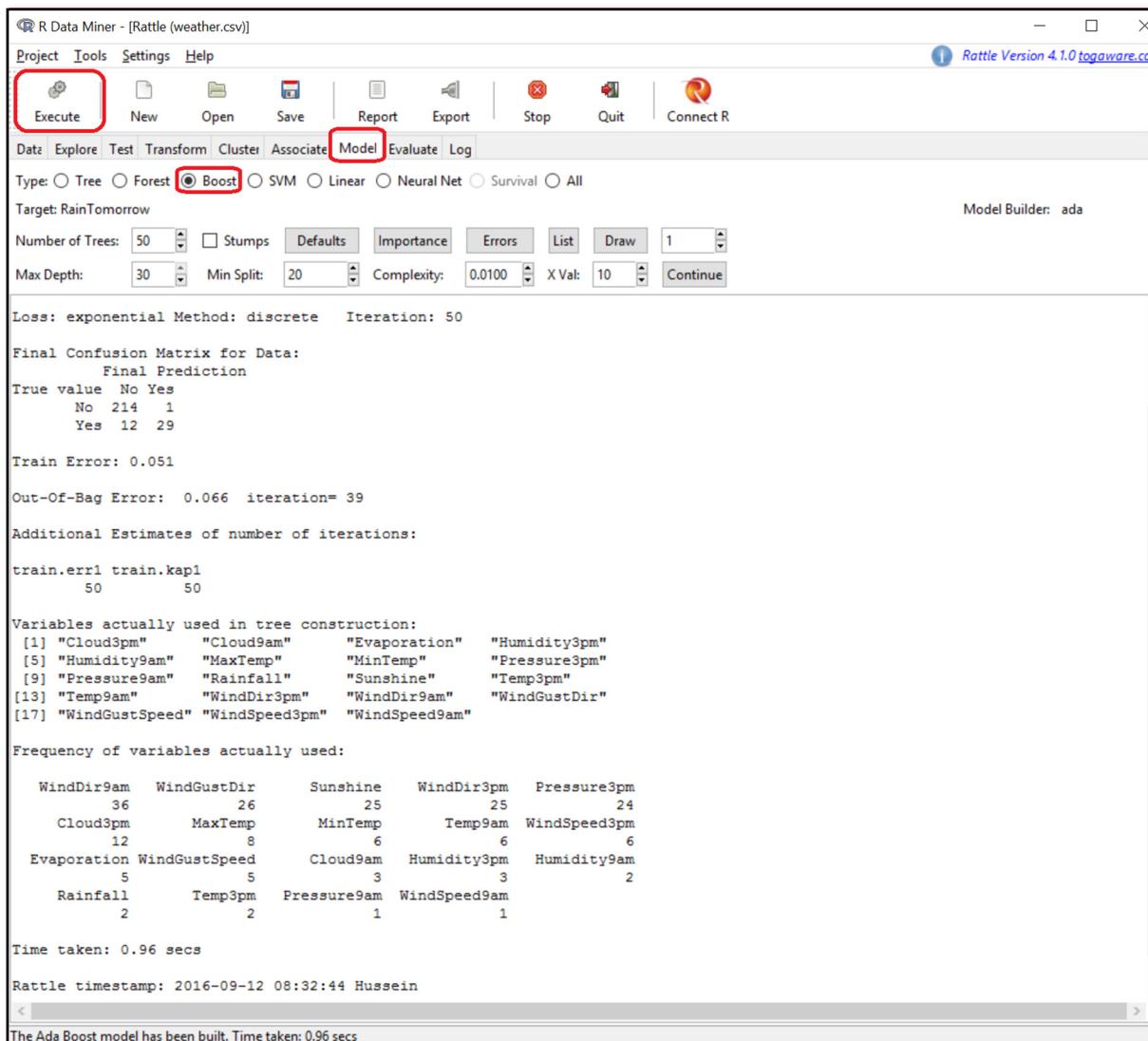


Figure-A III-8 Construction d'un modèle de *boosting* avec Rattle

Dans le Tableau-A III-1, on fait ressortir les poids de chaque arbre dans le modèle final ainsi que l'erreur commise dans la classification.

Tableau-A III-1 Poids et erreur de chaque arbre dans le modèle final

Numéro arbre	Poids	Erreur
1	1.0899914	0.13387978
2	0.8200923	0.13387978
3	0.9586191	0.10382514
4	0.9346704	0.09016393
5	0.9632484	0.06284153
6	0.9730307	0.07923497
7	0.8440946	0.05464481
8	0.9008403	0.05464481
9	0.7619330	0.04918033
10	0.8106903	0.05191257
11	0.8135138	0.04644809
12	0.7550444	0.04644809
13	0.5885324	0.03825137
14	0.8155280	0.04098361
15	0.6885956	0.03825137
16	0.6027676	0.04098361
17	0.6766033	0.05191257
18	0.5602803	0.04644809
19	0.7548848	0.04644809
20	0.7554945	0.04371585
21	0.8128059	0.04644809
22	0.8276871	0.04098361
23	0.6955648	0.04098361
24	0.8712918	0.03825137
25	0.9420133	0.03825137
26	0.7167832	0.03825137
27	0.7156933	0.03825137

Numéro arbre	Poids	Erreur
28	0.8469412	0.04098361
29	0.9132124	0.03825137
30	0.5851181	0.04098361
31	0.7880185	0.04098361
32	0.7590056	0.04098361
33	0.8748962	0.04098361
34	1.0129452	0.04098361
35	0.7040660	0.03551913
36	0.8229786	0.04371585
37	0.9016515	0.04098361
38	0.7342267	0.04918033
39	0.6952263	0.04918033
40	0.7903359	0.04918033
41	0.6599251	0.04918033
42	0.8986569	0.04644809
43	0.7935847	0.04371585
44	0.9451400	0.04644809
45	0.7054423	0.04644809
46	0.7869782	0.04371585
47	0.7462928	0.04371585
48	0.7048098	0.04371585
49	0.8818194	0.04371585
50	0.7926242	0.04371585

Du Tableau-A III-2 au Tableau-A III-5, on affiche la base de données qui est constituée des différentes observations recueillies dans la ville de Canberra en Australie à des dates différentes. Dans la première colonne, on retrouve toujours les dates de collecte des informations.

Tableau-A III-2 Données recueillies à Canberra relatives à la température, la chute de pluie, l'évaporation, l'ensoleillement et la direction des vents

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2007-11-01	8	24.3	0	3.4	6.3	NW
2007-11-02	14	26.9	3.6	4.4	9.7	ENE
2007-11-03	13.7	23.4	3.6	5.8	3.3	NW
2007-11-04	13.3	15.5	39.8	7.2	9.1	NW
2007-11-05	7.6	16.1	2.8	5.6	10.6	SSE
2007-11-06	6.2	16.9	0	5.8	8.2	SE
2007-11-07	6.1	18.2	0.2	4.2	8.4	SE
2007-11-08	8.3	17	0	5.6	4.6	E
2007-11-09	8.8	19.5	0	4	4.1	S
2007-11-10	8.4	22.8	16.2	5.4	7.7	E
2007-11-11	9.1	25.2	0	4.2	11.9	N
2007-11-12	8.5	27.3	0.2	7.2	12.5	E
2007-11-13	10.1	27.9	0	7.2	13	WNW
2007-11-14	12.1	30.9	0	6.2	12.4	NW
2007-11-15	10.1	31.2	0	8.8	13.1	NW
2007-11-16	12.4	32.1	0	8.4	11.1	E
2007-11-17	13.8	31.2	0	7.2	8.4	ESE
2007-11-18	11.7	30	1.2	7.2	10.1	S
2007-11-19	12.4	32.3	0.6	7.4	13	E
2007-11-20	15.6	33.4	0	8	10.4	NE
2007-11-21	15.3	33.4	0	8.8	9.5	WNW
2007-11-22	16.4	19.4	0.4	9.2	0	E
2007-11-23	12.8	18.5	25.8	2.8	0.6	ESE
2007-11-24	12	24.3	0.4	1.2	7.5	NNE
2007-11-25	15.4	28.4	0	4.4	8.1	ENE
2007-11-26	15.6	26.9	0	6.8	8.9	E
2007-11-27	13.3	22.2	0.2	6.6	2.3	ENE

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2007-11-28	12.9	28	0	4.4	10.7	S
2007-11-29	15.1	24.3	0	7	0.4	SE
2007-11-30	13.6	24.1	0.4	2.6	0.5	NNW
2007-12-01	15.1	20.4	22.6	2.4	0.2	SSE
2007-12-02	11.6	26.3	4.2	1.6	12	NNE
2007-12-03	16.6	24.2	0.2	6.6	4.7	NW
2007-12-04	13.3	26.5	6.6	3.8	11.8	NW
2007-12-05	14.5	21.8	0	8.4	9.8	ENE
2007-12-06	16.3	26.8	0	6	6.3	ENE
2007-12-07	17.2	25.8	0	4.2	8.8	SW
2007-12-08	16.5	28.2	4	4.2	8.8	NE
2007-12-09	15	29.4	0	6.6	11.1	NW
2007-12-10	14.9	24.8	0	10.4	10	NNW
2007-12-11	11.8	18.5	0.6	4.8	2.3	ENE
2007-12-12	11.7	21.5	0	4.2	7.3	ENE
2007-12-13	9.6	20.3	0	5	3.6	SE
2007-12-14	8.9	27.1	0	4.4	12.7	E
2007-12-15	10.1	29.9	0	6.8	8.8	E
2007-12-16	15.5	21.1	5.4	6.4	0.9	S
2007-12-17	10.8	21.7	1.4	2.8	10.6	ESE
2007-12-18	7.5	20.9	0	6.6	8.7	ENE
2007-12-19	12.8	21	0	6.4	0.8	NE
2007-12-20	12.6	23.1	3.4	1.6	2.3	NNW
2007-12-21	14.8	29.5	6.4	1.8	8.1	N
2007-12-22	19.9	22	11	4.4	5.9	NNW
2007-12-23	9.2	20.4	17.4	7.8	10.2	ENE
2007-12-24	12.4	24.4	0	6.2	12.1	NW
2007-12-25	11.3	21.7	3.4	8.2	5.6	E

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2007-12-26	9.8	26.3	0	5.2	13	NNW
2007-12-27	14.3	26.7	0	7.2	7.1	NNW
2007-12-28	15.1	28.3	14.4	8.8	13.2	NNW
2007-12-29	14.4	31.6	0	6.6	13.6	NNW
2007-12-30	15.4	35	0	9.6	13	E
2007-12-31	13.8	33.5	0	11.4	13.6	NE
2008-01-01	13.6	34.2	0	8.8	12.8	NNE
2008-01-02	14.3	35	0	7.6	10.5	ESE
2008-01-03	15.9	23.4	0	12.6	2.2	ESE
2008-01-04	16.7	25.3	0	6.2	12.5	ESE
2008-01-05	12.1	27.5	0	7.4	11.7	NE
2008-01-06	14.3	34.1	0	6.6	10.5	ENE
2008-01-07	16.5	33.9	0	9	12.6	ENE
2008-01-08	16.5	30.3	0	10	8.1	E
2008-01-09	17.5	29.9	0	6.6	8.8	E
2008-01-10	14.7	34.2	0	6.6	12.8	NE
2008-01-11	17.5	35.8	0	9.4	13.3	SSE
2008-01-12	20.9	35.7	0	13.8	6.9	SW
2008-01-13	17	33.8	2	9	13.5	WNW
2008-01-14	16	22.8	0	12.4	6	E
2008-01-15	15.4	33.8	0	5.2	11.1	W
2008-01-16	17.9	33.2	0	10.4	8.4	N
2008-01-17	15.2	25.1	4.8	6.4	11.6	E
2008-01-18	15.1	20.4	0	9.6	0.1	ESE
2008-01-19	15.3	19.6	18.8	5	0	NE
2008-01-20	17.2	24.7	12.2	1.4	8.1	NW
2008-01-21	15.9	19.9	0.8	5.6	1.6	E
2008-01-22	10	22.5	0	4.4	11.6	E

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-01-23	9.9	24.4	0	5.8	10.8	NE
2008-01-24	10.3	27.8	0	6.4	9.9	ENE
2008-01-25	15.4	25.7	0	6.8	8.2	E
2008-01-26	12.7	28.8	0	7.4	9	NNE
2008-01-27	13.2	31.3	0	6.6	11.6	WSW
2008-01-28	15.3	33.2	0	9.4	13.2	NNW
2008-01-29	17.9	33.9	0	10.4	11.8	ENE
2008-01-30	18	34.9	0	9.2	9.9	NW
2008-01-31	17.6	27.8	5.2	10.2	3.6	ESE
2008-02-01	16	23.8	2.2	5.4	6.2	SSE
2008-02-02	14.9	28.8	0	5.8	8.1	NNE
2008-02-03	17.1	29.6	0	5.8	9.2	E
2008-02-04	18.2	22.6	1.8	8	0	ENE
2008-02-05	16.8	22.8	9	2.8	0.3	ESE
2008-02-06	13.6	27.4	1	2.8	8	W
2008-02-07	14.5	24.2	0	6.8	5.9	SSW
2008-02-08	12.4	19.9	16.2	5.4	5.6	ENE
2008-02-09	10.4	20.9	0	4	8.9	SSE
2008-02-10	9.1	23.1	0	5.8	9.6	ENE
2008-02-11	8.9	26	0	5	10.7	NE
2008-02-12	14.5	24.2	4.4	6.6	5.9	W
2008-02-13	12.6	18.2	11	3.2	0.4	ENE
2008-02-14	8.6	24.2	0.2	2.8	12.7	E
2008-02-15	10.8	25.2	0	5.6	12.6	ENE
2008-02-16	11.2	26.1	0	7.2	12.6	ENE
2008-02-17	12.1	24.1	0	7.4	10.2	ENE
2008-02-18	10.8	25.8	0	6.8	11.7	E
2008-02-19	11.4	27.1	0	5.6	12.1	NE

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-02-20	12	28.9	0	7.2	8.2	ESE
2008-02-21	16.3	24.8	1.8	7.8	3.8	S
2008-02-22	12.7	28.6	6.6	3.2	8.6	W
2008-02-23	12.7	25.1	0	7.8	12.4	NW
2008-02-24	12	23.8	0	10.4	12.4	NW
2008-02-25	11.5	25.9	0	5.2	10.2	ENE
2008-02-26	13	28.2	0	5	11.6	NW
2008-02-27	11.7	27.6	0	7.8	8.1	W
2008-02-28	14.8	17.3	0	7.6	0.8	SSE
2008-02-29	7.7	18.4	10.4	5.2	12	S
2008-03-01	4.4	21.8	0	6.2	12.1	S
2008-03-02	7.4	24.4	0	6.2	11.6	E
2008-03-03	8.3	27.3	0	5.6	10.1	E
2008-03-04	10.1	28.2	0	6	7.3	ESE
2008-03-05	12	27.6	0	6	11	E
2008-03-06	12.9	31.8	0	6	11.3	WSW
2008-03-07	10.8	29.2	0	8.4	7.5	E
2008-03-08	9.5	27.4	3	4.8	11.5	NNW
2008-03-09	12.1	27.8	0	5	11.5	E
2008-03-10	12.5	31.7	0	6.6	11.2	WNW
2008-03-11	13.9	34.7	0	6.4	8.5	SSW
2008-03-12	13.3	31.7	0.2	7.8	11	WNW
2008-03-13	13.2	33.1	0	8.6	9.7	ENE
2008-03-14	12.3	33.8	0	7.2	11.3	W
2008-03-15	13.8	35.2	0	6.4	11.2	SE
2008-03-16	11.3	32.3	0	9.4	11.4	NE
2008-03-17	11.7	30.2	0	7.8	11.2	NE
2008-03-18	12.5	29.9	0	5.8	10.7	NW

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-03-19	15.1	26.2	0	9	9.8	NE
2008-03-20	11.5	29.3	0	5.2	8	NW
2008-03-21	13	14.8	0	8.2	0	SE
2008-03-22	11.6	19.6	0	2.8	1.6	E
2008-03-23	12.8	24.9	0	2.4	6.2	ENE
2008-03-24	15.5	22.4	0.6	4.8	1.9	NW
2008-03-25	13.1	17.4	6.4	2.8	0	NNW
2008-03-26	12.6	20.2	19.8	2.6	9.1	W
2008-03-27	4.4	18.3	0	4.6	9.6	N
2008-03-28	4.4	18.2	0	3.6	11	W
2008-03-29	7.1	18.5	0	3.4	10	NW
2008-03-30	4.2	18.9	0	6.4	10.8	WNW
2008-03-31	9.6	18.8	0	6.4	10	WNW
2008-04-01	3.5	21.8	0	6	10.3	NNW
2008-04-02	5.3	23.3	0	3.6	5.6	NNW
2008-04-03	7	14.3	2.6	9.6	9.7	WNW
2008-04-04	0.4	18.9	0	4.8	9.6	N
2008-04-05	3.2	21.4	0	3.2	10.6	E
2008-04-06	5.9	21.8	0	2.8	9.3	E
2008-04-07	8.1	20.5	0	3.8	7.8	ENE
2008-04-08	6.9	18.9	0	4.2	4.1	ESE
2008-04-09	5.6	19.5	0	2.8	6.8	SW
2008-04-10	7.2	22.9	0	2.2	9.5	E
2008-04-11	7.1	23.4	0	3.4	10.2	ESE
2008-04-12	6.1	24.1	0	4.6	6.1	WNW
2008-04-13	7.1	19.8	2	3.2	7.7	W
2008-04-14	5.6	18	5.2	3.8	9.3	S
2008-04-15	5.4	20.7	0	3.2	10.8	SSE

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-04-16	6.3	19.3	0	3.4	10.6	S
2008-04-17	5.3	21	0	4.6	6.3	WNW
2008-04-18	7.9	19.7	0	3.2	8.3	ESE
2008-04-19	8.4	16.1	0	3	4.9	SE
2008-04-20	8.1	18.7	0	4.4	7.1	ESE
2008-04-21	2.4	20.6	0	2.8	10.1	E
2008-04-22	5.6	18.9	0	3.8	8	SE
2008-04-23	7.5	19	0	4	6.8	ENE
2008-04-24	2.5	21.2	0	2	7.9	NE
2008-04-25	5	20.9	0	2	8.9	WNW
2008-04-26	3.8	21.7	0.2	2.8	6.5	NW
2008-04-27	7.9	18.7	0	5.8	5.8	NW
2008-04-28	4.3	11.3	7.2	4.4	5.6	W
2008-04-29	-2.1	13.8	0.2	1.8	9.5	NNW
2008-04-30	-1.8	14.8	0	1.4	7	N
2008-05-01	3.8	13.8	0	2.8	0.8	WNW
2008-05-02	2.1	17.3	0	1.6	9.2	W
2008-05-03	0.5	17.1	0	4	9.4	NW
2008-05-04	-0.9	16.7	0	2.4	9.3	NNW
2008-05-05	0.4	19	0	3.4	8.3	NW
2008-05-06	7.5	16.8	0	2.8	3	NW
2008-05-07	8.3	17.6	0	3.4	9.4	WNW
2008-05-08	-0.2	18.1	0	4.4	9.4	NW
2008-05-09	0.1	21	0	2.2	9.2	NNW
2008-05-10	1.5	20.9	0	2.4	9.3	NW
2008-05-11	8.3	17.4	0	2	1.6	E
2008-05-12	9.4	19.2	0	2.2	7.7	NA
2008-05-13	1.3	19	0	2.2	9.4	NNW

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-05-14	2.2	18.6	0	2	9.2	NNW
2008-05-15	-0.4	17.9	0	2.4	8.7	NW
2008-05-16	4.5	16	1.8	3	4.3	NW
2008-05-17	7.9	12.3	1	1.8	1.7	NW
2008-05-18	4.3	14.1	0.8	1.6	8.8	NW
2008-05-19	-2.7	18.1	0	2.2	9.3	W
2008-05-20	0.3	17	0	2.2	5.6	WNW
2008-05-21	3.8	17.4	0	2.6	7.3	E
2008-05-22	2.4	14.7	0	2.6	9.2	ESE
2008-05-23	1.2	14.8	0	2.4	6.7	SE
2008-05-24	1.2	14.5	0	1.2	8.9	SE
2008-05-25	-0.3	17.5	0	1.6	8.4	W
2008-05-26	4.7	18.5	3.8	2	5.1	ESE
2008-05-27	4.9	18.1	5.2	1.2	8.5	NNW
2008-05-28	1.4	16.8	0	1.6	8.2	NW
2008-05-29	2.2	17.5	0.2	1.2	8.5	NW
2008-05-30	-0.1	18	0	2	8.6	NA
2008-05-31	-0.9	18.5	0	2.2	8.9	E
2008-06-01	0.6	14	0	2.2	2.7	NE
2008-06-02	4.6	15.7	0	1	0.6	NNE
2008-06-03	9.8	14.4	0.8	0.6	0	ESE
2008-06-04	10.6	15.1	3.8	0.2	2.6	SSE
2008-06-05	7.8	15.1	0	2.2	2.7	S
2008-06-06	4.4	16.7	0	1.6	9	S
2008-06-07	-0.2	15.5	0	1.4	9	NE
2008-06-08	4.3	14.5	0	2	3	E
2008-06-09	7.4	16.3	0	1.2	NA	NNE
2008-06-10	8.6	13.7	6.2	2.2	0	NW

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-06-11	10.2	15	4.8	0.2	0.5	NW
2008-06-12	7.3	16.4	0.2	1.2	8.5	NW
2008-06-13	8.7	13	0.6	3.2	7.1	WNW
2008-06-14	1	11.8	0	2.2	7.2	S
2008-06-15	2.1	14.7	0	2.6	8.2	S
2008-06-16	6.4	16.9	0	2.6	5.9	SSW
2008-06-17	5.4	15.5	0	1.6	5.8	NNE
2008-06-18	0.4	15.5	0	1	4.7	N
2008-06-19	4	15.9	0	0.6	2	NNW
2008-06-20	8.4	11.7	4.8	0.6	0	W
2008-06-21	0.4	13.9	0.6	0.8	7.7	WNW
2008-06-22	4.2	14	0	1.4	7.6	NW
2008-06-23	0.9	12.9	0	2	6.8	NW
2008-06-24	0.8	13	0	1.4	4.1	WNW
2008-06-25	4.3	12.6	0	2	9	NW
2008-06-26	6.3	11.8	0	2.6	6.5	WNW
2008-06-27	3.5	14.3	0	3.4	8.9	NNW
2008-06-28	-1.5	14.8	0	2.2	8	NW
2008-06-29	1.2	16	0	1.4	6.1	NNW
2008-06-30	0.5	15.4	0	2.4	6.4	W
2008-07-01	5.3	11.7	2	3.6	5.6	NW
2008-07-02	6.6	13.1	0.2	2.8	8.2	WNW
2008-07-03	-1.6	11.5	0	2.8	8.9	N
2008-07-04	-3.1	12	0	1.8	3.9	ESE
2008-07-05	-0.1	14.2	0	1.4	7	W
2008-07-06	-0.6	14	0	1.2	7.1	NW
2008-07-07	3	11.1	0.8	1.4	0.2	W
2008-07-08	2.9	9.5	16.8	1.4	6.5	NW

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-07-09	-1.3	8.8	0	0.8	2.8	WNW
2008-07-10	1.8	8.7	0	1.8	1.2	NW
2008-07-11	2.9	8.4	1.6	1.4	7.7	NW
2008-07-12	-2.6	11.1	0.2	1.4	6.5	N
2008-07-13	0.5	11	0	1	0.9	NNW
2008-07-14	2.7	16.5	0	0.6	8.9	NNW
2008-07-15	-1.7	13.6	0	1.8	5.2	NNW
2008-07-16	-0.9	12.8	0.2	2	1.9	N
2008-07-17	-1.8	11.5	0	0.6	4.7	NNW
2008-07-18	1.3	10.6	0	0.8	5.6	NW
2008-07-19	2.4	11.6	1.2	2.2	8.1	NW
2008-07-20	-1.1	11	0.2	1.8	0	WNW
2008-07-21	2.3	11.6	19.2	1.8	7.5	WNW
2008-07-22	-2.2	11.6	0	1.4	9.2	S
2008-07-23	-3.5	11.2	0	1.6	7.7	ESE
2008-07-24	-1	12.2	0	1.6	8.4	ESE
2008-07-25	-2.1	12.9	0	1.2	8.1	S
2008-07-26	-2	11.3	0.2	2.2	5.9	WNW
2008-07-27	-2.3	9.7	0	1.4	1.9	SSE
2008-07-28	-1.6	10.7	1.4	0.8	9.1	S
2008-07-29	0.8	12.2	0	1.8	8.6	NA
2008-07-30	-2.8	12.2	0	2.6	8.7	NNW
2008-07-31	-2.8	14.1	0	2.2	6.8	WNW
2008-08-01	3	9.7	1	2.6	0.7	W
2008-08-02	4.4	11.5	6.6	2.2	9.3	WNW
2008-08-03	2.3	12.8	0	2.2	9.6	WNW
2008-08-04	-2	12.3	0	2.4	5.8	W
2008-08-05	-1.9	10.9	0	1.8	4.2	S

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-08-06	4.8	14.1	4	1.6	8.3	S
2008-08-07	-0.6	11.1	0	3	5	NNW
2008-08-08	3.1	12.5	1.2	1.4	7.2	N
2008-08-09	-2.9	9.6	0	1.8	7.3	NW
2008-08-10	-3.5	7.6	0.4	2.4	4.7	NW
2008-08-11	-0.3	9.3	0.4	1.4	9.9	N
2008-08-12	0.1	10.4	0	1.8	7.9	WNW
2008-08-13	2.3	12.2	0	2.8	9.8	WNW
2008-08-14	2.1	10.7	0	3.4	9.4	NW
2008-08-15	4.6	14.7	0	4.4	8.4	WNW
2008-08-16	3.7	14.2	0	3	10	NNW
2008-08-17	-1.3	11.6	0	4	10.4	N
2008-08-18	-3.4	12.5	0	3	6.8	SSE
2008-08-19	-5.3	13.1	0	2.2	7.9	NW
2008-08-20	0	14	0	2.4	4.7	WNW
2008-08-21	2.4	14.1	0	3	1.8	NW
2008-08-22	-0.6	12.2	0	2.6	7	S
2008-08-23	2.3	11.6	0	5.4	9.5	S
2008-08-24	-3.7	14.4	0	2.6	10.4	NNW
2008-08-25	-0.9	14.2	0	2.6	7.8	WNW
2008-08-26	-1.5	17.3	0	2.8	9	NW
2008-08-27	-3.3	15.1	0	3	NA	SSW
2008-08-28	-0.1	14.7	0	3.4	9.9	ENE
2008-08-29	-0.2	16.2	0	3.4	5.9	E
2008-08-30	0.5	16.3	0	1.8	4.1	NNW
2008-08-31	6.1	17.2	4	2.2	2.4	NW
2008-09-01	4.1	14.8	7.4	2.6	10.8	NW
2008-09-02	0.1	16.7	0	3.8	10.2	WNW

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-09-03	3.2	12.1	0	2.8	3.9	ESE
2008-09-04	5.4	11.3	0.2	2.2	0.6	ENE
2008-09-05	5.8	12.4	0	1.6	0	SE
2008-09-06	6.3	16.1	0	1.8	2.9	S
2008-09-07	-0.9	16.7	0	2.8	8.6	NW
2008-09-08	0.2	15.5	1	2.4	9.4	SE
2008-09-09	-3.7	14.7	0	3.4	10.9	SSE
2008-09-10	-2.7	15.2	0	4	9.7	N
2008-09-11	-2.5	16.6	0	3	9.9	NW
2008-09-12	-0.5	21.6	0	5	9.9	NW
2008-09-13	9	25.5	0	5.6	10.2	NW
2008-09-14	13.1	19.4	9.8	8.8	6	NNW
2008-09-15	8.7	19.7	1.6	5.2	8	NW
2008-09-16	3.9	13.2	3.4	6.6	11	WNW
2008-09-17	0.7	14.1	0	5.6	9	ENE
2008-09-18	1.1	18	0	1.6	8.6	NNW
2008-09-19	5.1	23.3	0	3.6	10.3	NW
2008-09-20	7.5	23.3	0	6.8	10.9	WNW
2008-09-21	4.7	19.5	0	10	11	NNW
2008-09-22	3.2	21.9	0	6.8	5.2	NW
2008-09-23	7.8	16.2	17.4	6.4	7.9	NW
2008-09-24	2.4	17.3	0	2.2	11.3	E
2008-09-25	3.2	18.7	0	2.6	11.1	N
2008-09-26	2.5	20.9	0	3.6	10.6	NNW
2008-09-27	6.5	25.7	0	4.8	10.5	NNW
2008-09-28	14.4	24.3	0	9.4	11.1	WNW
2008-09-29	4.9	18.9	0	9.6	9.4	NW
2008-09-30	2.3	16.8	0	4.8	11.4	N

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-10-01	1.4	20.6	0	5.4	11.1	N
2008-10-02	5.6	27.6	0	5.2	11	NW
2008-10-03	16.8	28.9	0	10	10.8	NNW
2008-10-04	14.4	20.7	7.6	9.4	4.9	NNW
2008-10-05	10.3	21.3	3	4.2	6.7	NNW
2008-10-06	11.2	18	0	4.8	8.4	W
2008-10-07	0.3	16	8.2	5.4	11.8	NW
2008-10-08	0.5	17.9	0	5.8	11.5	N
2008-10-09	0.5	20	0	6.2	11.5	NNW
2008-10-10	4.6	22	0	4.4	11	N
2008-10-11	8.2	22.4	0	5.4	11.2	NW
2008-10-12	4.5	23.9	0	4.8	11.7	NW
2008-10-13	6.7	26.1	0	6.2	7.5	SSW
2008-10-14	11.9	21.1	13.2	6.6	NA	NW
2008-10-15	9.2	19.6	0.6	3.4	10.4	ENE
2008-10-16	4.4	21	0	4.2	12.2	NW
2008-10-17	5	24.1	0	6.2	12	NNW
2008-10-18	6.7	24.7	0	5.4	8.6	NW
2008-10-19	8.3	28.5	0	5.8	9.8	NW
2008-10-20	11.3	27.4	0.2	7.6	12.1	NW
2008-10-21	9	20.6	0	9	6.2	ENE
2008-10-22	3.4	15	0.8	4.8	11.7	S
2008-10-23	3.2	18	0	7.4	12.2	SSE
2008-10-24	0.9	20.7	0	5.4	8.4	NNW
2008-10-25	3.3	25.5	0	5.2	10.8	N
2008-10-26	7.9	26.1	0	6.8	3.5	NNW
2008-10-27	9	30.7	0	7.6	12.1	NNW
2008-10-28	7.1	28.4	0	11.6	12.7	N

Date	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir
2008-10-29	12.5	19.9	0	8.4	5.3	ESE
2008-10-30	12.5	26.9	0	5	7.1	NW
2008-10-31	12.3	30.2	0	6	12.6	NW

Tableau-A III-3 Données recueillies à Canberra relatives à la direction et à la vitesse des vents à des heures précises

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2007-11-01	30	SW	NW	6	20
2007-11-02	39	E	W	4	17
2007-11-03	85	N	NNE	6	6
2007-11-04	54	WNW	W	30	24
2007-11-05	50	SSE	ESE	20	28
2007-11-06	44	SE	E	20	24
2007-11-07	43	SE	ESE	19	26
2007-11-08	41	SE	E	11	24
2007-11-09	48	E	ENE	19	17
2007-11-10	31	S	ESE	7	6
2007-11-11	30	SE	NW	6	9
2007-11-12	41	E	NW	2	15
2007-11-13	30	S	NW	6	7
2007-11-14	44	WNW	W	7	20
2007-11-15	41	S	W	6	20
2007-11-16	46	SE	WSW	7	9
2007-11-17	44	WSW	W	6	19
2007-11-18	52	SW	NE	6	11
2007-11-19	39	NNE	W	4	17
2007-11-20	33	NNW	NNW	2	13

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2007-11-21	59	N	NW	2	31
2007-11-22	26	ENE	E	6	11
2007-11-23	28	S	SE	13	13
2007-11-24	26	WSW	NE	6	9
2007-11-25	33	SSE	NE	9	15
2007-11-26	41	E	E	6	22
2007-11-27	39	E	E	20	17
2007-11-28	52	S	NNE	6	11
2007-11-29	39	SE	SE	7	17
2007-11-30	30	SSW	S	6	6
2007-12-01	41	E	S	6	20
2007-12-02	41	E	SW	6	9
2007-12-03	50	WNW	NW	13	35
2007-12-04	50	NW	WNW	20	26
2007-12-05	43	ESE	E	11	30
2007-12-06	39	ESE	ESE	13	9
2007-12-07	41	NW	N	6	6
2007-12-08	39	E	N	7	13
2007-12-09	43	N	W	9	24
2007-12-10	35	NNE	WNW	11	6
2007-12-11	35	ESE	E	9	15
2007-12-12	41	ESE	E	15	24
2007-12-13	39	ESE	E	22	20
2007-12-14	35	NNW	N	6	11
2007-12-15	41	SE	WNW	2	13
2007-12-16	31	SSE	NE	6	7
2007-12-17	48	SSE	ESE	13	20

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2007-12-18	39	SE	E	13	20
2007-12-19	22	NE	ENE	7	6
2007-12-20	30	N	NW	4	13
2007-12-21	41	NW	NW	6	22
2007-12-22	76	N	WNW	41	30
2007-12-23	39	N	N	9	20
2007-12-24	44	NNW	WNW	7	13
2007-12-25	41	SE	NE	11	11
2007-12-26	41	W	NNW	6	19
2007-12-27	65	N	NNW	7	48
2007-12-28	28	NNW	NW	6	13
2007-12-29	30	NW	N	6	7
2007-12-30	39	SSW	ESE	6	22
2007-12-31	31	SSE	NE	7	9
2008-01-01	35	ESE	W	2	11
2008-01-02	41	ESE	WSW	2	9
2008-01-03	50	ESE	ESE	20	24
2008-01-04	46	ESE	ESE	24	26
2008-01-05	35	SSE	E	7	15
2008-01-06	39	W	NNW	6	19
2008-01-07	39	E	NW	11	11
2008-01-08	46	E	N	7	9
2008-01-09	43	E	ENE	13	11
2008-01-10	33	WNW	NW	6	11
2008-01-11	57	NNW	NW	6	19
2008-01-12	50	E	WNW	4	17
2008-01-13	52	NNW	W	6	24

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-01-14	50	E	ENE	13	19
2008-01-15	35	E	NW	7	15
2008-01-16	59	N	NNE	15	20
2008-01-17	46	ESE	ESE	20	28
2008-01-18	39	ESE	ESE	17	19
2008-01-19	33	SSE	NNE	9	13
2008-01-20	50	NW	WNW	19	28
2008-01-21	48	SSE	ESE	17	22
2008-01-22	33	SE	S	11	7
2008-01-23	28	SE	E	7	7
2008-01-24	35	SE	NE	6	11
2008-01-25	41	E	NE	9	17
2008-01-26	28	NW	SSW	4	7
2008-01-27	46	N	WNW	4	26
2008-01-28	44	NNE	NW	4	24
2008-01-29	46	S	W	6	11
2008-01-30	69	N	W	6	13
2008-01-31	39	N	NNW	13	13
2008-02-01	30	SE	ESE	11	11
2008-02-02	30	SSE	NNE	7	11
2008-02-03	48	SE	ESE	9	17
2008-02-04	33	SSE	ENE	7	13
2008-02-05	30	S	SSE	7	11
2008-02-06	52	NW	WNW	6	24
2008-02-07	61	N	NNW	11	20
2008-02-08	41	ESE	ESE	7	20
2008-02-09	33	S	S	15	17

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-02-10	41	SSE	W	17	13
2008-02-11	31	SE	NE	7	13
2008-02-12	48	NNW	WNW	9	17
2008-02-13	30	SSE	SSE	13	15
2008-02-14	33	S	SE	11	9
2008-02-15	35	SE	E	7	7
2008-02-16	39	SE	E	7	19
2008-02-17	46	ESE	NNE	9	13
2008-02-18	31	S	NNE	7	11
2008-02-19	33	ESE	SE	6	15
2008-02-20	39	SSE	WNW	2	9
2008-02-21	50	NNW	N	4	7
2008-02-22	50	NA	W	NA	28
2008-02-23	46	WNW	WNW	20	26
2008-02-24	44	NW	WNW	15	30
2008-02-25	44	ENE	WSW	9	7
2008-02-26	44	WSW	NW	4	17
2008-02-27	48	NA	W	NA	30
2008-02-28	48	S	SE	13	13
2008-02-29	48	S	S	24	17
2008-03-01	35	SSE	SSE	15	15
2008-03-02	35	ESE	WNW	7	11
2008-03-03	35	ESE	WNW	6	9
2008-03-04	39	SSE	WNW	6	7
2008-03-05	46	SSE	WSW	7	6
2008-03-06	41	NNW	NNE	6	9
2008-03-07	50	NA	ENE	NA	9

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-03-08	24	SSE	NNW	6	9
2008-03-09	41	SE	NNE	7	11
2008-03-10	24	ENE	W	6	9
2008-03-11	46	SE	ENE	11	7
2008-03-12	44	ESE	WSW	6	28
2008-03-13	39	SE	NNW	7	7
2008-03-14	22	SE	W	9	13
2008-03-15	48	SE	ESE	7	4
2008-03-16	28	ENE	WNW	4	6
2008-03-17	33	ESE	NNW	9	13
2008-03-18	43	NA	WNW	0	19
2008-03-19	31	SE	NNE	4	15
2008-03-20	46	NA	WSW	0	26
2008-03-21	30	ESE	ESE	9	17
2008-03-22	33	SE	ESE	6	17
2008-03-23	30	NNW	SW	6	7
2008-03-24	28	SSE	WNW	6	19
2008-03-25	43	N	N	11	20
2008-03-26	46	NNW	W	17	20
2008-03-27	33	SE	N	7	17
2008-03-28	31	SE	WNW	7	19
2008-03-29	57	N	NW	9	22
2008-03-30	50	WSW	WNW	6	31
2008-03-31	57	NW	W	19	26
2008-04-01	22	ESE	WNW	7	11
2008-04-02	83	SSW	NW	2	37
2008-04-03	63	NW	W	26	33

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-04-04	22	NW	WSW	6	7
2008-04-05	26	SE	NNW	7	7
2008-04-06	35	SSW	NNW	7	11
2008-04-07	31	ESE	ENE	11	11
2008-04-08	39	SSE	SE	15	9
2008-04-09	17	S	WNW	9	6
2008-04-10	26	NNW	N	4	6
2008-04-11	39	SE	WSW	2	7
2008-04-12	35	SSW	WNW	6	11
2008-04-13	39	SSE	NW	13	15
2008-04-14	31	SSE	SSE	9	20
2008-04-15	28	SSE	S	7	9
2008-04-16	31	S	SE	9	13
2008-04-17	26	SSW	NW	7	7
2008-04-18	48	SSE	ENE	17	11
2008-04-19	35	SE	ESE	13	22
2008-04-20	33	SSE	ESE	11	30
2008-04-21	30	SE	SSE	9	7
2008-04-22	33	SSE	E	13	22
2008-04-23	26	NA	ESE	0	7
2008-04-24	28	ESE	NE	7	9
2008-04-25	22	NA	NW	0	7
2008-04-26	44	ESE	WNW	2	20
2008-04-27	59	NNW	NW	26	31
2008-04-28	57	WNW	WNW	26	22
2008-04-29	22	NA	NNW	0	9
2008-04-30	28	E	N	2	19

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-05-01	31	W	NNW	6	26
2008-05-02	43	SSW	WNW	6	28
2008-05-03	31	ESE	W	6	13
2008-05-04	30	SW	NNW	2	15
2008-05-05	39	NE	WNW	2	19
2008-05-06	41	W	NW	7	26
2008-05-07	43	NW	WNW	17	30
2008-05-08	24	NA	NW	0	9
2008-05-09	17	WNW	N	2	9
2008-05-10	20	NW	NNW	2	9
2008-05-11	20	WSW	NE	6	11
2008-05-12	24	E	NNW	4	15
2008-05-13	30	SE	NW	7	15
2008-05-14	31	NA	NNW	NA	20
2008-05-15	33	NNE	NW	2	20
2008-05-16	35	W	NW	9	15
2008-05-17	52	NNW	NW	30	24
2008-05-18	41	NNW	WNW	13	26
2008-05-19	52	NA	WNW	0	24
2008-05-20	31	S	WNW	6	24
2008-05-21	24	ESE	W	2	7
2008-05-22	20	SSE	E	7	6
2008-05-23	17	SE	SSW	11	6
2008-05-24	17	SE	SSW	6	6
2008-05-25	39	SE	NNW	6	15
2008-05-26	22	NNW	W	6	6
2008-05-27	22	NA	NW	0	13

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-05-28	20	NA	NW	0	6
2008-05-29	20	SSE	WNW	7	7
2008-05-30	NA	N	NNW	2	15
2008-05-31	22	NA	N	0	7
2008-06-01	17	NA	ENE	NA	7
2008-06-02	15	SSE	NNE	6	6
2008-06-03	20	S	ESE	9	11
2008-06-04	41	SE	SE	9	20
2008-06-05	31	SSE	S	6	20
2008-06-06	22	SE	NNE	11	4
2008-06-07	28	N	ESE	4	11
2008-06-08	30	SSE	ENE	6	13
2008-06-09	30	SSE	NE	6	19
2008-06-10	31	E	NA	2	0
2008-06-11	46	NNW	NW	22	17
2008-06-12	43	N	WNW	15	24
2008-06-13	54	W	WNW	19	28
2008-06-14	48	NW	S	7	24
2008-06-15	43	S	SSW	24	26
2008-06-16	26	SSW	ESE	17	9
2008-06-17	20	SSW	NNE	7	7
2008-06-18	13	N	NNW	6	9
2008-06-19	17	SE	SSE	6	7
2008-06-20	26	WNW	WNW	6	7
2008-06-21	41	E	WNW	6	19
2008-06-22	33	NNW	SE	9	9
2008-06-23	24	W	N	7	15

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-06-24	41	S	NW	2	28
2008-06-25	46	NA	WNW	0	28
2008-06-26	63	NNW	WNW	20	28
2008-06-27	35	N	NNW	7	24
2008-06-28	15	N	NW	6	9
2008-06-29	31	S	N	6	22
2008-06-30	70	NW	NNW	22	28
2008-07-01	78	NW	W	33	33
2008-07-02	61	NNW	WNW	31	30
2008-07-03	31	ESE	NNW	7	20
2008-07-04	35	SSW	NE	4	13
2008-07-05	13	NA	WNW	NA	2
2008-07-06	41	WNW	NNW	2	22
2008-07-07	35	NA	N	0	13
2008-07-08	35	NA	NW	0	15
2008-07-09	39	NW	NW	9	15
2008-07-10	65	NW	NW	31	35
2008-07-11	59	NW	WNW	20	31
2008-07-12	20	E	NNW	2	7
2008-07-13	31	NA	NW	0	15
2008-07-14	41	SW	NNW	7	22
2008-07-15	28	NA	NNW	NA	17
2008-07-16	17	SE	NNW	6	9
2008-07-17	41	NA	NNW	0	24
2008-07-18	46	NNW	WNW	24	33
2008-07-19	35	NNW	NW	15	22
2008-07-20	41	NA	NE	0	6

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-07-21	54	WNW	WNW	26	31
2008-07-22	39	E	SSW	2	28
2008-07-23	26	N	ESE	6	11
2008-07-24	30	SE	E	7	13
2008-07-25	31	SE	SE	7	13
2008-07-26	33	SE	WNW	2	19
2008-07-27	28	SSE	WSW	6	13
2008-07-28	59	SSW	S	28	37
2008-07-29	NA	S	S	19	22
2008-07-30	31	NA	NNW	0	15
2008-07-31	48	NA	WNW	0	26
2008-08-01	65	NW	NW	19	28
2008-08-02	57	NNW	WNW	28	31
2008-08-03	35	ENE	WNW	4	24
2008-08-04	31	ESE	W	4	19
2008-08-05	50	NA	SSW	0	26
2008-08-06	48	S	SSE	30	22
2008-08-07	41	SE	N	7	26
2008-08-08	43	WNW	N	6	30
2008-08-09	43	N	NW	6	24
2008-08-10	50	NW	NW	30	24
2008-08-11	39	NNW	N	9	30
2008-08-12	59	NW	WNW	22	24
2008-08-13	44	WNW	NW	20	30
2008-08-14	59	NNW	WNW	28	39
2008-08-15	52	WNW	NW	28	33
2008-08-16	46	WNW	WSW	24	28

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-08-17	30	ESE	NNW	9	17
2008-08-18	48	NNW	WNW	7	30
2008-08-19	33	N	NNW	4	20
2008-08-20	39	NNW	WNW	13	24
2008-08-21	46	NNW	NW	15	30
2008-08-22	59	S	S	15	26
2008-08-23	54	SSW	ESE	30	17
2008-08-24	22	N	NNW	2	11
2008-08-25	31	SSW	N	7	19
2008-08-26	48	NA	WNW	0	26
2008-08-27	30	SSE	SE	9	13
2008-08-28	30	SE	NE	6	15
2008-08-29	28	NA	ENE	0	20
2008-08-30	39	SSW	WSW	4	6
2008-08-31	59	NW	WNW	24	15
2008-09-01	46	NW	W	20	28
2008-09-02	28	ESE	NNE	2	13
2008-09-03	41	SE	SE	17	20
2008-09-04	35	SE	ESE	11	17
2008-09-05	26	SSE	ESE	13	13
2008-09-06	35	S	S	19	24
2008-09-07	35	WNW	NW	7	22
2008-09-08	31	NA	WNW	0	11
2008-09-09	43	SE	N	9	22
2008-09-10	30	SW	WNW	6	20
2008-09-11	31	NE	N	4	20
2008-09-12	48	NNE	NW	9	26

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-09-13	50	N	NW	20	26
2008-09-14	67	N	NW	39	30
2008-09-15	98	NW	NNW	31	52
2008-09-16	65	WNW	WNW	26	31
2008-09-17	20	SSW	NNW	6	7
2008-09-18	39	NNW	NNW	2	28
2008-09-19	44	N	NW	9	28
2008-09-20	65	N	NW	15	41
2008-09-21	46	NNE	NNW	13	33
2008-09-22	80	SSE	NW	2	20
2008-09-23	50	NW	N	15	22
2008-09-24	31	SSE	SW	4	6
2008-09-25	31	E	NNW	4	13
2008-09-26	44	NW	WNW	20	26
2008-09-27	50	NNW	NW	2	24
2008-09-28	52	NW	NW	31	30
2008-09-29	44	S	NW	6	22
2008-09-30	41	NNW	NNW	6	28
2008-10-01	46	SW	N	6	30
2008-10-02	46	NNW	NW	15	28
2008-10-03	70	NW	NW	31	41
2008-10-04	33	NNW	NNW	20	19
2008-10-05	43	ENE	N	7	19
2008-10-06	65	N	W	24	33
2008-10-07	57	NNW	N	11	28
2008-10-08	44	NNE	NNW	2	26
2008-10-09	31	S	N	2	9

Date	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm
2008-10-10	41	NNW	N	6	19
2008-10-11	31	SSW	NW	2	17
2008-10-12	30	ENE	NNW	4	11
2008-10-13	70	NE	NNW	6	33
2008-10-14	41	NNE	N	7	24
2008-10-15	31	SSE	NNW	4	7
2008-10-16	28	SW	NW	2	9
2008-10-17	52	NA	NNW	0	26
2008-10-18	43	N	NW	4	20
2008-10-19	46	W	NW	2	26
2008-10-20	52	SE	NW	6	31
2008-10-21	39	S	SW	11	11
2008-10-22	70	S	S	35	37
2008-10-23	48	SSE	S	26	15
2008-10-24	39	SSE	N	2	17
2008-10-25	43	N	NNW	4	19
2008-10-26	43	NA	WNW	0	19
2008-10-27	76	SSE	NW	7	50
2008-10-28	48	NNW	NNW	2	19
2008-10-29	43	ENE	ENE	11	9
2008-10-30	46	SSW	WNW	6	28
2008-10-31	78	NW	WNW	31	35

Tableau-A III-4 Données recueillies à Canberra relatives à l'humidité, la pression et l'ennuagement à des heures précises

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud 3pm
2007-11-01	68	29	1019.7	1015	7	7
2007-11-02	80	36	1012.4	1008.4	5	3
2007-11-03	82	69	1009.5	1007.2	8	7
2007-11-04	62	56	1005.5	1007	2	7
2007-11-05	68	49	1018.3	1018.5	7	7
2007-11-06	70	57	1023.8	1021.7	7	5
2007-11-07	63	47	1024.6	1022.2	4	6
2007-11-08	65	57	1026.2	1024.2	6	7
2007-11-09	70	48	1026.1	1022.7	7	7
2007-11-10	82	32	1024.1	1020.7	7	1
2007-11-11	74	34	1024.4	1021.1	1	2
2007-11-12	54	35	1023.8	1019.9	0	3
2007-11-13	62	29	1022	1017.1	0	1
2007-11-14	67	20	1017.3	1013.1	1	4
2007-11-15	45	16	1018.2	1013.7	0	1
2007-11-16	70	22	1017.9	1012.8	0	3
2007-11-17	72	23	1014.4	1009.8	7	6
2007-11-18	59	26	1016.4	1013	1	5
2007-11-19	60	25	1017.1	1013.3	1	3
2007-11-20	61	27	1018.5	1013.7	0	1
2007-11-21	60	26	1012.4	1006.5	1	5
2007-11-22	88	72	1010.7	1008.9	8	8
2007-11-23	91	79	1014	1014.9	8	8
2007-11-24	74	57	1020.7	1019.2	7	5
2007-11-25	85	31	1022.4	1018.6	8	2
2007-11-26	65	48	1019.7	1016.5	2	4

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2007-11-27	70	55	1021	1018.6	7	7
2007-11-28	61	31	1019.2	1014.8	5	7
2007-11-29	69	80	1019	1017.1	7	7
2007-11-30	83	49	1017.2	1013.3	8	7
2007-12-01	82	90	1015	1014.1	8	8
2007-12-02	66	46	1013.9	1009.5	2	2
2007-12-03	72	60	1007.9	1004.6	7	7
2007-12-04	74	40	1007.3	1006.3	5	2
2007-12-05	69	63	1015	1015.3	7	1
2007-12-06	78	54	1018.8	1016.1	8	7
2007-12-07	70	74	1014.5	1011.5	6	7
2007-12-08	76	42	1014.8	1011.4	8	1
2007-12-09	78	24	1010.8	1008.5	7	7
2007-12-10	58	37	1011.7	1010.4	7	7
2007-12-11	60	63	1018	1017.4	8	6
2007-12-12	66	51	1021.1	1019.6	7	5
2007-12-13	70	49	1021.9	1019.3	7	7
2007-12-14	63	37	1018.1	1012.5	1	2
2007-12-15	63	36	1012.6	1010.2	1	8
2007-12-16	93	86	1010.1	1008.6	8	8
2007-12-17	59	40	1017.6	1016.8	1	4
2007-12-18	61	43	1020.6	1018.3	6	7
2007-12-19	72	54	1018	1015.6	8	8
2007-12-20	97	74	1015.8	1014.1	8	7
2007-12-21	81	48	1012.9	1008.3	2	7
2007-12-22	92	62	996.5	996.8	8	3
2007-12-23	59	49	1009.5	1009.1	1	7

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2007-12-24	69	43	1012.8	1009.7	6	4
2007-12-25	67	52	1020.4	1019.1	8	3
2007-12-26	63	35	1023.2	1018	1	1
2007-12-27	68	64	1017.3	1014.3	7	7
2007-12-28	73	44	1016.8	1013.4	1	5
2007-12-29	65	30	1017.1	1014.2	0	2
2007-12-30	55	18	1017.4	1015	1	1
2007-12-31	44	20	1019.2	1015.6	0	1
2008-01-01	50	20	1015.8	1011.6	7	6
2008-01-02	43	14	1012.6	1008.7	0	1
2008-01-03	69	55	1016.8	1016.3	8	7
2008-01-04	56	45	1016.2	1013	6	7
2008-01-05	69	40	1011.8	1007.4	7	2
2008-01-06	77	34	1006.6	1003.3	1	1
2008-01-07	60	36	1012.3	1009.5	1	1
2008-01-08	77	44	1015.7	1013.4	7	1
2008-01-09	74	45	1018.5	1015.9	8	1
2008-01-10	64	25	1017.8	1013.3	1	1
2008-01-11	56	24	1011.8	1006.5	0	1
2008-01-12	61	28	1007.6	1003	7	2
2008-01-13	65	35	1004	1001.8	2	3
2008-01-14	70	57	1012.3	1012.3	8	6
2008-01-15	70	26	1014.2	1009	7	1
2008-01-16	58	62	1008.5	1006.1	6	7
2008-01-17	62	42	1016.8	1015	1	1
2008-01-18	63	55	1017.5	1015.8	8	8
2008-01-19	92	88	1008.7	1006	8	8

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-01-20	80	71	1004.9	1004	7	8
2008-01-21	73	58	1011.4	1012.1	8	7
2008-01-22	55	46	1018.7	1016.8	1	2
2008-01-23	60	38	1021.6	1019.2	1	4
2008-01-24	64	29	1021.3	1018.1	3	6
2008-01-25	67	32	1020.1	1016.9	7	7
2008-01-26	70	28	1016.5	1012.4	1	5
2008-01-27	71	28	1013.1	1009.5	1	4
2008-01-28	60	18	1013.9	1012	0	1
2008-01-29	72	22	1017.4	1014.6	1	1
2008-01-30	68	29	1018	1013.7	1	6
2008-01-31	66	49	1014.6	1012.8	7	8
2008-02-01	73	49	1018.9	1017.1	8	7
2008-02-02	71	44	1018.2	1014.5	7	2
2008-02-03	67	38	1016.7	1012.8	6	7
2008-02-04	92	76	1014.4	1011.5	8	8
2008-02-05	88	71	1010	1007.8	8	8
2008-02-06	77	41	1002.1	997.5	4	7
2008-02-07	76	76	999.4	998.9	7	7
2008-02-08	75	58	1007.6	1005	7	7
2008-02-09	65	51	1007.7	1006.8	6	6
2008-02-10	66	37	1013.6	1011.4	1	7
2008-02-11	74	34	1016.1	1011.7	1	5
2008-02-12	95	60	1010.8	1006.8	8	5
2008-02-13	81	73	1010	1011	7	8
2008-02-14	63	37	1017.5	1015.1	1	1
2008-02-15	69	32	1020.2	1016.7	1	1

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-02-16	71	40	1022.9	1020	1	1
2008-02-17	64	46	1025.2	1021.4	7	1
2008-02-18	72	42	1021.3	1017.4	5	4
2008-02-19	69	38	1017.6	1012.8	1	2
2008-02-20	78	31	1012.2	1009.1	2	7
2008-02-21	88	56	1011.8	1008.4	7	7
2008-02-22	78	42	1008.3	1002.3	5	5
2008-02-23	57	32	1007.4	1005.4	1	1
2008-02-24	54	34	1007.9	1007.2	0	1
2008-02-25	70	35	1016.4	1014.3	6	2
2008-02-26	74	27	1017.6	1014.4	3	6
2008-02-27	74	25	1013.6	1009.1	2	6
2008-02-28	74	78	1010	1009.9	7	7
2008-02-29	58	34	1015.5	1014.5	5	1
2008-03-01	60	31	1020.8	1019.3	1	1
2008-03-02	65	37	1026.3	1023.2	0	2
2008-03-03	78	28	1025.4	1020.4	1	1
2008-03-04	72	28	1021.1	1018.2	1	7
2008-03-05	69	35	1025.5	1022.2	1	1
2008-03-06	76	24	1023.9	1019.3	2	3
2008-03-07	56	34	1021.6	1017.7	0	7
2008-03-08	79	39	1024.7	1020.7	1	7
2008-03-09	66	33	1024.8	1023.1	0	1
2008-03-10	74	28	1024.9	1020.4	1	1
2008-03-11	50	18	1022.8	1019.3	0	3
2008-03-12	80	23	1021.6	1017.6	2	5
2008-03-13	91	18	1020.9	1017.7	7	1

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-03-14	57	15	1022.6	1019.1	0	1
2008-03-15	47	15	1022.5	1019.3	5	2
2008-03-16	44	17	1024.3	1020.7	5	2
2008-03-17	73	22	1024.3	1020.6	1	0
2008-03-18	76	30	1022.7	1018.8	0	1
2008-03-19	59	48	1022.1	1018.7	1	4
2008-03-20	76	35	1013.8	1009.4	7	7
2008-03-21	77	62	1017.4	1019.3	7	8
2008-03-22	71	64	1023.6	1021.8	7	8
2008-03-23	76	44	1023.2	1019.3	7	6
2008-03-24	89	68	1018.7	1015.4	8	7
2008-03-25	94	93	1010.2	1006.5	8	8
2008-03-26	76	50	1007.9	1008.4	5	4
2008-03-27	80	36	1014.8	1012.6	1	2
2008-03-28	78	35	1017.4	1015	1	3
2008-03-29	90	30	1015.1	1011.3	6	6
2008-03-30	60	34	1011.2	1009.3	3	2
2008-03-31	54	42	1012.7	1013	2	5
2008-04-01	72	34	1023.9	1021.2	0	1
2008-04-02	90	26	1018	1010.5	1	8
2008-04-03	68	40	1010.3	1012.7	1	3
2008-04-04	77	33	1021.2	1019.1	0	1
2008-04-05	76	27	1025.2	1022.3	0	1
2008-04-06	70	35	1028	1024.1	1	5
2008-04-07	66	48	1027.8	1024	1	5
2008-04-08	68	45	1028.2	1025.4	5	4
2008-04-09	69	49	1025.7	1021.8	6	4

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-04-10	82	28	1023.1	1019.2	1	1
2008-04-11	72	29	1021	1016.9	0	1
2008-04-12	75	32	1019.2	1015.8	7	3
2008-04-13	96	56	1018.4	1013.8	5	2
2008-04-14	70	40	1020.8	1019.6	1	1
2008-04-15	66	36	1027.2	1023.9	1	2
2008-04-16	65	39	1027.9	1024.4	1	1
2008-04-17	77	44	1027.2	1023.3	7	2
2008-04-18	58	43	1027.3	1024.1	3	6
2008-04-19	69	51	1029.5	1027.4	6	7
2008-04-20	66	65	1028.2	1024.2	4	7
2008-04-21	77	34	1024.8	1021	1	2
2008-04-22	79	49	1025.8	1023.5	7	6
2008-04-23	74	43	1025.5	1022.2	6	5
2008-04-24	82	55	1023.7	1019.8	4	4
2008-04-25	84	42	1019.9	1014.3	7	6
2008-04-26	99	34	1011.9	1006.7	7	7
2008-04-27	68	51	1003.2	997.7	6	6
2008-04-28	75	60	1006.3	1008.1	3	6
2008-04-29	69	40	1020.6	1019.6	0	1
2008-04-30	80	40	1024.2	1020.5	1	7
2008-05-01	67	54	1020.6	1018.5	7	7
2008-05-02	81	38	1020	1017.3	1	2
2008-05-03	74	42	1020.8	1017.4	1	1
2008-05-04	76	42	1022.7	1018.5	5	2
2008-05-05	76	41	1019.8	1015.8	6	5
2008-05-06	70	53	1018	1013.8	7	7

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-05-07	73	43	1015.8	1013.5	1	1
2008-05-08	80	44	1021.4	1018.9	1	1
2008-05-09	78	36	1023.2	1020.3	0	1
2008-05-10	80	41	1023.2	1020	1	1
2008-05-11	80	52	1024.4	1021.5	7	7
2008-05-12	73	47	1024.2	1020.3	7	1
2008-05-13	87	34	1024.7	1021.4	1	2
2008-05-14	84	44	1024.4	1021	1	3
2008-05-15	80	40	1023.5	1018.7	1	2
2008-05-16	88	69	1017.7	1013.7	7	7
2008-05-17	68	78	1006.3	1005.9	7	7
2008-05-18	71	53	1008.4	1008.3	6	6
2008-05-19	85	36	1017.8	1014.3	1	1
2008-05-20	99	52	1021.9	1018.6	8	1
2008-05-21	81	47	1024.5	1022.9	6	2
2008-05-22	74	43	1030.4	1026	1	1
2008-05-23	81	49	1028.5	1024.8	7	1
2008-05-24	74	43	1026.7	1023.5	0	1
2008-05-25	81	48	1025.5	1021.8	1	5
2008-05-26	99	50	1026.6	1023.1	7	4
2008-05-27	93	55	1025.7	1022.3	1	1
2008-05-28	91	50	1027.2	1024.7	1	4
2008-05-29	92	45	1032.1	1029.6	2	1
2008-05-30	83	46	1033.2	1028.7	1	1
2008-05-31	84	38	1030.2	1027.4	2	3
2008-06-01	99	67	1032.3	1028.9	7	7
2008-06-02	85	67	1030.3	1027.1	7	7

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-06-03	99	86	1028.5	1027.2	8	8
2008-06-04	78	65	1028.2	1025.7	8	7
2008-06-05	84	65	1025.9	1023.4	6	7
2008-06-06	74	42	1025.2	1022.1	1	1
2008-06-07	86	42	1027.6	1027	1	1
2008-06-08	76	70	1034.3	1031.7	7	7
2008-06-09	75	65	1031.4	1027.9	7	7
2008-06-10	97	94	1027.8	1024.3	8	8
2008-06-11	90	65	1022.9	1022.1	7	7
2008-06-12	86	48	1023.7	1019.7	1	1
2008-06-13	82	58	1013.7	1014.1	4	2
2008-06-14	73	49	1025.7	1025.6	4	1
2008-06-15	72	57	1029	1026.3	1	2
2008-06-16	78	51	1027.8	1024.8	7	2
2008-06-17	80	58	1028.6	1025.5	7	1
2008-06-18	97	70	1026.7	1023.5	5	7
2008-06-19	95	63	1022.5	1019.7	8	7
2008-06-20	91	86	1021.2	1019.2	8	7
2008-06-21	99	59	1023	1021	3	6
2008-06-22	74	44	1026.4	1025.7	5	4
2008-06-23	92	53	1025.1	1022.1	7	7
2008-06-24	95	69	1028.2	1024.2	8	6
2008-06-25	75	54	1026.3	1021.9	3	1
2008-06-26	81	48	1018.7	1020.4	6	1
2008-06-27	71	51	1027.8	1025	1	1
2008-06-28	90	41	1029.1	1026	4	3
2008-06-29	84	38	1024.6	1020.9	7	6

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-06-30	53	68	1017.2	1010.2	1	7
2008-07-01	73	56	1012.1	1012.3	3	1
2008-07-02	62	53	1013.5	1014	5	1
2008-07-03	82	53	1024.3	1021.9	1	1
2008-07-04	80	52	1030.5	1030	1	6
2008-07-05	81	55	1032.9	1028.9	1	2
2008-07-06	86	59	1030	1026.3	5	5
2008-07-07	99	96	1024.4	1021.1	7	8
2008-07-08	93	72	1021.6	1019.6	6	6
2008-07-09	95	68	1020	1016.8	7	6
2008-07-10	83	85	1014.6	1010.5	7	7
2008-07-11	76	66	1016.7	1018.2	7	3
2008-07-12	87	54	1027.1	1022.7	1	7
2008-07-13	92	69	1024.4	1021.9	7	7
2008-07-14	93	51	1021.8	1019	5	1
2008-07-15	83	46	1023	1020.1	1	7
2008-07-16	99	62	1024.1	1020.8	7	7
2008-07-17	99	55	1019.7	1013.9	7	3
2008-07-18	75	73	1010.8	1009.4	3	5
2008-07-19	71	48	1016.7	1014.9	1	3
2008-07-20	92	87	1014.4	1009	7	8
2008-07-21	78	56	1006.9	1010.3	7	3
2008-07-22	68	33	1025.7	1025.8	1	1
2008-07-23	76	48	1033.5	1031.1	1	5
2008-07-24	82	51	1032.2	1027.4	3	5
2008-07-25	92	25	1025.7	1020.9	2	1
2008-07-26	88	50	1020	1015.3	7	7

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-07-27	92	66	1013.8	1010	7	7
2008-07-28	74	56	1014.1	1014.9	2	4
2008-07-29	69	49	1020.9	1016.8	3	1
2008-07-30	84	48	1018.3	1014.6	1	1
2008-07-31	81	38	1017.3	1012.9	1	7
2008-08-01	64	82	1005.1	1001.3	8	7
2008-08-02	67	56	1013.9	1014.9	1	1
2008-08-03	70	48	1021.2	1018.1	1	1
2008-08-04	81	60	1020.9	1018.3	1	7
2008-08-05	78	72	1020.8	1019.2	3	8
2008-08-06	67	48	1018.2	1017.6	5	1
2008-08-07	89	78	1018.4	1015.9	7	7
2008-08-08	69	62	1019.8	1019.1	6	2
2008-08-09	76	70	1022.4	1019	1	6
2008-08-10	92	68	1014.8	1012.6	8	6
2008-08-11	72	60	1017.4	1016.6	1	1
2008-08-12	69	63	1021.3	1020.1	1	6
2008-08-13	58	41	1024.2	1022.4	1	1
2008-08-14	59	45	1022.2	1018.2	1	1
2008-08-15	54	51	1014.7	1012.4	1	3
2008-08-16	52	27	1012.7	1011.8	1	1
2008-08-17	59	35	1023.3	1022	0	1
2008-08-18	89	30	1025.2	1021.8	6	7
2008-08-19	94	47	1029.6	1025.6	6	6
2008-08-20	60	36	1027.3	1024.4	7	7
2008-08-21	50	27	1021.4	1017.8	8	8
2008-08-22	66	44	1018.5	1016.8	8	5

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-08-23	49	38	1025.9	1025.5	1	1
2008-08-24	64	28	1031	1028	0	0
2008-08-25	82	34	1030.2	1024.4	7	6
2008-08-26	77	23	1023.9	1020.3	5	7
2008-08-27	57	23	1027.7	1025.5	1	1
2008-08-28	60	43	1029.9	1025.9	1	5
2008-08-29	74	45	1027.6	1022.9	7	6
2008-08-30	81	45	1023.7	1017.5	7	8
2008-08-31	89	85	1011.2	1010.4	7	7
2008-09-01	60	34	1023.2	1021.7	1	1
2008-09-02	69	44	1027	1024.8	1	4
2008-09-03	69	51	1033.6	1033.2	7	5
2008-09-04	65	61	1035.7	1031.9	7	7
2008-09-05	79	73	1027.3	1023.5	8	8
2008-09-06	63	33	1019.7	1016.2	7	7
2008-09-07	84	33	1018.1	1013.9	7	7
2008-09-08	79	37	1020.6	1018.2	1	1
2008-09-09	53	25	1025.8	1020.9	0	0
2008-09-10	62	35	1022.8	1017.9	1	2
2008-09-11	68	20	1018.8	1013.9	1	1
2008-09-12	44	27	1014.6	1010.9	7	5
2008-09-13	47	31	1015.3	1011.4	3	1
2008-09-14	71	43	1007.3	1006	7	1
2008-09-15	63	33	1009.1	1001.5	3	6
2008-09-16	53	33	1017	1017.6	3	1
2008-09-17	69	43	1026.7	1022.1	7	1
2008-09-18	82	36	1021	1015.6	4	1

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-09-19	54	30	1016.4	1011.7	1	1
2008-09-20	46	15	1012.5	1010.9	1	0
2008-09-21	42	25	1018.7	1015.1	1	0
2008-09-22	51	50	1012.1	1006.5	1	8
2008-09-23	73	44	1016.7	1017.8	7	1
2008-09-24	67	33	1030.5	1027.7	1	1
2008-09-25	70	32	1032.2	1026.9	1	1
2008-09-26	48	35	1025.4	1020.6	2	6
2008-09-27	62	25	1022.8	1016.4	5	1
2008-09-28	36	25	1010.5	1009.9	1	1
2008-09-29	41	32	1015.2	1013.9	2	7
2008-09-30	52	34	1019.8	1017	2	1
2008-10-01	56	16	1020.8	1015.4	5	1
2008-10-02	44	21	1017.7	1014.1	0	0
2008-10-03	44	22	1016.3	1011.8	1	1
2008-10-04	94	65	1015.5	1013.2	8	4
2008-10-05	79	46	1018.1	1013.6	8	1
2008-10-06	65	40	1009.5	1005.3	5	4
2008-10-07	68	45	1016.8	1013.3	1	1
2008-10-08	72	33	1019.1	1017.5	0	1
2008-10-09	69	22	1026.2	1024.2	0	1
2008-10-10	82	25	1028.8	1024.9	1	2
2008-10-11	70	30	1027.8	1023.8	1	3
2008-10-12	67	27	1025.8	1021.5	0	4
2008-10-13	56	47	1020.9	1016	4	7
2008-10-14	82	61	1019.2	1016.7	7	3
2008-10-15	71	42	1022.3	1019.7	7	4

Date	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm
2008-10-16	73	30	1025.7	1022.3	1	1
2008-10-17	73	34	1024.5	1020.7	6	1
2008-10-18	80	31	1025.7	1022.2	1	7
2008-10-19	66	30	1024.1	1019.8	1	6
2008-10-20	60	20	1021.4	1017.5	1	1
2008-10-21	54	28	1022.3	1018.6	7	5
2008-10-22	43	24	1023.4	1023.1	1	5
2008-10-23	47	25	1026.6	1022.8	1	2
2008-10-24	71	29	1023.2	1018.4	3	8
2008-10-25	57	16	1018.8	1014.6	0	3
2008-10-26	45	20	1017.6	1014.2	5	8
2008-10-27	38	15	1016.1	1010.8	1	3
2008-10-28	45	22	1020	1016.9	0	1
2008-10-29	63	47	1024	1022.8	3	2
2008-10-30	69	39	1021	1016.2	6	7
2008-10-31	43	13	1009.6	1009.2	1	1

Tableau-A III-5 Données recueillies à Canberra relatives à la température à des heures précises et aux risques de pluie

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2007-11-01	14.4	23.6	No	3.6	Yes
2007-11-02	17.5	25.7	Yes	3.6	Yes
2007-11-03	15.4	20.2	Yes	39.8	Yes
2007-11-04	13.5	14.1	Yes	2.8	Yes
2007-11-05	11.1	15.4	Yes	0	No
2007-11-06	10.9	14.8	No	0.2	No
2007-11-07	12.4	17.3	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2007-11-08	12.1	15.5	No	0	No
2007-11-09	14.1	18.9	No	16.2	Yes
2007-11-10	13.3	21.7	Yes	0	No
2007-11-11	14.6	24	No	0.2	No
2007-11-12	16.8	26	No	0	No
2007-11-13	17	27.1	No	0	No
2007-11-14	19.7	30.7	No	0	No
2007-11-15	18.7	30.4	No	0	No
2007-11-16	19.1	30.7	No	0	No
2007-11-17	20.2	29.8	No	1.2	Yes
2007-11-18	20.1	28.6	Yes	0.6	No
2007-11-19	20.2	31.2	No	0	No
2007-11-20	22.8	32	No	0	No
2007-11-21	22.2	32.8	No	0.4	No
2007-11-22	16.5	18.3	No	25.8	Yes
2007-11-23	14	16.8	Yes	0.4	No
2007-11-24	17.8	22.8	No	0	No
2007-11-25	16.8	27.3	No	0	No
2007-11-26	19.8	25.1	No	0.2	No
2007-11-27	16.5	21.2	No	0	No
2007-11-28	18.8	26.7	No	0	No
2007-11-29	18.9	19.7	No	0.4	No
2007-11-30	17.3	23.2	No	22.6	Yes
2007-12-01	17	16.3	Yes	4.2	Yes
2007-12-02	16.7	25.2	Yes	0.2	No
2007-12-03	20.5	19.9	No	6.6	Yes
2007-12-04	18.8	25.1	Yes	0	No
2007-12-05	18.6	20.5	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2007-12-06	18.6	24.5	No	0	No
2007-12-07	21.5	22.6	No	4	Yes
2007-12-08	18	26.3	Yes	0	No
2007-12-09	20.6	28.6	No	0	No
2007-12-10	19.3	22.9	No	0.6	No
2007-12-11	13.8	16.4	No	0	No
2007-12-12	15.7	19.4	No	0	No
2007-12-13	15.8	19.2	No	0	No
2007-12-14	15.8	25	No	0	No
2007-12-15	17.4	29.1	No	5.4	Yes
2007-12-16	16.6	20	Yes	1.4	Yes
2007-12-17	15.3	20.8	Yes	0	No
2007-12-18	15.3	19.8	No	0	No
2007-12-19	16.1	20	No	3.4	Yes
2007-12-20	15.3	20.4	Yes	6.4	Yes
2007-12-21	21	28.2	Yes	11	Yes
2007-12-22	20.6	19.6	Yes	17.4	Yes
2007-12-23	14.9	18.8	Yes	0	No
2007-12-24	15	22.5	No	3.4	Yes
2007-12-25	12.4	21	Yes	0	No
2007-12-26	15.3	23.9	No	0	No
2007-12-27	19.8	19	No	14.4	Yes
2007-12-28	18.3	27.4	Yes	0	No
2007-12-29	21.5	30.3	No	0	No
2007-12-30	23.4	34.3	No	0	No
2007-12-31	21.9	32.2	No	0	No
2008-01-01	21.9	31.8	No	0	No
2008-01-02	23	33.6	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-01-03	19.2	22.3	No	0	No
2008-01-04	20.3	23.9	No	0	No
2008-01-05	18.6	26.8	No	0	No
2008-01-06	18.9	31.9	No	0	No
2008-01-07	20.8	31.3	No	0	No
2008-01-08	18.1	28	No	0	No
2008-01-09	19.3	27.9	No	0	No
2008-01-10	21.3	32.2	No	0	No
2008-01-11	24.7	34.5	No	0	No
2008-01-12	23.6	34	No	2	Yes
2008-01-13	23	31.8	Yes	0	No
2008-01-14	17	21.6	No	0	No
2008-01-15	18	31.7	No	0	No
2008-01-16	24.5	23.5	No	4.8	Yes
2008-01-17	17.1	23.9	Yes	0	No
2008-01-18	17.8	19.4	No	18.8	Yes
2008-01-19	18	18.6	Yes	12.2	Yes
2008-01-20	19.4	21.9	Yes	0.8	No
2008-01-21	17.4	18.6	No	0	No
2008-01-22	14.7	22.2	No	0	No
2008-01-23	16.2	22.7	No	0	No
2008-01-24	17	26.3	No	0	No
2008-01-25	17.6	24.6	No	0	No
2008-01-26	18.3	28.3	No	0	No
2008-01-27	19.7	30.7	No	0	No
2008-01-28	22.4	32.1	No	0	No
2008-01-29	19.8	32.3	No	0	No
2008-01-30	22.2	33.1	No	5.2	Yes

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-01-31	21.4	26.3	Yes	2.2	Yes
2008-02-01	16.5	22.9	Yes	0	No
2008-02-02	19.1	27	No	0	No
2008-02-03	21.7	29.1	No	1.8	Yes
2008-02-04	18.5	22.1	Yes	9	Yes
2008-02-05	18.4	21.5	Yes	1	No
2008-02-06	18.7	26.5	No	0	No
2008-02-07	17.9	20.3	No	16.2	Yes
2008-02-08	13.6	18.7	Yes	0	No
2008-02-09	13.8	18.7	No	0	No
2008-02-10	14	21.4	No	0	No
2008-02-11	14.6	24.2	No	4.4	Yes
2008-02-12	15.8	22.6	Yes	11	Yes
2008-02-13	16.3	16.8	Yes	0.2	No
2008-02-14	13.7	22	No	0	No
2008-02-15	15	24.3	No	0	No
2008-02-16	15.3	24.7	No	0	No
2008-02-17	15.4	22.5	No	0	No
2008-02-18	17.2	24.1	No	0	No
2008-02-19	16.9	26.1	No	0	No
2008-02-20	17.4	28.6	No	1.8	Yes
2008-02-21	17.5	24	Yes	6.6	Yes
2008-02-22	16.5	27.4	Yes	0	No
2008-02-23	15.7	24.5	No	0	No
2008-02-24	15.1	22.9	No	0	No
2008-02-25	15.8	24.3	No	0	No
2008-02-26	17.4	27	No	0	No
2008-02-27	17.6	26.8	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-02-28	15.7	15.3	No	10.4	Yes
2008-02-29	10.6	16.8	Yes	0	No
2008-03-01	11.7	20	No	0	No
2008-03-02	12.2	22.1	No	0	No
2008-03-03	13.8	25.5	No	0	No
2008-03-04	15.6	27	No	0	No
2008-03-05	15.7	26.2	No	0	No
2008-03-06	17.4	30	No	0	No
2008-03-07	17.5	27.7	No	3	Yes
2008-03-08	14	24.9	Yes	0	No
2008-03-09	17.8	27	No	0	No
2008-03-10	17.2	30.1	No	0	No
2008-03-11	20.2	32.7	No	0.2	No
2008-03-12	17.5	30.7	No	0	No
2008-03-13	16.1	31.7	No	0	No
2008-03-14	17.9	32.7	No	0	No
2008-03-15	21.4	34.1	No	0	No
2008-03-16	18.2	30.5	No	0	No
2008-03-17	17.7	28.8	No	0	No
2008-03-18	16.8	29.2	No	0	No
2008-03-19	19.7	24.5	No	0	No
2008-03-20	17.1	27.6	No	0	No
2008-03-21	13.3	13.9	No	0	No
2008-03-22	14	17.9	No	0	No
2008-03-23	15.5	24.2	No	0.6	No
2008-03-24	17.1	21.1	No	6.4	Yes
2008-03-25	16.4	16.6	Yes	19.8	Yes
2008-03-26	15.1	17.4	Yes	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-03-27	8.4	17.4	No	0	No
2008-03-28	9.2	17.2	No	0	No
2008-03-29	8.8	18.2	No	0	No
2008-03-30	11.2	18.1	No	0	No
2008-03-31	13.1	16.5	No	0	No
2008-04-01	8.7	21.2	No	0	No
2008-04-02	9.5	22.5	No	2.6	Yes
2008-04-03	9.1	13.7	Yes	0	No
2008-04-04	5.2	17.6	No	0	No
2008-04-05	8.2	20.5	No	0	No
2008-04-06	12.9	20.8	No	0	No
2008-04-07	14.9	19.5	No	0	No
2008-04-08	13.8	17.8	No	0	No
2008-04-09	14.3	18.6	No	0	No
2008-04-10	14	22.2	No	0	No
2008-04-11	12.7	22.4	No	0	No
2008-04-12	11.5	23.7	No	2	Yes
2008-04-13	10.7	19.1	Yes	5.2	Yes
2008-04-14	12.4	17.9	Yes	0	No
2008-04-15	12.3	20.4	No	0	No
2008-04-16	12.6	18.8	No	0	No
2008-04-17	12	20.4	No	0	No
2008-04-18	14	19.1	No	0	No
2008-04-19	13.6	15.2	No	0	No
2008-04-20	13.5	15.6	No	0	No
2008-04-21	10.6	19.8	No	0	No
2008-04-22	12.5	17.7	No	0	No
2008-04-23	12.6	18.1	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-04-24	9.2	17.6	No	0	No
2008-04-25	10.1	20.6	No	0.2	No
2008-04-26	7.9	20.2	No	0	No
2008-04-27	14.4	16.6	No	7.2	Yes
2008-04-28	7.1	8.8	Yes	0.2	No
2008-04-29	6.3	13.2	No	0	No
2008-04-30	5.3	13.9	No	0	No
2008-05-01	10.5	12.4	No	0	No
2008-05-02	8.7	15.9	No	0	No
2008-05-03	7.4	16.2	No	0	No
2008-05-04	6.2	15.4	No	0	No
2008-05-05	7.7	18.5	No	0	No
2008-05-06	12.5	15.4	No	0	No
2008-05-07	12.4	16.5	No	0	No
2008-05-08	6.7	16.9	No	0	No
2008-05-09	7.6	20.7	No	0	No
2008-05-10	8.4	20.9	No	0	No
2008-05-11	13.5	17.2	No	0	No
2008-05-12	12.1	18.8	No	0	No
2008-05-13	7.9	18.6	No	0	No
2008-05-14	8.5	17.8	No	0	No
2008-05-15	6.8	17.7	No	1.8	Yes
2008-05-16	10	15.2	Yes	1	No
2008-05-17	11.3	7.3	No	0.8	No
2008-05-18	7.9	13.3	No	0	No
2008-05-19	3.5	16.7	No	0	No
2008-05-20	3.8	16.2	No	0	No
2008-05-21	8	16.6	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-05-22	5.9	13.8	No	0	No
2008-05-23	5.3	14.5	No	0	No
2008-05-24	6.5	13.9	No	0	No
2008-05-25	6.2	17.2	No	3.8	Yes
2008-05-26	7	17.9	Yes	5.2	Yes
2008-05-27	8.9	17.7	Yes	0	No
2008-05-28	6.2	16.3	No	0.2	No
2008-05-29	5.5	17.3	No	0	No
2008-05-30	5.6	17.4	No	0	No
2008-05-31	4.7	18.2	No	0	No
2008-06-01	4.6	13.9	No	0	No
2008-06-02	9.9	15.5	No	0.8	No
2008-06-03	12.1	13.6	No	3.8	Yes
2008-06-04	11.5	13.3	Yes	0	No
2008-06-05	10.3	14.8	No	0	No
2008-06-06	9.6	16.2	No	0	No
2008-06-07	4.9	14.5	No	0	No
2008-06-08	7.9	13	No	0	No
2008-06-09	12.1	14.5	No	6.2	Yes
2008-06-10	10.3	11.6	Yes	4.8	Yes
2008-06-11	12.9	13.7	Yes	0.2	No
2008-06-12	11.2	14.7	No	0.6	No
2008-06-13	8.8	11.1	No	0	No
2008-06-14	6.5	11	No	0	No
2008-06-15	9	13.8	No	0	No
2008-06-16	9.5	16	No	0	No
2008-06-17	9.4	14.9	No	0	No
2008-06-18	4	15	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-06-19	8.4	14.8	No	4.8	Yes
2008-06-20	9.7	11.3	Yes	0.6	No
2008-06-21	4.2	12.1	No	0	No
2008-06-22	8.8	12.5	No	0	No
2008-06-23	5.5	12.3	No	0	No
2008-06-24	4.3	11.7	No	0	No
2008-06-25	7.8	11.3	No	0	No
2008-06-26	7	11.5	No	0	No
2008-06-27	6.7	12.7	No	0	No
2008-06-28	2.1	14.5	No	0	No
2008-06-29	6.5	14.6	No	0	No
2008-06-30	9.7	12.9	No	2	Yes
2008-07-01	6.9	11.1	Yes	0.2	No
2008-07-02	10.6	12.3	No	0	No
2008-07-03	3	11.1	No	0	No
2008-07-04	5.9	10.2	No	0	No
2008-07-05	4.7	12.6	No	0	No
2008-07-06	3.7	13.6	No	0.8	No
2008-07-07	6.3	8.6	No	16.8	Yes
2008-07-08	5.5	7.4	Yes	0	No
2008-07-09	1.8	8	No	0	No
2008-07-10	6	6.9	No	1.6	Yes
2008-07-11	5.5	7.1	Yes	0.2	No
2008-07-12	0.8	9.9	No	0	No
2008-07-13	2.8	10.7	No	0	No
2008-07-14	6.7	15.9	No	0	No
2008-07-15	3.9	13.3	No	0.2	No
2008-07-16	2.7	12.1	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-07-17	1.3	11	No	0	No
2008-07-18	7.7	8.2	No	1.2	Yes
2008-07-19	6	10.7	Yes	0.2	No
2008-07-20	2.4	8.7	No	19.2	Yes
2008-07-21	7.6	11.2	Yes	0	No
2008-07-22	3.4	11	No	0	No
2008-07-23	3.6	10.1	No	0	No
2008-07-24	3.7	10.8	No	0	No
2008-07-25	1.4	12.3	No	0.2	No
2008-07-26	1.2	10.9	No	0	No
2008-07-27	1	7.8	No	1.4	Yes
2008-07-28	5.6	9.6	Yes	0	No
2008-07-29	6.5	11.2	No	0	No
2008-07-30	2.7	11.5	No	0	No
2008-07-31	3	12.8	No	1	No
2008-08-01	9.6	8.6	No	6.6	Yes
2008-08-02	6.9	10.5	Yes	0	No
2008-08-03	7.8	11.6	No	0	No
2008-08-04	2.6	9.1	No	0	No
2008-08-05	5.5	7.2	No	4	Yes
2008-08-06	8.9	13.6	Yes	0	No
2008-08-07	3.8	5.7	No	1.2	Yes
2008-08-08	7.6	11.3	Yes	0	No
2008-08-09	5.6	8.1	No	0.4	No
2008-08-10	3.3	5.1	No	0.4	No
2008-08-11	4.7	8.3	No	0	No
2008-08-12	7	8.6	No	0	No
2008-08-13	6.5	10.4	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-08-14	6.2	9.5	No	0	No
2008-08-15	9.2	12	No	0	No
2008-08-16	8.7	13.5	No	0	No
2008-08-17	4.5	10.8	No	0	No
2008-08-18	1.4	11.2	No	0	No
2008-08-19	0.1	12.2	No	0	No
2008-08-20	7.9	12.4	No	0	No
2008-08-21	8.6	13.2	No	0	No
2008-08-22	6.9	10.7	No	0	No
2008-08-23	6.7	11.1	No	0	No
2008-08-24	5.2	14	No	0	No
2008-08-25	5.3	12.8	No	0	No
2008-08-26	4.8	15.8	No	0	No
2008-08-27	7.9	14.4	No	0	No
2008-08-28	7.2	12.3	No	0	No
2008-08-29	7.3	14.6	No	0	No
2008-08-30	6.2	15.4	No	4	Yes
2008-08-31	13.5	12.6	Yes	7.4	Yes
2008-09-01	8.8	14.3	Yes	0	No
2008-09-02	9	14.9	No	0	No
2008-09-03	9.3	11.1	No	0.2	No
2008-09-04	8.3	10.2	No	0	No
2008-09-05	8.7	10.8	No	0	No
2008-09-06	10	15.7	No	0	No
2008-09-07	6.4	15.5	No	1	No
2008-09-08	7.2	15.3	No	0	No
2008-09-09	5.5	13.4	No	0	No
2008-09-10	6.3	14.9	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-09-11	6	16.3	No	0	No
2008-09-12	11.9	21.5	No	0	No
2008-09-13	17.5	25.1	No	9.8	Yes
2008-09-14	16.7	18.2	Yes	1.6	Yes
2008-09-15	15.6	18.5	Yes	3.4	Yes
2008-09-16	7.1	12.2	Yes	0	No
2008-09-17	7.4	13.7	No	0	No
2008-09-18	7	17.8	No	0	No
2008-09-19	16.4	22.4	No	0	No
2008-09-20	18.3	21.3	No	0	No
2008-09-21	14.5	18.3	No	0	No
2008-09-22	15.9	21	No	17.4	Yes
2008-09-23	8	14.3	Yes	0	No
2008-09-24	10.7	16.5	No	0	No
2008-09-25	11	18.1	No	0	No
2008-09-26	15.4	20	No	0	No
2008-09-27	15.9	24.7	No	0	No
2008-09-28	20.4	23	No	0	No
2008-09-29	14.9	16	No	0	No
2008-09-30	9.8	16.3	No	0	No
2008-10-01	11.3	19.5	No	0	No
2008-10-02	19	26.7	No	0	No
2008-10-03	22.5	28.4	No	7.6	Yes
2008-10-04	14.5	19.3	Yes	3	Yes
2008-10-05	11.7	19.8	Yes	0	No
2008-10-06	12.8	16.2	No	8.2	Yes
2008-10-07	6.9	14.6	Yes	0	No
2008-10-08	7.2	16.6	No	0	No

Date	Temp9am	Temp3pm	RainToday	RISK_MM	RainTomorrow
2008-10-09	8.1	18.8	No	0	No
2008-10-10	10	21.4	No	0	No
2008-10-11	13.6	20.6	No	0	No
2008-10-12	12.6	22.3	No	0	No
2008-10-13	16.3	23.2	No	13.2	Yes
2008-10-14	14.5	19.4	Yes	0.6	No
2008-10-15	11.6	18.4	No	0	No
2008-10-16	9.6	19.2	No	0	No
2008-10-17	11.6	21.9	No	0	No
2008-10-18	12.7	23.7	No	0	No
2008-10-19	16.8	27.4	No	0.2	No
2008-10-20	16.4	26.3	No	0	No
2008-10-21	11.4	18.5	No	0.8	No
2008-10-22	8.3	14.3	No	0	No
2008-10-23	9.1	16.3	No	0	No
2008-10-24	9.4	19.1	No	0	No
2008-10-25	12	24.8	No	0	No
2008-10-26	16.3	25.9	No	0	No
2008-10-27	20.4	30	No	0	No
2008-10-28	17.2	28.2	No	0	No
2008-10-29	14.5	18.3	No	0	No
2008-10-30	15.8	25.9	No	0	No
2008-10-31	23.8	28.6	No	0	No

BIBLIOGRAPHIE

- Abbasy, N. H., et H. M. Ismail. 2009. « A Unified Approach for the Optimal PMU Location for Power System State Estimation ». *IEEE Transactions on Power Systems*, vol. 24, n° 2, p. 806-813.
- Aboreshaid, Saleh, Roy Billinton et Mahmud Fotuhi-Firuzabad. 1996. « Probabilistic transient stability studies using the method of bisection [power systems] ». *IEEE Transactions on Power Systems*, vol. 11, n° 4, p. 1990-1995.
- Adibi, M. M., P. M. Hirsch et J. A. Jordan. 1974. « Solution methods for transient and dynamic stability ». *Proceedings of the IEEE*, vol. 62, n° 7, p. 951-958.
- Aggarwal, Charu C. 2015. *Data Mining : The Textbook* (2015). Cham: Springer International Publishing, 1 ressource en ligne (XXIX, 734 p.) p.
- AIEE. 1926. *First report of power stability*. Coll. « Transactions of the American Institute of Electrical Engineers »: American Institute of Electrical Engineers Subcommittee on Interconnections and Stability Factors, 51-80 p.
- AL-Masri, Ahmed Naufal, Mohd Zainal Abidin Ab Kadir, Hashim Hizam et Norman Mariun. 2013. « A novel implementation for generator rotor angle stability prediction using an adaptive artificial neural network application for dynamic security assessment ». *IEEE Transactions on Power Systems*, vol. 28, n° 3, p. 2516-2525.
- Alcheikh-Hamoud, K., N. Hadjsaid, Y. Besanger et J. P. Rognon. 2009. « Decision tree based filter for control area external contingencies screening ». In *PowerTech, 2009 IEEE Bucharest*. (June 28 2009-July 2 2009), p. 1-8.
- Alpaydin, Ethem. 2014. *Introduction to machine learning*. MIT press.
- Aminifar, F., A. Khodaei, M. Fotuhi-Firuzabad et M. Shahidehpour. 2010. « Contingency-Constrained PMU Placement in Power Networks ». *IEEE Transactions on Power Systems*, vol. 25, n° 1, p. 516-523.
- Amjadi, N. 2005. « Application of a new fuzzy neural network to transient stability prediction ». In *Power Engineering Society General Meeting, 2005. IEEE*. (12-16 June 2005), p. 636-643 Vol. 1.
- Amjadi, N., et S. A. Banihashemi. 2010. « Transient stability prediction of power systems by a new synchronism status index and hybrid classifier ». *Generation, Transmission & Distribution, IET*, vol. 4, n° 4, p. 509-518.

- Amjadi, N., et S. F. Majedi. 2007. « Transient Stability Prediction by a Hybrid Intelligent System ». *Power Systems, IEEE Transactions on*, vol. 22, n° 3, p. 1275-1283.
- Anders, G.J. 1990. *Probability concepts in electric power systems*. Wiley.
- Anderson, P. M., et A. A. Fouad. 2003. *Power system control and stability* (2003), 2nd ed. Piscataway, N.J.: IEEE Press , Wiley-Interscience, 1 texte électronique (xiv, 658 p.) p.
- Anderson, Paul M, et Paul M Anderson. 1995. *Analysis of faulted power systems*, 445. IEEE press Piscataway.
- Anderson, Paul M, et Anjan Bose. 1983. « A probabilistic approach to power system stability analysis ». *IEEE Transactions on Power Apparatus and Systems*, n° 8, p. 2430-2439.
- Apraez, Camilo. 2012. « Étude comparative de méthodes de simulation de la stabilité transitoire ». École de technologie supérieure.
- Arani, M. F. M., et E. F. El-Saadany. 2013. « Implementing Virtual Inertia in DFIG-Based Wind Power Generation ». *IEEE Transactions on Power Systems*, vol. 28, n° 2, p. 1373-1384.
- Atanackovic, Djordje, Jose H Clapauch, Greg Dwernychuk, Jim Gurney et Harry Lee. 2008. « First steps to wide area control ». *IEEE Power and Energy Magazine*, vol. 6, n° 1, p. 61-68.
- Athay, T., R. Podmore et S. Virmani. 1979. « A Practical Method for the Direct Analysis of Transient Stability ». *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, n° 2, p. 573-584.
- Bahbah, A. G., et A. A. Girgis. 2004. « New method for generators' angles and angular velocities prediction for transient stability assessment of multimachine power systems using recurrent artificial neural network ». *Power Systems, IEEE Transactions on*, vol. 19, n° 2, p. 1015-1022.
- Barbier, C, L Carpentier et F Saccomanno. 1978. « CIGRE SC32 Report: "Tentative classification and terminologies relating to stability problems of power systems," ». *Electra*, n° 56.
- Becker, Richard A, John M Chambers et Allan R Wilks. 1988. « The new S language: a programming environment for data analysis and graphics, Wadsworth and Brooks ». *Cole: Pacific Grove*.
- Begovic, Miroslav, Damir Novosel, Daniel Karlsson, Charlie Henville et Gary Michel. 2005. « Wide-area protection and emergency control ». *Proceedings of the IEEE*, vol. 93, n° 5, p. 876-891.

- Bergen, Arthur R. 2000. *Power systems analysis*, 2nd ed.. Upper Saddle River, NJ: Prentice Hall, xii, 619 p. p.
- Bernabeu, E. E., J. S. Thorp et V. Centeno. 2012. « Methodology for a Security/Dependability Adaptive Protection Scheme Based on Data Mining ». *IEEE Transactions on Power Delivery*, vol. 27, n° 1, p. 104-111.
- Bevrani, Hassan. 2009. *Robust power system frequency control*, 85. Springer.
- Bhargava, B. 1999. « Synchronized Phasor Measurement System project at Southern California Edison Co ». In *Power Engineering Society Summer Meeting, 1999. IEEE*. (18-22 Jul 1999) Vol. 1, p. 16-22 vol.1.
- Bhargava, Bharat, Armando Salazar et Transmission Distribution Business Unit. 2007. « Synchronized phasor measurement system (SPMS) for monitoring transmission system at SCE ». In *NASPI Meeting, Carson, CA, May*. p. 967-979.
- Billinton, R, et S Aboreshaid. 1995. « Stochastic modelling of high-speed reclosing in probabilistic transient stability studies ». *IEE Proceedings-Generation, Transmission and Distribution*, vol. 142, n° 4, p. 350-354.
- Billinton, R, PRS Kuruganty et MF Carvalho. 1979. « An approximate method for probabilistic assessment of transient stability ». *IEEE Transactions on Reliability*, vol. 28, n° 3, p. 255-258.
- Billinton, Roy. 1970. *Power system reliability evaluation*. Taylor & Francis.
- Borlase, Stuart. 2013. *Smart grids infrastructure, technology, and solutions* (2013). Boca Raton, Fla.: CRC Press, 1 ressource en ligne (xviii, 577 p., [12] p. de pl.) p.
- Breiman, Leo. 1996a. « Bagging predictors ». *Machine learning*, vol. 24, n° 2, p. 123-140.
- Breiman, Leo. 1996b. « Bias, variance, and arcing classifiers ». <<http://statistics.berkeley.edu/tech-reports/460>>. Consulté le 07 septembre 2016.
- Breiman, Leo. 1998. « Arcing classifier (with discussion and a rejoinder by the author) ». *The annals of statistics*, vol. 26, n° 3, p. 801-849.
- Breiman, Leo, et Adele Cutler. 2001. « Random forests ». *Machine learning*, vol. 45, n° 1, p. 5-32.
- Breiman, Leo, Jerome H Friedman, Richard A Olshen et Charles J Stone. 1999. *Classification and Regression Trees*. CRC Press, New York.

- Breslow, Leonard A, et David W Aha. 1997. « Simplifying decision trees: A survey ». *The Knowledge Engineering Review*, vol. 12, n° 01, p. 1-40.
- Bretas, A. S., et A. G. Phadke. 2003. « Artificial neural networks in power system restoration ». *IEEE Transactions on Power Delivery*, vol. 18, n° 4, p. 1181-1186.
- Brueni, Dennis J, et Lenwood S Heath. 2005. « The PMU placement problem ». *SIAM Journal on Discrete Mathematics*, vol. 19, n° 3, p. 744-761.
- Burnett, R. O., M. M. Butts, T. W. Cease, V. Centeno, G. Michel, R. J. Murphy et A. G. Phadke. 1994. « Synchronized phasor measurements of a power system event ». *IEEE Transactions on Power Systems*, vol. 9, n° 3, p. 1643-1650.
- Caron, Stéphane. 2011. « Une introduction aux arbres de décision ». < <https://scaron.info/doc/intro-arbres-decision/intro.pdf> >. Consulté le 6 décembre 2016.
- Caruana, Rich, et Alexandru Niculescu-Mizil. 2006. « An empirical comparison of supervised learning algorithms ». In *Proceedings of the 23rd international conference on Machine learning*. p. 161-168. ACM.
- Castillo, Enrique, Antonio J Conejo, Pablo Pedregal, Ricardo Garcia et Natalia Alguacil. 2011. *Building and solving mathematical programming models in engineering and science*, 62. John Wiley & Sons.
- Cepeda, Jaime C., José L. Rueda, Delia G. Colomé et Diego E. Echeverria. 2014. « Real-time transient stability assessment based on centre-of-inertia estimation from phasor measurement unit records ». *Generation, Transmission & Distribution, IET*, vol. 8, n° 8, p. 1363-1376.
- Chakrabarti, S., et E. Kyriakides. 2008. « Optimal Placement of Phasor Measurement Units for Power System Observability ». *IEEE Transactions on Power Systems*, vol. 23, n° 3, p. 1433-1440.
- Chakrabarti, Saikat, Elias Kyriakides et Demetrios G Eliades. 2009. « Placement of synchronized measurements for power system observability ». *IEEE Transactions on Power Delivery*, vol. 24, n° 1, p. 12-19.
- Chiang, Hsiao-Dong, Felix F Wu et Pravin P Varaiya. 1994. « A BCU method for direct analysis of power system transient stability ». *IEEE Transactions on Power Systems*, vol. 9, n° 3, p. 1194-1208.
- Chih-Wen, Liu, Su Mu-Chun, Tsay Shuenn-Shing et Wang Yi-Jen. 1999. « Application of a novel fuzzy neural network to real-time transient stability swings prediction based on

- synchronized phasor measurements ». *Power Systems, IEEE Transactions on*, vol. 14, n° 2, p. 685-692.
- Chiodo, E, F Gagliardi et D Lauria. 1994. « Probabilistic approach to transient stability evaluation ». *IEE Proceedings-Generation, Transmission and Distribution*, vol. 141, n° 5, p. 537-544.
- Chirici, Gherardo, Matteo Mura, Daniel McInerney, Nicolas Py, Erkki O. Tomppo, Lars T. Waser, Davide Travaglini et Ronald E. McRoberts. 2016. « A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data ». *Remote Sensing of Environment*, vol. 176, p. 282-294.
- Chow, Joe H. 2013. *Power system coherency and model reduction*. Springer.
- Chow, Joe, et Graham Rogers. 2000. « Power system toolbox ». *Cherry Tree Scientific Software, [Online] Available : http://www.eps.ee.kth.se/personal/vanfretti/pst/Power_System_Toolbox_Webpage/PS_T.html*. Consulté le 9 décembre 2016.
- Christie, Richard D, SN Talukdar et James C Nixon. 1990. « CQR: a hybrid expert system for security assessment ». *IEEE Transactions on Power Systems*, vol. 5, n° 4, p. 1503-1509.
- Cigré, SC38. 1996. « Analysis and control of power system oscillations ». *Final Report, Task Force*, vol. 7.
- Committee, IEEE Power System Engineering. 1982. *Proposed terms and definitions for power system stability*. Coll. « Power Apparatus and Systems »: Technical report, IEEE Trans, 1894-1897 p.
- Crary, SB, I Herlitz et B Favez. 1948. « CIGRE SC32 Report: “System stability and voltage, power and frequency control,” ». *CIGRE, Appendix*, vol. 1.
- Dasgupta, Sambarta, Magesh Paramasivam, Umesh Vaidya et Venkataramana Ajjarapu. 2015. « PMU-based model-free approach for real-time rotor angle monitoring ». *IEEE Transactions on Power Systems*, vol. 30, n° 5, p. 2818-2819.
- De Boeck, Steven, Kaushik Das, Vincenzo Trovato, Jukka Turunen, Milenko Halat, Poul Sorensen et Dirk Van Hertem. 2016. « Review of Defence Plans in Europe: Current Status, Strengths and Opportunities ». *CIGRE Science & Engineering*, vol. 5, p. 2-11.
- De La Ree, Jaime, Virgilio Centeno, James S Thorp et Arun G Phadke. 2010. « Synchronized phasor measurement applications in power systems ». *IEEE Transactions on Smart Grid*, vol. 1, n° 1, p. 20-27.

De Micheaux, Pierre Lafaye, Rémy Drouilhet et Benoît Liquet. 2011. « Présentation du logiciel R ». In *Le logiciel R: Maîtriser le langage Effectuer des analyses statistiques*. de Micheaux2011. p. 1-6. Paris: Springer Paris. <http://dx.doi.org/10.1007/978-2-8178-0115-5_1>. Consulté le 3 décembre 2016.

Del Angel, Alberto, Pierre Del Angel, Damien Geurts, Mevludin Ernst, Louis Glavic et Wehenkel. 2007. « Estimation of rotor angles of synchronous machines using artificial neural networks and local PMU-based quantities ». *Neurocomputing*, vol. 70, n° 16, p. 2668-2678.

Del Angel, Alberto, Mevludin Glavic et Louis Wehenkel. 2003. « Using artificial neural networks to estimate rotor angles and speeds from phasor measurements ». *Proceedings of Intelligent Systems Applications to Power Systems, ISAP03*.

Delivery, US DOE Electricity. 2013. *Synchrophasor technologies and their deployment in the recovery act smart grid programs*. Department of Energy USA. <https://www.smartgrid.gov/files/Synchrophasor_Report_08_09_2013_DOE_2_version_0.pdf>. Consulté le 18 mai 2015.

Devroye, Luc, et Laszlo Gyorfi. 1985. *Nonparametric density estimation: the L1 view*, 119. John Wiley & Sons Incorporated.

Dietterich, Thomas G. 2000. « An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization ». *Machine learning*, vol. 40, n° 2, p. 139-157.

Dmitrova, E., M. L. Wittrock, J. H. x00F, hannsson et A. H. Nielsen. 2015. « Early Prevention Method for Power System Instability ». *IEEE Transactions on Power Systems*, vol. 30, n° 4, p. 1784-1792.

DOE. 2014. *Factors affecting PMU installation costs*. U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability. <https://www.smartgrid.gov/files/PMU-cost-studyfinal-10162014_1.pdf>. Consulté le 30 juillet 2016.

DOE, et NRC. *Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*. U.S.- Canada Power System Outage Task Force. <<http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinalImplementationReport%282%29.pdf>>. Consulté le 9 août 2016.

Dommel, HW, et N Sato. 1972. « Fast transient stability soultions ». *IEEE Transactions on Power Apparatus and Systems*, n° 4, p. 1643-1650.

Dua, Devesh, Sanjay Dambhare, Rajeev Kumar Gajbhiye et SA Soman. 2008. « Optimal multistage scheduling of PMU placement: An ILP approach ». *IEEE Transactions on Power Delivery*, vol. 23, n° 4, p. 1812-1820.

- Duda, Richard O., Peter E. Hart et David G. Stork. 2001. *Pattern classification* (2001), 2nd ed. New York, Toronto: John Wiley & Sons, xx, 654 p. p.
- Efron, Bradley, et B Efron. 1982. *The jackknife, the bootstrap and other resampling plans*, 38. SIAM.
- Efron, Bradley, et Robert J Tibshirani. 1994. *An introduction to the bootstrap*. CRC press.
- Ekanayake, J. B. 2012. *Smart grid technology and applications* (2012). Chichester, West Sussex, U.K.: John Wiley & Sons Ltd.
- Ekanayake, Janaka, et Nick Jenkins. 2004. « Comparison of the response of doubly fed and fixed-speed induction generator wind turbines to changes in network frequency ». *IEEE Transactions on Energy conversion*, vol. 19, n° 4, p. 800-802.
- El-Abiad, Ahmed H, et K Nagappan. 1966. « Transient stability regions of multimachine power systems ». *IEEE Transactions on Power Apparatus and Systems*, n° 2, p. 169-179.
- El Aouni, Wassil. 2014. « Équivalents dynamiques pour l'étude de la stabilité transitoire ». École de technologie supérieure.
- El Guindi, Mahmoud, et M Mansour. 1982. « Transient stability of a power system by the Liapunov method considering the transfer conductances ». *IEEE Transactions on Power Apparatus and Systems*, n° 5, p. 1088-1094.
- Emami, Roozbeh, et Ali Abur. 2010. « Robust measurement design by placing synchronized phasor measurements on network branches ». *IEEE Transactions on power systems*, vol. 25, n° 1, p. 38-43.
- EPRI. 2016. « Online Measurement-based Voltage Stability Assessment Tool (MBVSA) 2.1». <<http://www.ePRI.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000003002006901>>. Consulté le 12 août 2016.
- ERGEG. 2007. *The lessons to be learned from the large disturbance in the European power system on the 4th of November 2006*. European Regulators' Group for Electricity and Gas. <http://www.energy-regulators.eu/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Electricity/2007/E06-BAG-01-06_Blackout-FinalReport_2007-02-06.pdf>. Consulté le 10 août 2016.
- Ernst, Damien, Daniel Ruiz-Vega, Mania Pavella, Peter M Hirsch et Dejan Sobajic. 2001. « A unified approach to transient stability contingency filtering, ranking and assessment ». *IEEE Transactions on Power Systems*, vol. 16, n° 3, p. 435-443.

- Fayyad, Usama, Gregory Piatetsky-Shapiro et Padhraic Smyth. 1996. « From data mining to knowledge discovery in databases ». *AI magazine*, vol. 17, n° 3, p. 37.
- Fink, Lester H, et Kjell Carlsen. 1978. « Operating under stress and strain ». *IEEE Spectrum; (United States)*, vol. 15, n° 3.
- Fink, LH, et TA Trygar. 1979. *Systems engineering for power: emergency operating state control*. Department of Energy, Washington, DC (USA). Office of Energy Technology.
- Force, Task. 1999. *38.02. 14, "Analysis and Modeling Needs of Power Systems Under Major Frequency Disturbances,"*. technical report, CIGRE.
- Fouad, AA, S Vekataraman et JA Davis. 1991. « An expert system for security trend analysis of a stability-limited power system ». *IEEE Transactions on Power Systems*, vol. 6, n° 3, p. 1077-1084.
- Fouad, Abdel-Azia, et Vijay Vittal. 1991. *Power system transient stability analysis using the transient energy function method*. Pearson Education.
- Freund, Yoav, et Robert E Schapire. 1996. « Experiments with a new boosting algorithm ». In *Icml*. Vol. 96, p. 148-156.
- Friedman, Jerome, Trevor Hastie et Robert Tibshirani. 2001. *The elements of statistical learning*, 1. Springer series in statistics Springer, Berlin.
- Fu, Chengjun, et Anjan Bose. 1999. « Contingency ranking based on severity indices in dynamic security analysis ». *IEEE Transactions on power systems*, vol. 14, n° 3, p. 980-985.
- Gao, B, GK Morison, P Kundur, CA CANIZARES, CW TAYLOR, C PARKER, MK PAL, K TOMSOVIC, V VENKATASUBRAMANIAN et A BOSE. 1996. « Towards the development of a systematic approach for voltage stability assessment of large-scale power systems ». *IEEE Transactions on Power Systems*, vol. 11, n° 3, p. 1314-1324.
- Gingras, JP, S Breault, R Brodeur, C Crevier, A Déry et A Vallée. 2000. « Stratégie de renforcement du réseau d'Hydro-Québec à la suite du verglas exceptionnel de janvier 1998: une démarche pour sécuriser davantage l'alimentation électrique ». *CIGR É Session, Paris, Paper*, p. 37-101.
- Glover, J Duncan, Mulukutla S Sarma et Thomas Overbye. 2012. *Power System Analysis & Design, SI Version*. Cengage Learning.
- Glover, J. Duncan. 2002. *Power system analysis and design*, 3rd ed.. Pacific Grove, Calif.: Brooks/Cole, xiv, 656 p. p.

- Gómez-Expósito, Antonio, Antonio J Conejo et Claudio Cañizares. 2016. *Electric energy systems: analysis and operation*. CRC Press.
- Gomez, F. R., A. D. Rajapakse, U. D. Annakkage et I. T. Fernando. 2011. « Support Vector Machine-Based Algorithm for Post-Fault Transient Stability Status Prediction Using Synchronized Measurements ». *Power Systems, IEEE Transactions on*, vol. 26, n° 3, p. 1474-1483.
- Gou, B. 2008. « Generalized Integer Linear Programming Formulation for Optimal PMU Placement ». *IEEE Transactions on Power Systems*, vol. 23, n° 3, p. 1099-1104.
- Grigsby, Leonard L. 2016. *Power Systems Stability and Control*, Third. CRC Press.
- Guo, Hao-Zhi, Huan Xie, Bao-Hui Zhang, Guang-liang Yu, Peng Li, Zhi-qian Bo et A. Klimek. 2010. « Study on power system transient instability detection based on wide area measurement system ». *European Transactions on Electrical Power*, vol. 20, n° 2, p. 184-205.
- Haque, MH, et AHMA Rahim. 1989. « Determination of first swing stability limit of multimachine power systems through Taylor series expansions ». In *IEE Proceedings C-Generation, Transmission and Distribution*. Vol. 136, p. 373-380. IET.
- Hastie, Trevor, Robert Tibshirani et JH Friedman. 2003. « The elements of statistical learning, corrected edition ». Berlin: Springer. Haxby, JV, Gobbini, MI, Furey, ML, Ishai, A., Schouten, JL, & Pietrini, P.(2001). *Distributed and overlapping representations of faces and objects in ventral temporal cortex*. *Science*, vol. 293, n° 5539, p. 24252430.
- Hatziaargyriou, N, E Karapidakis et D Hatzifotis. 1998. « Frequency stability of power systems in large islands with high wind power penetration ». In *Bulk Power Syst. Dynamics Control Symp.—IV Restructuring*. Vol. 102.
- Hawkins, Douglas M, Subhash C Basak et Denise Mills. 2003. « Assessing model fit by cross-validation ». *Journal of chemical information and computer sciences*, vol. 43, n° 2, p. 579-586.
- He, L., et C. C. Liu. 2014. « Parameter Identification With PMUs for Instability Detection in Power Systems With HVDC Integrated Offshore Wind Energy ». *IEEE Transactions on Power Systems*, vol. 29, n° 2, p. 775-784.
- Hill, David J, MK PAL, XU WILSON, YAKOUT MANSOUR, CO NWANKPA, L Xu et R Fischl. 1993. « Nonlinear dynamic load models with recovery for voltage stability studies. discussion. authors' response ». *IEEE transactions on power systems*, vol. 8, n° 1, p. 166-176.

- Horowitz, S. H., A. G. Phadke et J. S. Thorpe. 1988. « Adaptive transmission system relaying ». *IEEE Transactions on Power Delivery*, vol. 3, n° 4, p. 1436-1445.
- Huang, Z., P. Du, D. Kosterev et S. Yang. 2013. « Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection ». *IEEE Transactions on Power Systems*, vol. 28, n° 2, p. 1939-1949.
- Hui, Deng, Zhao Jinquan, Wu Xiaochen et Men Kun. 2012. « Real time transient instability detection based on trajectory characteristics and transient energy ». In *Power and Energy Society General Meeting, 2012 IEEE*. (22-26 July 2012), p. 1-7.
- Hyafil, Laurent, et Ronald L Rivest. 1976. « Constructing optimal binary decision trees is NP-complete ». *Information Processing Letters*, vol. 5, n° 1, p. 15-17.
- IEEE. « New England 10 Generators 39 Busses Test System ». < <http://sys.elec.kitami-it.ac.jp/ueda/demo/WebPF/39-New-England.pdf> >. Consulté le 2 avril 2014.
- IEEE, et Standard. 2006. « IEEE Recommended Practice for Excitation System Models for Power System Stability Studies ». *IEEE Std 421.5-2005 (Revision of IEEE Std 421.5-1992)*, p. 1-93.
- IEEE Standard for Synchrophasors for Power Systems. 1995. « ». *IEEE Std 1344-1995(R2001)*, p. i.
- IEEE Standard for Synchrophasor Data Transfer for Power Systems. 2011. « ». *IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005)*, p. 1-53.
- IEEE Standard for Synchrophasor Measurements for Power Systems. 2011. « ». *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, p. 1-61.
- IEEE Standard for Synchrophasors for Power Systems. 2006. « ». *IEEE Std C37.118-2005 (Revision of IEEE Std 1344-1995)*, p. 0_1-57.
- Ihaka, Ross, et Robert Gentleman. 1996. « R: a language for data analysis and graphics ». *Journal of computational and graphical statistics*, vol. 5, n° 3, p. 299-314.
- James, Gareth, Daniela Witten, Trevor Hastie et Robert Tibshirani. 2013. *An introduction to statistical learning*, 6. Springer.
- Jiang, W., V. Vittal et G. T. Heydt. 2007. « A Distributed State Estimator Utilizing Synchronized Phasor Measurements ». *IEEE Transactions on Power Systems*, vol. 22, n° 2, p. 563-571.

- Kaci, A., I. Kamwa, L. A. Dessaint et S. Guillon. 2014a. « Synchrophasor Data Baseline and Mining for Online Monitoring of Dynamic Security Limits ». *Power Systems, IEEE Transactions on*, vol. PP, n° 99, p. 1-15.
- Kaci, A., I. Kamwa, L. A. Dessaint et S. Guillon. 2014b. « Synchrophasor Data Baseline and Mining for Online Monitoring of Dynamic Security Limits ». *Power Systems, IEEE Transactions on*, vol. 29, n° 6, p. 2681-2695.
- Kai, Sun, S. T. Lee et Zhang Pei. 2011. « An Adaptive Power System Equivalent for Real-Time Estimation of Stability Margin Using Phase-Plane Trajectories ». *Power Systems, IEEE Transactions on*, vol. 26, n° 2, p. 915-923.
- Kai, Sun, S. Likhate, V. Vittal, V. S. Kolluri et S. Mandal. 2007. « An Online Dynamic Security Assessment Scheme Using Phasor Measurements and Decision Trees ». *Power Systems, IEEE Transactions on*, vol. 22, n° 4, p. 1935-1943.
- Kalyani, S., et K. S. Swarup. 2011. « Classification and Assessment of Power System Security Using Multiclass SVM ». *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 41, n° 5, p. 753-758.
- Kamwa, I., J. Beland et D. Mcnabb. 2006. « PMU-Based Vulnerability Assessment Using Wide-Area Severity Indices and Tracking Modal Analysis ». In *2006 IEEE PES Power Systems Conference and Exposition*. (Oct. 29 2006-Nov. 1 2006), p. 139-149.
- Kamwa, I., J. Beland, G. Trudel, R. Grondin, C. Lafond et D. McNabb. 2006. « Wide-area monitoring and control at Hydro-Quebec: past, present and future ». In *2006 IEEE Power Engineering Society General Meeting*. (0-0 0), p. 12 pp.
- Kamwa, I., et R. Grondin. 2002a. « PMU configuration for system dynamic performance measurement in large multi-area power systems ». In *Power Engineering Society Summer Meeting, 2002 IEEE*. (25-25 July 2002) Vol. 1, p. 239 vol.1.
- Kamwa, I., R. Grondin et L. Loud. 2001a. « Time-varying contingency screening for dynamic security assessment using intelligent-systems techniques ». In *Power Industry Computer Applications, 2001. PICA 2001. Innovative Computing for Power - Electric Energy Meets the Market. 22nd IEEE Power Engineering Society International Conference on*. (20-24 May 2001), p. 196-200.
- Kamwa, I., R. Grondin et L. Loud. 2001b. « Time-varying contingency screening for dynamic security assessment using intelligent-systems techniques ». *Power Systems, IEEE Transactions on*, vol. 16, n° 3, p. 526-536.
- Kamwa, I., A. K. Pradhan, G. Joos et S. R. Samantaray. 2009a. « Fuzzy Partitioning of a Real Power System for Dynamic Vulnerability Assessment ». *IEEE Transactions on Power Systems*, vol. 24, n° 3, p. 1356-1365.

- Kamwa, I., S. R. Samantaray et G. Joos. 2009. « Development of Rule-Based Classifiers for Rapid Stability Assessment of Wide-Area Post-Disturbance Records ». *Power Systems, IEEE Transactions on*, vol. 24, n° 1, p. 258-270.
- Kamwa, I., S. R. Samantaray et G. Joos. 2012. « On the Accuracy Versus Transparency Trade-Off of Data-Mining Models for Fast-Response PMU-Based Catastrophe Predictors ». *Smart Grid, IEEE Transactions on*, vol. 3, n° 1, p. 152-161.
- Kamwa, I., S. R. Samantaray et G. Joos. 2013. « Compliance Analysis of PMU Algorithms and Devices for Wide-Area Stabilizing Control of Large Power Systems ». *IEEE Transactions on Power Systems*, vol. 28, n° 2, p. 1766-1778.
- Kamwa, Innocent, et Robert Grondin. 2002b. « PMU configuration for system dynamic performance measurement in large, multiarea power systems ». *IEEE Transactions on Power Systems*, vol. 17, n° 2, p. 385-394.
- Kamwa, Innocent, Ashok Kumar Pradhan et Géza Joós. 2007. « Automatic segmentation of large power systems into fuzzy coherent areas for dynamic vulnerability assessment ». *IEEE Transactions on Power Systems*, vol. 22, n° 4, p. 1974-1985.
- Kamwa, Innocent, Ashok Kumar Pradhan, Geza Joos et SR Samantaray. 2009b. « Fuzzy partitioning of a real power system for dynamic vulnerability assessment ». *IEEE Transactions on Power Systems*, vol. 24, n° 3, p. 1356-1365.
- Karady, George G, Ahmed A Daoud et Mansour A Mohamed. 2002. « On-line transient stability enhancement using multi-agent technique ». In *Power Engineering Society Winter Meeting, 2002. IEEE*. Vol. 2, p. 893-899. IEEE.
- Kass, Gordon V. 1980. « An exploratory technique for investigating large quantities of categorical data ». *Applied statistics*, p. 119-127.
- Kearns, Michael. 1988. « Thoughts on hypothesis boosting ». *Unpublished manuscript*, vol. 45, p. 105.
- Kearns, Michael, et Leslie Valiant. 1994. « Cryptographic limitations on learning Boolean formulae and finite automata ». *Journal of the ACM (JACM)*, vol. 41, n° 1, p. 67-95.
- Keller, J. M., M. R. Gray et J. A. Givens. 1985. « A fuzzy K-nearest neighbor algorithm ». *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, n° 4, p. 580-585.
- Khalil, Hassan K. 2002. *Nonlinear systems* (2002), 3rd ed. Upper Saddle River, N.J.: Prentice Hall, xv, 750 p. p.
- Khan, Laiq, Tariq Saeed et KL Lo. 2008. « Robust damping control system design for TCSC using particle swarm optimization ». *COMPEL-The international journal for*

- computation and mathematics in electrical and electronic engineering*, vol. 27, n° 3, p. 593-612.
- Khoi; Vu, Damir Novosel; et Bill Synder. 2007. *A business case study on applying synchrophasor measurement technology and applications in the California and the western electricity coordinating council grid.*, 500-99-013 BOA-130. California, USA: KEMA, Inc. <<http://www.energy.ca.gov/2013publications/CEC-500-2013-045/CEC-500-2013-045.pdf>>. Consulté le 21 septembre 2016.
- Kim, T. T., et H. V. Poor. 2011. « Strategic Protection Against Data Injection Attacks on Power Grids ». *IEEE Transactions on Smart Grid*, vol. 2, n° 2, p. 326-333.
- Kimbark, Edward Wilson. 1995. *Power system stability*, 1. John Wiley & Sons.
- Kohavi, Ronny, et J Ross Quinlan. 2002a. « Data mining tasks and methods: Classification: decision-tree discovery ». In *Handbook of data mining and knowledge discovery*. p. 267-276. Oxford University Press, Inc.
- Kohavi, Ronny, et J. Ross Quinlan. 2002b. « Data mining tasks and methods: Classification: decision-tree discovery ». In *Handbook of data mining and knowledge discovery*, sous la dir. de Willi, Kl, sgen et M. Zytkow Jan. p. 267-276. Oxford University Press, Inc.
- Korkali, Mert, et Ali Abur. 2010. « Impact of network sparsity on strategic placement of phasor measurement units with fixed channel capacity ». In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. p. 3445-3448. IEEE.
- Krishna, S., et K. R. Padiyar. 2000. « Transient stability assessment using artificial neural networks ». In *Industrial Technology 2000. Proceedings of IEEE International Conference on*. (19-22 Jan. 2000) Vol. 1, p. 627-632 vol.2.
- Kundur, Prabha, Neal J Balu et Mark G Lauby. 1994. *Power system stability and control*, 7. McGraw-hill New York.
- Kundur, Prabha, John Paserba, Venkat Ajjarapu, Göran Andersson, Anjan Bose, Claudio Canizares, Nikos Hatziargyriou, David Hill, Alex Stankovic et Carson Taylor. 2004. « Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions ». *IEEE transactions on Power Systems*, vol. 19, n° 3, p. 1387-1401.
- Larose, Daniel T., et Chantal D. Larose. 2014. *Discovering knowledge in data an introduction to data mining* (2014), 2nd ed. Hoboken, N.J.: John Wiley & Sons, 1 ressource en ligne. p.
- Lee, Kwang Y, et Mohamed A El-Sharkawi. 2008. *Modern heuristic optimization techniques: theory and applications to power systems*, 39. John Wiley & Sons.

- Li, D. H., et Y. J. Cao. 2005. « SOFM based support vector regression model for prediction and its application in power system transient stability forecasting ». In *Power Engineering Conference, 2005. IPEC 2005. The 7th International*. (Nov. 29 2005-Dec. 2 2005), p. 765-770 Vol. 2.
- Li, Meiyang, Anamitra Pal, Arun G. Phadke et James S. Thorp. 2014. « Transient stability prediction based on apparent impedance trajectory recorded by PMUs ». *International Journal of Electrical Power & Energy Systems*, vol. 54, p. 498-504.
- Liacco, TE Dy. 1974. « Real-time computer control of power systems ». *Proceedings of the IEEE*, vol. 62, n° 7, p. 884-891.
- Liancheng, Wang, et A. A. Girgis. 1997. « A new method for power system transient instability detection ». *Power Delivery, IEEE Transactions on*, vol. 12, n° 3, p. 1082-1089.
- Liang, H., H. Zhang et Y. Yan. 2006. « Decision Trees for Probability Estimation: An Empirical Study ». In *2006 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06)*. (Nov. 2006), p. 756-764.
- Liaw, Andy, et Matthew Wiener. 2002. « Classification and regression by randomForest ». *R news*, vol. 2, n° 3, p. 18-22.
- Liu, C-W, et J Thorp. 1995. « Application of synchronised phasor measurements to real-time transient stability prediction ». *IEE Proceedings-Generation, Transmission and Distribution*, vol. 142, n° 4, p. 355-360.
- Liu, Chen-Ching, et Kevin Tomsovic. 1986. « An expert system assisting decision-making of reactive power/voltage control ». *IEEE Transactions on Power Systems*, vol. 1, n° 3, p. 195-201.
- Liu, Chih-Wen, Mu-Chun Su, Shuenn-Shing Tsay et Yi-Jen Wang. 1999a. « Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements ». *Power Systems, IEEE Transactions on*, vol. 14, n° 2, p. 685-692.
- Liu, Chih-Wen, Mu-Chun Su, Shuenn-Shing Tsay et Yi-Jen Wang. 1999b. « Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements ». *IEEE Transactions on Power Systems*, vol. 14, n° 2, p. 685-692.
- Liu, Chih-Wen, et James S Thorp. 2000. « New methods for computing power system dynamic response for real-time transient stability prediction ». *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, n° 3, p. 324-337.

- Liu, Chih-Wen, Shuenn-Shing Tsay, Yi-Jen Wang et Mu-Chun Su. 1999c. « Neuro-fuzzy approach to real-time transient stability prediction based on synchronized phasor measurements ». *Electric Power Systems Research*, vol. 49, n° 2, p. 123-127.
- Liu, X., Q. Jiang et Y. Cao. 2008. « Novel fast transient stability prediction method based on perturbed trajectories fitting of rotor angle ». *Dianli Xitong Zidonghua/Automation of Electric Power Systems*, vol. 32, n° 19, p. 5-9.
- Liu, Yutian, et Fei Lin. 2000. « Application of PMU and fuzzy radial basis function network to power system transient stability prediction ». *Zhongguo Dianji Gongcheng Xuebao/Proceedings of the Chinese Society of Electrical Engineering*, vol. 20, n° 2, p. 19-23.
- Llamas, A, et J De La Ree. 1993. « Stability and the transient energy method for the classroom ». In *System Theory, 1993. Proceedings SSST'93., Twenty-Fifth Southeastern Symposium on*. p. 79-85. IEEE.
- Llamas, A, J De La Ree Lopez, L Mili, AG Phadke et JS Thorp. 1995. « Clarifications of the BCU method for transient stability analysis ». *IEEE Transactions on Power Systems*, vol. 10, n° 1, p. 210-219.
- Ma, Jian, Yuri Makarov et Zhaoyang Dong. 2010. « Phasor Measurement Unit and Its Application in Modern Power Systems ». In *Emerging Techniques in Power System Analysis*. Ma2010. p. 147-184. Berlin, Heidelberg: Springer Berlin Heidelberg. < http://dx.doi.org/10.1007/978-3-642-04282-9_6 >. Consulté le 2 août 2016.
- Machias, AV. 1989. « Transient stability evaluation by a pattern recognition approach using multivalent logics ». *Electric power systems research*, vol. 17, n° 3, p. 209-217.
- Machias, AV, et JL Souflis. 1990. « A fuzzy approach in the determination of unstable machines ». In *IEE Proceedings C-Generation, Transmission and Distribution*. Vol. 137, p. 115-122. IET.
- Machowski, Jan, Janusz Bialek et Jim Bumby. 2011. *Power system dynamics: stability and control*. John Wiley & Sons.
- Madani, Vahid, Damir Novosel et Roger King. 2014. « Wide-Area Monitoring, Protection and Control Needs, Applications, and Benefits ». *Computational Intelligence Applications in Smart Grids: Enabling Methodologies for Proactive and Self-Organizing Power Systems*, p. 1.
- Maimon, Oded, et Lior Rokach. 2005. *Data mining and knowledge discovery handbook*, 2. Springer.

- Mansour, Y., A. Y. Chang, J. Tamby, E. Vaahedi, B. R. Corns et M. A. El-Sharkawi. 1997. « Large scale dynamic security screening and ranking using neural networks ». *Power Systems, IEEE Transactions on*, vol. 12, n° 2, p. 954-960.
- Martin, K. E. 2011. « Synchrophasor Standards Development - IEEE C37.118 & IEC 61850 ». In *System Sciences (HICSS), 2011 44th Hawaii International Conference on*. (4-7 Jan. 2011), p. 1-8.
- Mauch, Konrad, et Aidan Foss. 2005. « Smart grid technology overview ». *Natural Resources, Canada*.
- McCalley, James D, AA Fouad, Vijay Vittal, AA Irizarry-Rivera, BL Agrawal, RG Farmer, R BILLINTON, S ABORESHAID et M FOTUHI-FIRUZABAD. 1997. « A risk-based security index for determining operating limits in stability-limited electric power systems. Discussion ». *IEEE Transactions on Power Systems*, vol. 12, n° 3, p. 1210-1219.
- Meyer, David, Friedrich Leisch et Kurt Hornik. 2003. « The support vector machine under test ». *Neurocomputing*, vol. 55, n° 1, p. 169-186.
- Milano, Federico. 2010. *Power system modelling and scripting*. Springer Science & Business Media.
- Milosevic, Borka, et Miroslav Begovic. 2003. « Nondominated sorting genetic algorithm for optimal phasor measurement placement ». *IEEE Transactions on Power Systems*, vol. 18, n° 1, p. 69-75.
- Momoh, James A. 2012. *Smart grid fundamentals of design and analysis* (2012). Hoboken: John Wiley & Sons, 1 texte électronique (234 p.) p.
- Morison, GK, B Gao et P Kundur. 1993. « Voltage stability analysis using static and dynamic approaches ». *IEEE Transactions on Power Systems*, vol. 8, n° 3, p. 1159-1171.
- Morison, Kip, Lei Wang et Prabha Kundur. 2004. « Power system security assessment ». *IEEE Power and Energy Magazine*, vol. 2, n° 5, p. 30-39.
- Morren, J., S. W. H. de Haan et J. A. Ferreira. 2005. « Contribution of DG units to primary frequency control ». In *2005 International Conference on Future Power Systems*. (18-18 Nov. 2005), p. 6 pp.-6.
- Moulin, L. S., A. P. Alves da Silva, M. A. El-Sharkawi et R. J. Marks, II. 2002. « Support vector and multilayer perceptron neural networks applied to power systems transient stability analysis with input dimensionality reduction ». In *Power Engineering Society Summer Meeting, 2002 IEEE*. (25-25 July 2002) Vol. 3, p. 1308-1313 vol.3.

- Moulin, L. S., A. P. Alves da Silva, M. A. El-Sharkawi et R. J. Marks, II. 2004. « Support vector machines for transient stability analysis of large-scale power systems ». *Power Systems, IEEE Transactions on*, vol. 19, n° 2, p. 818-825.
- Muenchen, Robert A. 2011. *R for SAS and SPSS users*. Springer Science & Business Media.
- Muenchen, Robert A, et Joseph Hilbe. 2010. *R for Stata users*. Springer.
- Naik, Prem Kumar, Nirmal-Kumar C. Nair et Akshya Kumar Swain. 2016. « Impact of reduced inertia on transient stability of networks with asynchronous generation ». *International Transactions on Electrical Energy Systems*, vol. 26, n° 1, p. 175-191.
- Narasimhamurthi, N., et M. Musavi. 1984. « A generalized energy function for transient stability analysis of power systems ». *IEEE Transactions on Circuits and Systems*, vol. 31, n° 7, p. 637-645.
- NASPI. 2007. *A Guide for PMU Installation, Commissioning and Maintenance Part II PMU Installation Procedures*. North American SynchroPhasor Initiative (NASPI). < <https://www.naspi.org/File.aspx?fileID=557> >. Consulté le 09 août 2016.
- NASPI. 2016. « North American SynchroPhasor Initiative ». < <https://www.naspi.org/> >. Consulté le 09 août 2016.
- Nayak, Neeraj, Heng Chen, Wayne Schmus et Ryan Quint. 2016. « Generator parameter validation and calibration process based on PMU data ». In *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, p. 1-5. IEEE.
- Nisbet, Robert, Gary Miner et John Elder IV. 2009. *Handbook of statistical analysis and data mining applications*. Academic Press.
- Norio, Okada, Tao Ye, Yoshio Kajitani, Peijun Shi et Hirokazu Tatano. 2011. « The 2011 eastern Japan great earthquake disaster: Overview and comments ». *International Journal of Disaster Risk Science*, vol. 2, n° 1, p. 34-42.
- Novosel, Damir, Vahid Madani, Bharat Bhargava, Khoi Vu et Jim Cole. 2008. « Dawn of the grid synchronization ». *IEEE Power and Energy Magazine*, vol. 1, n° 6, p. 49-60.
- Nuqui, R. F., A. G. Phadke, R. P. Schulz et N. Bhatt. 2001. « Fast on-line voltage security monitoring using synchronized phasor measurements and decision trees ». In *Power Engineering Society Winter Meeting, 2001. IEEE*. (2001) Vol. 3, p. 1347-1352 vol.3.
- Oates, Tim, et David Jensen. 1997. « The Effects of Training Set Size on Decision Tree Complexity ». In *Proceedings of the Fourteenth International Conference on Machine Learning*. p. 254-262. 657136: Morgan Kaufmann Publishers Inc.

- Opitz, David, et Richard Maclin. 1999. « Popular ensemble methods: An empirical study ». *Journal of Artificial Intelligence Research*, vol. 11, p. 169-198.
- Pai, Anantha. 2012. *Energy function analysis for power system stability*. Springer Science & Business Media.
- Pai, M.A. 1989. *Energy function analysis for power system stability*. Medium: X; Size: Pages: (256 p) p.
- Pal, Anamitra, James S Thorp, Santosh S Veda et VA Centeno. 2013. « Applying a robust control technique to damp low frequency oscillations in the WECC ». *International Journal of Electrical Power & Energy Systems*, vol. 44, n° 1, p. 638-645.
- Park, RH, et EH Bancker. 1929. « System stability as a design problem ». *Transactions of the American Institute of Electrical Engineers*, vol. 48, n° 1, p. 170-193.
- Paserba, J. 1996. « Analysis and control of power system oscillation ». *CIGRE special publication*, vol. 38, n° 07.
- Patel, M. 2010. *Eastern Interconnection Wide Area SynchroPhasor Angles Baseline Study*. <<https://www.naspi.org/File.aspx?fileID=519>>. Consulté le 11 août 2016.
- Pavella, Mania, Damien Ernst et Daniel Ruiz-Vega. 2012. *Transient stability of power systems: a unified approach to assessment and control*. Springer Science & Business Media.
- Phadke, Arun G. 2002. « Synchronized phasor measurements-a historical overview ». In *Transmission and Distribution Conference and Exhibition 2002: Asia Pacific*. IEEE/PES. Vol. 1, p. 476-479. IEEE.
- Phadke, Arun G, et James S Thorp. 2009. *Computer relaying for power systems*. John Wiley & Sons.
- Phadke, Arun G, et John Samuel Thorp. 2008. *Synchronized phasor measurements and their applications*. Springer Science & Business Media.
- Portier, Natacha. 1998. « Résolutions universelles pour des problèmes NP-complets ». *Theoretical Computer Science*, vol. 201, n° 1, p. 137-150.
- Powers, David Martin. 2011. « Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation ».
- Quinlan, J Ross. 2014. *C4. 5: programs for machine learning*. Elsevier.
- Quinlan, J. Ross. 1986. « Induction of decision trees ». *Machine learning*, vol. 1, n° 1, p. 81-106.

- Qun, Gao, et S. M. Rovnyak. 2011. « Decision Trees Using Synchronized Phasor Measurements for Wide-Area Response-Based Control ». *Power Systems, IEEE Transactions on*, vol. 26, n° 2, p. 855-861.
- Rahmatian, Matin, Yu Christine Chen, Atefeh Palizban, Ali Moshref et William G. Dunford. 2017. « Transient stability assessment via decision trees and multivariate adaptive regression splines ». *Electric Power Systems Research*, vol. 142, p. 320-328.
- Rajapakse, A. D., F. Gomez, K. Nanayakkara, P. A. Crossley et V. V. Terzija. 2010a. « Rotor Angle Instability Prediction Using Post-Disturbance Voltage Trajectories ». *Power Systems, IEEE Transactions on*, vol. 25, n° 2, p. 947-956.
- Rajapakse, Athula D, Francisco Gomez, Kasun Nanayakkara, Peter A Crossley et Vladimir V Terzija. 2010b. « Rotor angle instability prediction using post-disturbance voltage trajectories ». *IEEE Transactions on Power Systems*, vol. 25, n° 2, p. 947-956.
- Rakpenthai, Chawasaki, Suttichai Premrudeepreechacharn, Sermsak Uatrongjit et Neville R Watson. 2007. « An optimal PMU placement method against measurement loss and branch outage ». *IEEE Transactions on Power Delivery*, vol. 22, n° 1, p. 101-107.
- Rampurkar, V., P. Pentayya, H. A. Mangalvedekar et F. Kazi. 2016. « Cascading Failure Analysis for Indian Power Grid ». *IEEE Transactions on Smart Grid*, vol. 7, n° 4, p. 1951-1960.
- Ray, Dennis. 2003. « Blackout of 2003: Description and responses ». *Power Systems Engineering Research Centre (PSERC)*.
- Ribbens-Pavella, M, et FJ Evans. 1985. « Direct methods for studying dynamics of large-scale electric power systems—A survey ». *Automatica*, vol. 21, n° 1, p. 1-21.
- Rogers, Graham. 2012. *Power system oscillations*. Springer Science & Business Media.
- Rouche, Nicolas, Patrick Habets et Michel Laloy. 1977. *Stability theory by Liapunov's direct method*, 4. Springer.
- Rovnyak, S., S. Kretsinger, J. Thorp et D. Brown. 1994a. « Decision trees for real-time transient stability prediction ». *IEEE Transactions on Power Systems*, vol. 9, n° 3, p. 1417-1426.
- Rovnyak, S., S. Kretsinger, J. Thorp et D. Brown. 1994b. « Decision trees for real-time transient stability prediction ». *Power Systems, IEEE Transactions on*, vol. 9, n° 3, p. 1417-1426.

- Roy, B. K. S., A. K. Sinha et A. K. Pradhan. 2013. « Synchrophasor-assisted Prediction of Stability/Instability of a Power System ». *International Journal of Emerging Electric Power Systems*, vol. 14, n° 1, p. 7 pp.
- Sastry, S Shankar. 2013. *Nonlinear systems: analysis, stability, and control*, 10. Springer Science & Business Media.
- Sauer, Peter W., et M. A. Pai. 1998. *Power system dynamics and stability* (1998). Upper Saddle River, N.J.: Prentice-Hall, x, 357 p. p.
- Savulescu, Savu. 2014. *Real-time stability in power systems: techniques for early detection of the risk of blackout*. Springer.
- Savulescu, Savu Crivat. 2009. *Real-time stability assessment in modern power system control centers* (2009). Hoboken, N.J.: John Wiley & Sons, Inc., 1 texte électronique (xxviii, 425 p.) p.
- Schapire, Robert E. 1990. « The strength of weak learnability ». *Machine learning*, vol. 5, n° 2, p. 197-227.
- Schapire, Robert E. 1999. « A brief introduction to boosting ». In *Ijcai*. Vol. 99, p. 1401-1406.
- Sharma, C., et B. Tyagi. 2013. « Transient stability status predictor based on relevance vector machine ». In *Innovative Smart Grid Technologies - Asia (ISGT Asia), 2013 IEEE*. (10-13 Nov. 2013), p. 1-6.
- Silverman, Bernard W. 1986. *Density estimation for statistics and data analysis*, 26. CRC press.
- Simonoff, Jeffrey S. 2012. *Smoothing methods in statistics*. Springer Science & Business Media.
- Smith, MM, RS Powell, MR Irving et MJH Sterling. 1991. « Robust algorithm for state estimation in electrical networks ». In *IEE Proceedings C-Generation, Transmission and Distribution*. Vol. 138, p. 283-288. IET.
- Sobajic, D. J., et Y. H. Pao. 1989. « Artificial neural-net based dynamic security assessment for electric power systems ». *Power Systems, IEEE Transactions on*, vol. 4, n° 1, p. 220-228.
- Song, Fangfang, Tianshu Bi et Qixun Yang. 2006. « Perturbed trajectory prediction method based on wide area measurement systems ». *Automation of Electric Power Systems*, vol. 23, p. 006.

- Soni, N., S. Doolla et M. C. Chandorkar. 2013. « Improvement of Transient Response in Microgrids Using Virtual Inertia ». *IEEE Transactions on Power Delivery*, vol. 28, n° 3, p. 1830-1838.
- Souflis, JL, AV Machias et BC Papadias. 1989. « An application of fuzzy concepts to transient stability evaluation [of power systems] ». *IEEE Transactions on Power Systems*, vol. 4, n° 3, p. 1003-1009.
- Stagg, Glenn W, et Ahmed H El-Abiad. 1968. *Computer methods in power system analysis*. McGraw-Hill.
- Stanton, S. E., C. Slivinsky, K. Martin et J. Nordstrom. 1995. « Application of phasor measurements and partial energy analysis in stabilizing large disturbances ». *Power Systems, IEEE Transactions on*, vol. 10, n° 1, p. 297-306.
- Steinmetz, Charles P. 1920. « Power control and stability of electric generating stations ». *Transactions of the American Institute of Electrical Engineers*, vol. 39, n° 2, p. 1215-1287.
- Steinmetz, Charles Proteus. 1893. « Complex quantities and their use in electrical engineering ». In *Proceedings of the International Electrical Congress, Chicago*. p. 33-74.
- Stéphane, Tufféry. 2012. *Data Mining et statistique décisionnelle: L'intelligence des données*. Éditions Technip.
- Strobl, Carolin, James Malley et Gerhard Tutz. 2009. « An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests ». *Psychological methods*, vol. 14, n° 4, p. 323.
- Sun, K., S. Likhate, V. Vittal, V. S. Kolluri et S. Mandal. 2007. « An Online Dynamic Security Assessment Scheme Using Phasor Measurements and Decision Trees ». *IEEE Transactions on Power Systems*, vol. 22, n° 4, p. 1935-1943.
- Suprême, Hussein. 2012. « Analyse et classement des contingences d'un réseau électrique pour la stabilité transitoire ». École de technologie supérieure.
- Suprême, Hussein, et Louis- A. Dessaint. 2012. « Analyse et classement des contingences d'un réseau électrique pour la stabilité transitoire ». École de technologie supérieure. <http://espace.etsmtl.ca/1100/1/SUPR%C3%8AME_Hussein.pdf>.
- Suprême, Hussein, Louis-A Dessaint, Innocent Kamwa et Annissa Heniche-Oussédik. 2016. « Development of New Predictors Based on the Concept of Center of Power for Transient and Dynamic Instability Detection ». *IEEE Transactions on Smart Grid, accepted paper for publication*.

- Suprême, Hussein, Annissa Heniche-Oussédik, Innocent Kamwa et Louis-A. Dessaint. 2016. « A novel approach for instability detection based on wide-area measurements and new predictors ». In *Conference on Electrical and Computer Engineering*. (Vancouver, British Columbia, Canada).
- Susuki, Yoshihiko, Igor Mezić et Takashi Hikihara. 2011. « Coherent swing instability of power grids ». *Journal of nonlinear science*, vol. 21, n° 3, p. 403-439.
- Svetnik, Vladimir, Andy Liaw, Christopher Tong, J Christopher Culberson, Robert P Sheridan et Bradley P Feuston. 2003. « Random forest: a classification and regression tool for compound classification and QSAR modeling ». *Journal of chemical information and computer sciences*, vol. 43, n° 6, p. 1947-1958.
- Tao, Xu, RM He, Wang Peng et DJ Xu. 2004. « Phasor measurement placement for transient stability analysis of power systems ». In *Developments in Power System Protection, 2004. Eighth IEE International Conference on*. Vol. 1, p. 28-31. IET.
- Theodorakatos, Nikolaos, Nikolaos Manousakis et George Korres. 2015. « Optimal PMU Placement Using Nonlinear Programming ». *arXiv preprint arXiv:1507.05258*.
- Tin Kam, Ho. 1998. « The random subspace method for constructing decision forests ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, n° 8, p. 832-844.
- Trudel, Gilles, Jean-Pierre Gingras et Jean-Robert Pierre. 2005. « Designing a reliable power system: Hydro-Quebec's integrated approach ». *Proceedings of the IEEE*, vol. 93, n° 5, p. 907-917.
- Tufféry, Stéphane. 2005. *Data mining et statistique décisionnelle: l'intelligence dans les bases de données*. Editions Technip.
- Tuffery, Stéphane, et Rod Riesco. 2011. *Data mining and statistics for decision making* (2011). Chichester, West Sussex, U.K.: Wiley.
- Vaahedi, E., W. Li, T. Chia et H. Dommel. 2000. « Large scale probabilistic transient stability assessment using BC Hydro's on-line tool ». *IEEE Transactions on Power Systems*, vol. 15, n° 2, p. 661-667.
- Van Cutsem, Thierry, et Richard Mailhot. 1997. « Validation of a fast voltage stability analysis method on the Hydro-Quebec system ». *IEEE Transactions on Power Systems*, vol. 12, n° 1, p. 282-292.
- Van Cutsem, Thierry, et Costas Vournas. 1998. *Voltage stability of electric power systems*, 441. Springer Science & Business Media.

- Vapnik, Vladimir Naumovich. 2000. *The nature of statistical learning theory*, 2nd ed.. New York: Springer, xix, 314 p. p.
- Vorotnikov, Vladimir Il'ich. 2012. *Partial stability and control*. Springer Science & Business Media.
- Wallace, Chris S., et JD Patrick. 1993. « Coding decision trees ». *Machine Learning*, vol. 11, n° 1, p. 7-22.
- Wehenkel, Louis, Mania Pavella, Edwige Euxibie et Bertrand Heilbronn. 1994. « Decision tree based transient stability method a case study ». *IEEE Transactions on Power Systems*, vol. 9, n° 1, p. 459-469.
- Williams, Graham. 2011. *Data mining with rattle and R: the art of excavating data for knowledge discovery*. Springer Science & Business Media.
- Witten, I. H., Mark A. Hall et Eibe Frank. 2011. *Data mining : practical machine learning tools and techniques* (2011), 3rd ed. Burlington, Mass.: Morgan Kaufmann, xxxiii, 629 p. p.
- Xie, Huan, Bao-hui Zhang, Guang-liang Yu, Ying-hui Li, Peng Li, Decai Zhou et Feng Yao. 2005. « Power System Transient Stability Detection Based on Characteristic Concave or Convex of Trajectory ». In *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*. (2005), p. 1-6.
- Xue, Y, Z Gao, J Zhu et J Liu. 1989. « A planning decision support expert system for transient and dynamic security of power systems ». In *Proceedings on the Second Symposium on Expert Systems Application to Power Systems*. p. 218-224.
- Xue, Y., T. Van Cutsem et M. Ribbens-Pavella. 1988. « A simple direct method for fast transient stability assessment of large power systems ». *IEEE Transactions on Power Systems*, vol. 3, n° 2, p. 400-412.
- Xue, Y., T. Huang, K. Li, Z. Dong, D. Yue, F. Xue et J. Huang. 2015. « An efficient and robust case sorting algorithm for transient stability assessment ». In *2015 IEEE Power & Energy Society General Meeting*. (26-30 July 2015), p. 1-5.
- Xue, Yusheng, Louis Wehenkel, Regine Belhomme, Patricia Rousseaux, Mania Pavella, Edwige Euxibie, Bertrand Heilbronn et J-F Lesigne. 1992. « Extended equal area criterion revisited [EHV power systems] ». *IEEE Transactions on Power Systems*, vol. 7, n° 3, p. 1012-1022.
- Yager, Ronald R, et Lotfi A Zadeh. 2012. *An introduction to fuzzy logic applications in intelligent systems*, 165. Springer Science & Business Media.

- Yuan-Yih, Hsu, et Chang Chung-Liang. 1988. « Probabilistic transient stability studies using the conditional probability approach ». *IEEE transactions on power systems*, vol. 3, n° 4, p. 1565-1572.
- Yusof, SB, GJ Rogers et RTH Alden. 1993. « Slow coherency based network partitioning including load buses ». *Power Systems, IEEE Transactions on*, vol. 8, n° 3, p. 1375-1382.
- Zabaiou, Tarik, Louis-A Dessaint et Innocent Kamwa. 2014. « Preventive control approach for voltage stability improvement using voltage stability constrained optimal power flow based on static line voltage stability indices ». *IET Generation, Transmission & Distribution*, vol. 8, n° 5, p. 924-934.
- Zamora-C, E. A., x00E, rdenas et C. R. Fuerte-Esquivel. 2012. « Static state estimation of FACTS containing synchronized phasor measurement units ». In *2012 IEEE Power and Energy Society General Meeting*. (22-26 July 2012), p. 1-7.
- Zerigui, A., L. A. Dessaint, R. Hannat, R. T. F. Ah King et I. Kamwa. 2015. « Statistical approach for transient stability constrained optimal power flow ». *IET Generation, Transmission & Distribution*, vol. 9, n° 14, p. 1856-1864.
- Zerigui, Amel. 2015. *Répartition optimale de la production électrique avec une contrainte de stabilité transitoire établie par une approche statistique* (2015). Montréal: École de technologie supérieure, 1 ressource en ligne (136 pages) p.
- Zhang, Y., E. Muljadi, D. Kosterev et M. Singh. 2015. « Wind Power Plant Model Validation Using Synchrophasor Measurements at the Point of Interconnection ». *IEEE Transactions on Sustainable Energy*, vol. 6, n° 3, p. 984-992.
- Zhang, Yang, et Anjan Bose. 2008. « Design of wide-area damping controllers for interarea oscillations ». *IEEE Transactions on Power Systems*, vol. 23, n° 3, p. 1136-1143.
- Zheng, C., V. Malbasa et M. Kezunovic. 2013. « Regression tree for stability margin prediction using synchrophasor measurements ». *IEEE Transactions on Power Systems*, vol. 28, n° 2, p. 1978-1987.
- Zhonghong, Yu, Zhou Xiaoxin et Wu Zhongxi. 2005. « Fast Transient Stability Assessment Based on Data Mining for Large-Scale Power System ». In *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*. (2005), p. 1-6.
- Zhou, Ming, Virgilio A Centeno, James S Thorp et Arun G Phadke. 2006. « An alternative for including phasor measurements in state estimators ». *IEEE transactions on power systems*, vol. 21, n° 4, p. 1930-1937.

Zhu, Xiaorong, Yi Wang, Lie Xu, Xiangyu Zhang et Heming Li. 2011. « Virtual inertia control of DFIG-based wind turbines for dynamic grid frequency support ». In *Renewable Power Generation (RPG 2011), IET Conference on*. p. 1-6. IET.