

# **SURVEILLANCE CONTINUE DES TRANSFORMATEURS DE PUISSANCE PAR RÉSEAUX DE NEURONES AUTO-ORGANISÉS**

THÈSE N° 2390 (2001)

PRÉSENTÉE AU DÉPARTEMENT D'ÉLECTRICITÉ

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES TECHNIQUES

PAR

**Alain SCHENK**

Ingénieur électricien diplômé EPF  
de nationalité suisse et originaire d'Onex (GE)

acceptée sur proposition du jury:

Prof. A. Germond, directeur de thèse  
Dr Th. Aschwanden, rapporteur  
M. P. Boss, rapporteur  
Prof. W. Gerstner, rapporteur  
Prof. D. Niebur, rapporteur

Lausanne, EPFL  
2001

*À mes parents*

*À Céline*

# Résumé

L'indisponibilité d'un transformateur de puissance a généralement de lourdes conséquences financières pour les exploitants des réseaux électriques. Afin de prévenir leurs pannes et d'optimiser leur maintenance, un nombre croissant de paramètres de fonctionnement est mesuré en continu. L'afflux de données lié au nombre toujours plus élevé d'unités à surveiller, nécessite l'emploi de dispositifs d'alarme automatiques.

Compte tenu de la complexité des transformateurs de puissance, l'évolution de leurs paramètres de fonctionnement est encore mal connue et difficile à modéliser. C'est pourquoi les alarmes sont généralement déclenchées sur la base de seuils fixés indépendamment pour chacune des variables enregistrées. Avec une telle approche, ni les variations relatives aux conditions de fonctionnement, ni des relations existant entre les différentes grandeurs surveillées ne sont prises en compte.

Cette thèse est consacrée à l'étude d'un système de surveillance dit "intelligent" capable de s'adapter en fonction de ces variations. La solution développée met en œuvre les cartes auto-organisées (SOM). Il s'agit de réseaux de neurones non supervisés inspirés de l'organisation des zones corticales du cerveau des mammifères. Elles ont la propriété de représenter un ensemble complexe de données dans un espace généralement bidimensionnel tout en préservant au mieux la topologie de l'ensemble original.

Plus spécifiquement, ce travail montre que ces réseaux permettent de modéliser les différents états d'un transformateur en conservant non seulement la structure générale des données mais aussi leurs corrélations locales. La qualité de conservation topologique des cartes implique que des états de fonctionnement proches seront également proches sur la

carte. Sur la base de cette constatation, une interface graphique puissante est proposée pour l'exploitation qualitative des mesures.

La surveillance est considérée dans le contexte réel pour lequel, en toute généralité, le comportement du transformateur en cas de défaut n'est pas connu. Cette contrainte a donné lieu à l'étude de la réponse des cartes auto-organisées pour des états situés hors de l'espace original modélisé. Un indicateur de confiance permettant de quantifier la qualité de la réponse du modèle a ainsi été proposé. On montre que celui-ci permet de définir un seuil d'alarme adaptatif qui complète l'outil d'interprétation qualitative.

L'approche générale développée est finalement validée pour des données collectées sur un transformateur pilote fonctionnant normalement dans le réseau électrique suisse. Les résultats obtenus montrent le bien-fondé de la technique et son aptitude à déceler des changements rapides de comportement d'une ou plusieurs variables avec une sensibilité relativement fine.

# Summary

An outage of a power transformer, generally, has heavy financial consequences for electric power systems utilities. In order to prevent any failure and to optimize their maintenance, a growing number of operating parameters are measured on-line. The important amount of data together with the ever increasing number of units to be monitored calls for automatic alarm devices.

Due to the complexity of power transformers, the evolution of the operating parameters is hardly known and, in fact, difficult to model. That is why the alarms, generally, start on thresholds defined independently for each recorded variable. With such an approach, neither variations in relation to operating conditions, nor existing relations between different monitored quantities are taken into consideration.

This thesis is devoted to the study of an "intelligent" on-line monitoring system, able to adapt itself depending on those variations. The developed solution uses self-organizing maps (SOM). It consists of unsupervised neural networks inspired by the organization of cortical zones in the brains of mammals. Their characteristics enable them to represent a complex set of data in a generally bidimensional space preserving at the same time, the topology of the original set as truthfully as possible.

This work shows, more specifically, that these systems allow to model the different states of a transformer preserving not only the general structure of data, but also their local correlations. The topological preservation feature of the maps implies that close operating states will also be close on the map. Based on this observation, a powerful graphic interface is proposed to exploit qualitatively the measurement results.

The monitoring is considered in the real context for which, in case of a defect, the transformer behavior is generally unknown. This constraint resulted in the study of the self-organizing map response for states located outside the original model space. A confidence indicator quantifying the accuracy of the model response was proposed. It is shown that this indicator allows to define an adaptive alarm threshold that complements the developed tool for qualitative interpretation.

The developed approach is finally validated using data collected on a pilot-transformer, operating in the Swiss electric network. The obtained results show the pertinence of the proposed technique and its ability to reveal any sudden change of one or several variables with a relatively sharp sensitivity.

# Remerciements

Par ces quelques lignes, je tiens à remercier toutes les personnes sans lesquelles ce travail n'aurait pas été possible. Ma gratitude va au Professeur Alain Germond qui m'a donné l'opportunité de réaliser un doctorat sous sa direction dans son laboratoire. Je le remercie pour son soutien inconditionnel et la confiance qu'il m'a toujours témoignée. Merci également au Professeur Michel Ianoz qui a été à l'origine de ce projet de recherche et qui m'a encadré au début de mon travail.

J'aimerais remercier le Professeur Dagmar Niebur, le Dr Thomas Aschwanden, Monsieur Pierre Boss et le Professeur Wulfram Gerstner d'avoir participé à mon jury de thèse ainsi que le Professeur Alfred Ruffer de l'avoir présidé. Je les remercie pour l'intérêt qu'ils ont manifesté à l'égard de mon travail et leurs critiques pertinentes.

Ma reconnaissance va également aux partenaires impliqués dans ce projet : le PSEL et la commission RDP-CREE pour le financement ; ABB Sécheron pour le soutien logistique ; les entreprises électriques BKW, ENSA et GWK pour la mise à disposition des données.

Parmi les collègues qui se sont succédés au laboratoire, je tiens tout d'abord à remercier le Dr Antonio Piras qui, toujours débordant d'idées, a su me guider au début de mes recherches. Ensuite, je veux témoigner ma gratitude à Stéphane Gerbex qui, en tant qu'ami et collègue a toujours été disponible pour discuter aussi bien des aspects scientifiques de mon travail que de certains points de détail dont j'ai le souci. Je le remercie également pour la lecture et la critique du manuscrit.

Je remercie Emmanuel Marthe pour l'aide qu'il m'a apportée lors la rédaction. J'ai également toujours pu compter sur lui pour la résolution des inévitables problèmes informatiques. Je veux aussi témoigner ma re-

connaissance au Dr Rachid Cherkaoui pour son soutien. Son expérience et ses qualités humaines m'ont aidé à canaliser mon énergie sur l'essentiel. Merci également au Dr Pierre Zweiacker, au Dr Bruno Buggiani et au Dr Farhad Rachidi de m'avoir fait profiter de leurs compétences techniques et scientifiques pour la partie expérimentale de mon travail.

J'aimerais encore associer à ces remerciements le Dr Tanja Kostic, Marcela Lenz, Andrée Moinat, Sonia Natale, le Dr Nell Reimann, José Luis Bermúdez, Vincent Borde, le Dr David Orzan, Alexandre Oudalov, Emanuel Petrache, Gaëtan Pitteloud, Abraham Rubinstein et le Dr Daniel Tabara qui, par leur amitié ont contribué à une ambiance de travail détendue et motivante.

Je ne terminerais pas sans remercier spécialement mes proches pour leur soutien. En particulier, merci à mes parents de m'avoir donné la chance d'étudier et de m'avoir toujours encouragé. Merci à mon frère David qui, depuis longtemps me fait profiter de ses compétences en informatique. Merci également à Liliane, ma belle-mère, pour ses précieux conseils en anglais. Enfin merci à Céline, mon épouse, pour ses conseils, sa patience et son tendre soutien.



# Table des matières

<b>Résumé</b>	<b>i</b>
<b>Summary</b>	<b>iii</b>
<b>Remerciements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Généralités . . . . .	1
1.2 Surveillance continue . . . . .	2
1.3 Interprétation intelligente . . . . .	3
1.4 Organisation de la thèse . . . . .	3
<b>2 Transformateurs de puissance</b>	<b>5</b>
2.1 Classification . . . . .	5
2.2 Constitution . . . . .	6
2.2.1 Éléments principaux . . . . .	6
2.2.2 Papier et carton . . . . .	8
2.2.3 Huile . . . . .	8
2.3 Évacuation de chaleur . . . . .	11
2.3.1 Principe . . . . .	11
2.3.2 Circulation de l'huile . . . . .	12
2.3.3 Point chaud . . . . .	13
2.4 Vieillissement . . . . .	14
2.4.1 Définitions . . . . .	14
2.4.2 Dégradation des isolants . . . . .	15
2.4.3 Calcul du vieillissement . . . . .	16
2.4.4 Éléments externes . . . . .	17

2.5	Défaillances . . . . .	18
2.5.1	Répartition typique . . . . .	18
2.5.2	Causes . . . . .	19
2.5.3	Température . . . . .	20
2.5.4	Humidité . . . . .	20
2.5.5	Oxygène . . . . .	22
2.5.6	Bulles de gaz . . . . .	22
<b>3</b>	<b>Évaluation de l'état des transformateurs</b>	<b>25</b>
3.1	Introduction . . . . .	25
3.2	Paramètres de l'huile . . . . .	25
3.2.1	Analyse visuelle (coloration) . . . . .	25
3.2.2	Analyses physico-chimiques . . . . .	26
3.2.3	Rigidité et facteur de pertes diélectriques . . . . .	26
3.2.4	Mesure de la température . . . . .	27
3.2.5	Gaz dissous . . . . .	29
3.2.6	Bulles de gaz . . . . .	34
3.2.7	Humidité . . . . .	34
3.2.8	Produits de décomposition . . . . .	35
3.3	Paramètres électriques . . . . .	35
3.3.1	Valeurs efficaces et instantanées . . . . .	35
3.3.2	Décharges partielles . . . . .	36
3.3.3	Fonction de transfert . . . . .	38
3.3.4	Réponse diélectrique . . . . .	40
3.4	Mesure des vibrations . . . . .	41
3.4.1	Origine du bruit . . . . .	41
3.4.2	Dispositif de mesure . . . . .	42
3.5	Équipement auxiliaire . . . . .	42
3.5.1	Bornes de traversées . . . . .	42
3.5.2	Changeurs de prises . . . . .	43
3.6	Synthèse . . . . .	43
<b>4</b>	<b>Surveillance continue</b>	<b>47</b>
4.1	Introduction . . . . .	47
4.1.1	Motivations . . . . .	47
4.1.2	Interprétation des mesures . . . . .	48
4.2	Analyses statistiques . . . . .	52

4.2.1	Ensemble de données . . . . .	52
4.2.2	Concepts de base . . . . .	52
4.2.3	Mesures de corrélation . . . . .	53
4.2.4	Analyse en composantes principales . . . . .	54
4.2.5	Méthodes de coalescence . . . . .	57
4.3	Évaluation des tendances . . . . .	58
4.3.1	Constantes de temps . . . . .	58
4.3.2	Décomposition structurelle . . . . .	58
4.3.3	Évolution à long terme . . . . .	59
4.4	Méthode des résidus . . . . .	60
4.4.1	Remarques préliminaires . . . . .	60
4.4.2	Principe . . . . .	61
4.4.3	Modélisation physique . . . . .	63
4.4.4	Modélisation linéaire . . . . .	63
4.4.5	Modélisation neuronale . . . . .	64
4.5	Détection des changements . . . . .	65
4.5.1	États de sortie . . . . .	65
4.5.2	Seuil d'alarme intelligent . . . . .	66
4.5.3	Rapport de vraisemblance . . . . .	67
4.5.4	Mesure robuste du changement . . . . .	67
4.6	Surveillance intelligente . . . . .	70
4.6.1	Systèmes experts . . . . .	70
4.6.2	Logique floue . . . . .	71
4.6.3	Réseaux de neurones . . . . .	72
<b>5</b>	<b>Cartes auto-organisées</b> . . . . .	<b>73</b>
5.1	Introduction . . . . .	73
5.2	Quantification vectorielle . . . . .	75
5.3	Fondements biologiques . . . . .	76
5.4	Réseau de Kohonen . . . . .	79
5.4.1	Hypothèses de base . . . . .	79
5.4.2	Architecture . . . . .	79
5.4.3	Apprentissage . . . . .	80
5.5	Considérations statistiques . . . . .	84
5.5.1	Préservation de la topologie . . . . .	84
5.5.2	Normalisation . . . . .	84
5.5.3	Distribution de la probabilité des poids . . . . .	86

5.6	Autres algorithmes . . . . .	87
5.7	Réseau kMER . . . . .	88
5.7.1	Maximisation de l'entropie . . . . .	88
5.7.2	Quantificateur équiprobable . . . . .	89
5.7.3	Structure du réseau kMER . . . . .	91
5.7.4	Apprentissage . . . . .	92
5.7.5	Choix des paramètres . . . . .	94
5.7.6	Densité de probabilité . . . . .	96
5.7.7	Détection de dépassement . . . . .	98
<b>6</b>	<b>Surveillance par les cartes auto-organisées</b>	<b>99</b>
6.1	Principe . . . . .	99
6.1.1	Application des cartes auto-organisées . . . . .	99
6.1.2	Cadre d'application . . . . .	102
6.2	Mise en forme des mesures . . . . .	104
6.2.1	Base de données . . . . .	104
6.2.2	Vecteur d'entrée . . . . .	105
6.2.3	Normalisation . . . . .	105
6.3	Apprentissage et adaptation du modèle . . . . .	107
6.3.1	Validation de l'algorithme d'apprentissage . . . . .	107
6.3.2	Adaptation périodique . . . . .	112
6.4	Visualisation . . . . .	113
6.4.1	Généralités . . . . .	113
6.4.2	Décomposition en couches . . . . .	114
6.4.3	Matrice des distances . . . . .	115
6.4.4	Histogramme d'activation . . . . .	116
6.5	Point de fonctionnement et trajectoire . . . . .	118
6.5.1	Neurone gagnant . . . . .	118
6.5.2	Représentation graphique . . . . .	118
6.6	Processus décisionnel . . . . .	119
6.6.1	Interprétation directe . . . . .	119
6.6.2	Distance et indicateur de confiance . . . . .	120
6.6.3	Interprétation robuste de l'indicateur <i>IC</i> . . . . .	123

<b>7</b>	<b>Évaluation du système de surveillance</b>	<b>125</b>
7.1	Installation pilote . . . . .	125
7.1.1	Caractéristiques principales . . . . .	125
7.1.2	Grandeurs mesurées . . . . .	126
7.2	Définition du vecteur d'entrée . . . . .	127
7.2.1	Mesures brutes . . . . .	127
7.2.2	Analyses préliminaires des données . . . . .	129
7.2.3	Sélection et prétraitement des variables . . . . .	134
7.3	Implémentation . . . . .	137
7.3.1	Console de surveillance . . . . .	137
7.3.2	Zones de fonctionnement . . . . .	137
7.3.3	Taille de la base de données . . . . .	140
7.3.4	Apprentissage et adaptation de la carte . . . . .	141
7.4	Sensibilité du système . . . . .	142
7.4.1	Réponse de la carte pour une valeur faussée . . . . .	142
7.4.2	Seuil de détection . . . . .	144
7.4.3	Influence du nombre de neurones . . . . .	147
7.5	Efficacité du système . . . . .	148
7.5.1	Procédure de test . . . . .	148
7.5.2	Erreur de classification . . . . .	150
7.5.3	Rapidité de détection . . . . .	152
7.6	Études de cas . . . . .	156
7.6.1	Apparition soudaine d'un décalage . . . . .	156
7.6.2	Dérive rapide d'une variable . . . . .	161
7.7	Discussion . . . . .	163
<b>8</b>	<b>Conclusion</b>	<b>165</b>
8.1	Résultats obtenus . . . . .	165
8.2	Perspectives d'avenir . . . . .	167
	<b>Bibliographie</b>	<b>169</b>
	<b>Index</b>	<b>187</b>
	<b>Curriculum vitae</b>	<b>191</b>



# Chapitre 1

## Introduction

### 1.1 Généralités

Dans un réseau d'énergie électrique, les transformateurs de puissance constituent des maillons sensibles situés entre la chaîne de production et celle de transport. Leur coût d'investissement élevé et les pertes financières associées à leur indisponibilité justifient largement la mise en place de moyens de surveillance (aussi appelée "monitoring"). L'objectif visé consiste à évaluer l'état effectif des transformateurs afin de limiter leurs risques de défaillances et d'éviter toute indisponibilité prolongée. Par ailleurs, dans l'optique de la réduction des coûts d'exploitation, la connaissance de l'état effectif du transformateur doit permettre d'optimiser la planification de la maintenance.

Deux familles de transformateurs sont particulièrement concernées par le "monitoring" :

- les transformateurs de puissance stratégiquement critiques (transformateurs élévateurs à la sortie des centrales, transformateurs situés dans un site où l'accès est difficile, transformateurs travaillant sans réserve pour des raisons économiques) ;
- les transformateurs de puissance en service depuis un grand nombre d'années dont il serait avantageux de pouvoir réduire les coûts de maintenance et planifier le remplacement.

## 1.2 Surveillance continue

Il existe deux approches de surveillance : les méthodes dites "off-" et "on-line". Les mesures du premier type (par exemple l'analyse des gaz dissous) sont les plus précises et les plus riches en information. Elles nécessitent une procédure parfois lourde à suivre qui, dans tous les cas, n'est mise en œuvre que périodiquement. Par contre, les mesures "on-line" ont l'avantage d'être effectuées automatiquement et en permanence ; elles donnent des indications pendant le fonctionnement du transformateur permettant de déceler des défauts dès leur apparition. Compte tenu de cet avantage déterminant, un grand nombre de techniques de mesures continues ont été, et continuent d'être développées. Un nombre important de paramètres de fonctionnement peut aujourd'hui être surveillé en continu. Dans la plupart des cas, les mesures effectuées sont gérées séparément les unes des autres en provoquant une alarme lorsqu'un certain seuil, défini comme la limite admissible, est dépassé.

L'interprétation des valeurs fournies par les capteurs est un problème relativement complexe. Suivant le mode de fonctionnement, les niveaux d'alarmes devraient être adaptés. Ceci implique que les relations existant entre les différents paramètres surveillés doivent être prises en compte. Malheureusement, les transformateurs de puissance sont des systèmes complexes. Tous les phénomènes régissant leur fonctionnement et leur vieillissement ne sont pas connus avec précision. De plus, il n'existe pas de transformateur de puissance standard. Chaque unité est conçue spécialement pour l'environnement dans lequel elle travaille. C'est pourquoi il n'existe aucun modèle de fonctionnement général.

Les connaissances actuelles relatives à l'évaluation de l'état des transformateurs proviennent essentiellement de l'expérience acquise sur les appareils en fonctionnement. Compte tenu de la diversité des transformateurs, les seules informations connues, pour les cas de fonctionnement critiques, sont les défaillances survenues accidentellement. En raison des coûts prohibitifs que cela engendrerait, des tests destructifs sur des unités saines ne sont tout simplement pas envisageables. De ce fait, la connaissance expérimentale des développements des défauts est elle aussi relativement limitée.



### 1.3 Interprétation intelligente

Le problème posé se résume donc à l'identification d'états défaillants sur la base de mesures dont les évolutions ne sont connues que pour des états sains. À ce stade-là, la notion de diagnostic, toujours intimement liée à la surveillance, est très limitée. Formellement, elle pourrait se limiter à : "état sain" ou "état défaillant". L'approche généralement adoptée dans des cas analogues consiste à modéliser le fonctionnement normal et à détecter toute évolution du système par rapport au modèle considéré. L'exemple le plus connu dans le domaine des transformateurs a été développé au début des années 1990 par le MIT [125]. Il met en œuvre des modèles linéaires qui doivent être adaptés fréquemment.

La nature non-linéaire et mal connue des phénomènes surveillés laisse suggérer le fait que l'approche peut être améliorée par l'utilisation de réseaux de neurones artificiels. Ces techniques sont aujourd'hui largement répandues et sont appliquées avec succès dans le domaine de l'énergie électrique depuis de nombreuses années [106]. Les réseaux les plus communs ont une structure en couches et permettent d'apprendre de façon supervisée une fonction mathématique continue sur la base d'exemples. Ils peuvent être envisagés pour remplacer avantageusement les modèles linéaires utilisés dans le système décrit plus haut. Cette thèse explore une alternative originale mettant en œuvre un autre type de réseau de neurones : les cartes auto-organisées.

### 1.4 Organisation de la thèse

Le chapitre 2 est consacré à la description générale des transformateurs de puissance et de leurs paramètres de fonctionnement. Le phénomène de vieillissement est discuté et les principales causes de défaillance sont traitées.

Les différentes méthodes mises en œuvre pour caractériser l'état des transformateurs de puissance sont détaillées au chapitre 3. Les méthodes "on-line" et "off-line" sont considérées. Elles sont synthétisées à la fin du chapitre sous forme d'un tableau qui permet de comparer le domaine d'utilisation de chaque technique.

Le chapitre 4 est dédié au traitement des informations fournies par les dispositifs de mesure. Après avoir dressé un inventaire complet des moyens utilisés dans le domaine des transformateurs de puissance, les méthodes généralement appliquées dans le cadre de la surveillance de processus sont développées.

La théorie des cartes auto-organisées est abordée au chapitre 5. Le modèle de Kohonen est tout d'abord introduit et détaillé. Les liens entre cette méthode et la quantification vectorielle d'une part, ainsi que l'organisation du cortex cérébral d'autre part, sont établis. Dans la deuxième partie du chapitre, une modification de la règle d'apprentissage appelée kMER est décrite. Cette règle introduit la notion de surface d'activation qui sera utilisée pour le système de monitoring.

Le chapitre 6 est entièrement consacré au développement de la méthode de surveillance utilisant les cartes auto-organisées. Le concept original, s'articulant autour d'une interface graphique synthétisant tous les états sur une seule carte accompagnée d'une fonction de décision numérique, est présenté. Les principales contributions de cette thèse sont développées dans ce chapitre. Au nombre de celles-ci doivent être mentionnées l'étude du rayon d'activation en tant qu'indicateur de détection d'erreur, la méthodologie d'interprétation des résultats graphiques à l'aide des histogrammes d'activation et la technique de réapprentissage contrôlé des cartes.

Au chapitre 7, l'approche développée est mise en œuvre pour des données réelles fournies par un transformateur pilote. Sur la base des enseignements tirés de l'apprentissage des cartes kMER, une méthodologie de sélection des variables est proposée. L'aptitude du système à détecter des états inconnus est également démontrée par l'étude de quelques cas.

Le chapitre 8 contient finalement les conclusions qui peuvent être tirées et propose des voies précises à explorer dans le cadre de développements futurs.

# Chapitre 2

## Transformateurs de puissance

### 2.1 Classification

Pour des raisons d'isolation électrique et de transfert de chaleur, les transformateurs rencontrés dans le domaine de l'énergie électrique contiennent généralement de l'huile minérale. Ils peuvent être classifiés selon les quatre catégories décrites ci-après [152, 170].

- a. Les *transformateurs de grande puissance* sont des unités dont la puissance nominale dépasse 100 MVA en triphasé. Les installations de ce type sont utilisées pour élever la tension à la sortie des générateurs afin d'injecter le courant dans le réseau de transport d'énergie.
- b. Les *transformateurs ou autotransformateurs de moyenne puissance* assurent généralement l'interconnexion entre les réseaux de transport et de distribution d'énergie. Leur puissance ne dépasse pas 100 MVA dans le cas des transformateurs triphasés.
- c. Les *transformateurs de distribution* ont pour mission de faire l'interface entre l'utilisateur et le réseau de distribution. Leur puissance nominale maximale est de 2500 kVA en triphasé (ou 833 kVA par noyau monophasé) et la haute tension a une valeur nominale ne dépassant pas 33 kV. Au contraire des unités des deux premières catégories, les transformateurs de distribution ne sont pas équipés de dispositif de changeur de prises en charge.
- d. La dernière catégorie est constituée par les *transformateurs à usage industriel* ou particulier. Parmi eux, les transformateurs des fours,

des redresseurs, des engins de traction peuvent être cités. La conception de ces unités dépend fortement de leur application ; il est par conséquent difficile de mettre en évidence une structure typique.

Selon les normes, le terme de transformateur de puissance s'applique à l'ensemble des transformateurs décrits ci-dessus [152] ou aux seules unités reliant les centrales au réseau [161]. Dans cette thèse, il désigne plus particulièrement les unités appartenant aux deux premières catégories.

## 2.2 Constitution

### 2.2.1 Éléments principaux

Les transformateurs de puissance rencontrés dans la pratique sont presque exclusivement triphasés. Ceux-ci peuvent être constitués soit d'une unité triphasée, soit de trois unités monophasées. Cette dernière possibilité n'est généralement utilisée que pour les très grandes puissances.

Quel que soit son type, un transformateur de puissance se compose généralement des parties suivantes (fig. 2.1) :

1. la cuve contenant l'huile ;
2. le noyau constitué de tôles d'acier isolées entre elles par du vernis ;
3. les enroulements multispire en cuivre (parfois en aluminium) montés sur le noyau ;
4. les bornes de traversée basse et haute tensions qui peuvent être respectivement isolées par de la porcelaine et du papier imprégné d'huile ou de résine époxy ;
5. le conservateur d'huile permettant le remplissage complet de la cuve en huile et réduisant au minimum la surface entre l'huile et l'air ;
6. le changeur de prises en charge ou hors charge ;
7. les groupes de refroidissement.

Pour que cette liste soit complète, il convient encore d'ajouter les diélectriques (huile, papier, carton) qui ne sont pas représentés sur la figure 2.1.

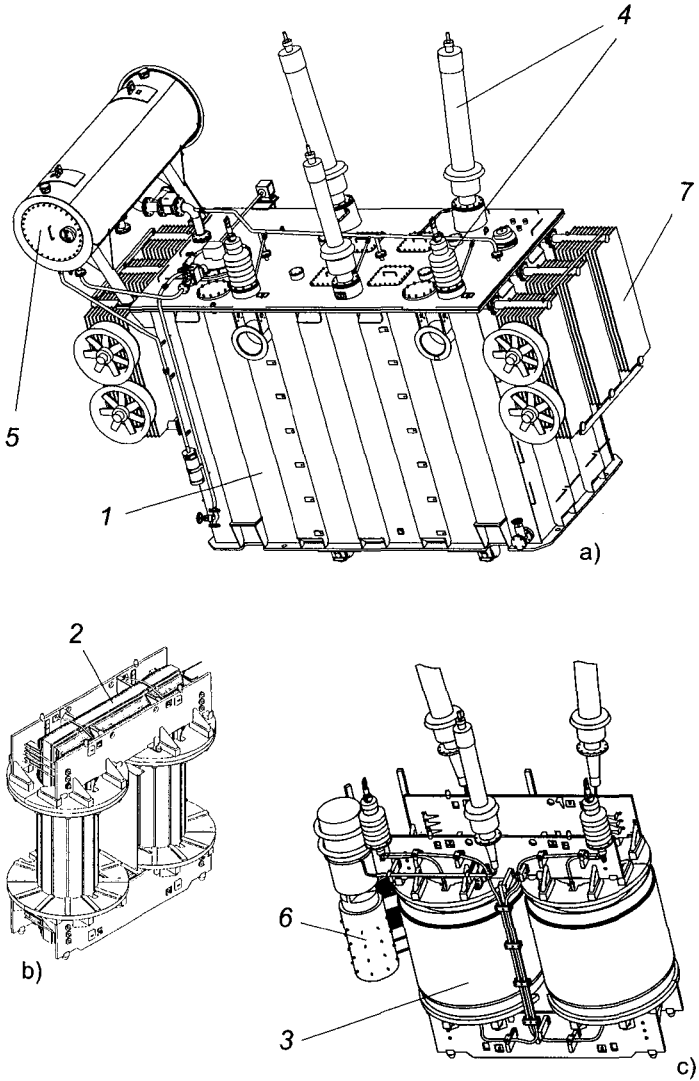


FIG. 2.1 – Constitution typique d'un transformateur de puissance : (a) vue générale, (b) circuit magnétique et (c) détail de la partie active (transformateur monophasé reproduit avec l'autorisation d'ABB Sécheron SA).

### 2.2.2 Papier et carton

Le papier et le carton (ou "pressboard") sont des matériaux isolants bon marché, possédant de surcroît des caractéristiques diélectriques excellentes. Ils sont constitués de fibres de cellulose (chaînes de glucose, fig. 2.2) extraites de la pulpe de bois ou de fibres de coton. Selon leur mode de fabrication, ils possèdent des caractéristiques mécaniques et électriques particulières expliquées principalement par la densité, la longueur et la diversité des fibres de cellulose (propriété des différents bois et cotons) [61, 145, 147].

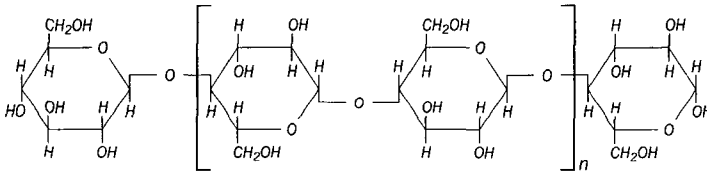


FIG. 2.2 – Formule chimique de la cellulose.

Typiquement, la constante diélectrique du papier est deux fois plus élevée que celle de l'huile. Le papier et le carton seront donc utilisés dans les zones où règne un fort champ électrique pour renforcer l'isolation de l'huile. Dans la partie active, ils interviennent seuls ou combinés dans l'isolation de la manière suivante :

- papier seul lorsqu'il s'agit d'une isolation entre deux pièces métalliques nues (plots de commutateurs par exemple) ;
- couche solide (papier ou carton), par exemple dans le cas de l'isolation entre deux conducteurs voisins d'un même enroulement ;
- isolation mixte (combinaison de lames de diélectrique liquide et de barrières d'isolant solide) pour assurer l'isolation entre deux enroulements concentriques de tensions différentes.

### 2.2.3 Huile

Les huiles minérales sont utilisées pour l'isolation des équipements électriques depuis la fin du 19<sup>ème</sup> siècle. À ce jour, l'huile constitue en-

core un composant important dans les installations électriques et les transformateurs de puissance en particulier.

Mis à part les transformateurs de distribution soumis à des contraintes de fonctionnement particulières, les transformateurs à isolation liquide sont généralement remplis d'huile minérale. Comme cela sera vu plus loin (sec. 2.3), l'huile a non seulement un rôle d'isolation électrique mais aussi de liquide de refroidissement. Les principales qualités exigées sont donc [132]:

- une faible viscosité et un bas point d'écoulement<sup>1</sup> pour assurer sa circulation;
- un point d'éclair (ou point d'inflammabilité) élevé;
- une bonne stabilité chimique essentiellement vis-à-vis des phénomènes d'oxydation et de décomposition;
- une rigidité diélectrique élevée.

L'huile minérale est un composé obtenu à partir du raffinage du pétrole brut. Sans entrer dans les détails de l'opération d'extraction, une fois exempte d'impuretés, l'huile minérale est composée essentiellement d'hydrocarbures. Ces molécules ne comportent, comme leur nom l'indiquent, que du carbone et de l'hydrogène. Elles sont divisées en trois groupes: les alcanes (ou paraffines), les cyclanes (ou naphènes) et les composés aromatiques. La quantité relative de molécules appartenant aux différents groupes varie en fonction de l'origine du gisement de pétrole. La figure 2.3 contient les formules chimiques développées de quelques exemples choisis pour chacun de ces groupes; bien que le méthane et le butane soient des gaz, leurs structures représentent exactement celles de chaînes plus longues présentes dans l'huile.

Deux types de structures chimiques sont identifiables: en chaîne ou cyclique. Rattachés à cette dernière catégorie, les composés aromatiques, compte tenu de leurs doubles liaisons, sont plus instables et donc plus réactifs que les composés paraffiniques et naphéniques. Comme cela sera vu plus loin, leur dosage optimal permet de contribuer à la stabilité chimique de l'huile.

---

1. Le point d'écoulement d'une huile est la plus basse température à laquelle l'huile coule encore lorsqu'elle est refroidie au repos dans certaines conditions bien déterminées.

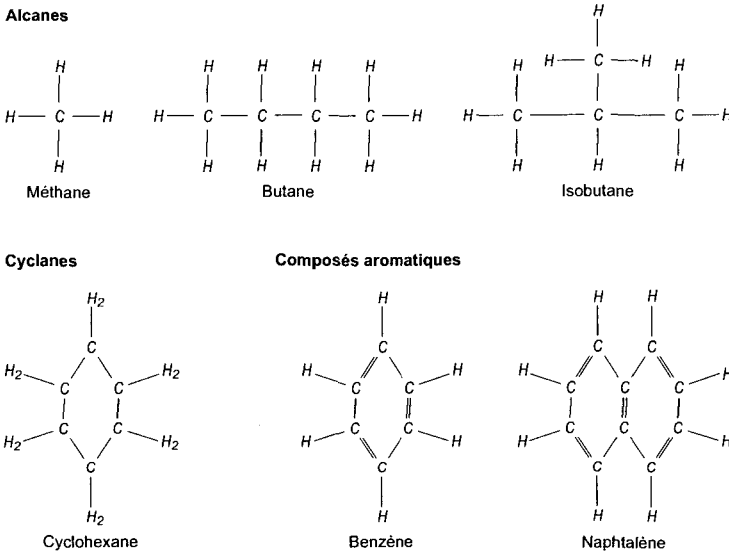


FIG. 2.3 – Structures moléculaires typiques des hydrocarbures.

En fonction des caractéristiques requises pour les installations électriques, les organisations de normalisation ont défini trois classes d'huiles isolantes pures (soit sans additif) [150]. Celles-ci se distinguent entre elles essentiellement par leur viscosité et leur point d'écoulement. Selon le type choisi, elles peuvent être utilisées pour des températures extrêmement basses (point d'écoulement minimum de  $-60\text{ }^{\circ}\text{C}$ ) et jusqu'à une température maximum de  $120\text{ }^{\circ}\text{C}$  pour de courtes durées. En utilisation continue, la limite est fixée à  $100\text{ }^{\circ}\text{C}$ .

Pour améliorer la stabilité chimique des huiles isolantes, les fabricants incorporent des additifs à très faible dose. Ceux-ci sont principalement des anti-oxydants ayant pour effet de réagir directement avec les agents oxydants, cassant ainsi la réaction d'oxydation. Ces additifs ne font que prolonger la phase d'utilisation de l'huile précédant l'altération. Une fois leur action terminée, quand les additifs anti-oxydants sont consommés, des boues et des acides se forment dans l'huile à peu près à la même vitesse que s'il n'y avait pas eu d'additif dans l'huile (§ 2.4.2 et § 2.5.5). Il convient donc de disposer d'une huile de base de haute qualité. Paral-



lèlement, une série de mesures sont prises pour limiter ce phénomène. Outre la minimisation des surfaces en contact avec l'air qui tient de la structure même des installations électriques, le dosage relatif des composés aromatiques joue un rôle important. De part leur nature réactive, ils jouent le rôle d'auto-catalyseur et, on peut montrer qu'il existe un dosage optimal minimisant les produits d'oxydation [61].

Les "abaisseurs de point d'écoulement" sont d'autres additifs fréquemment utilisés pour les huiles paraffiniques. Ils ont été introduits dans les années 1970 et permettent aux huiles alcalines, normalement peu performantes à basse température, de satisfaire les contraintes les plus sévères dans ces conditions. Grâce à cet additif, les huiles paraffiniques ont un comportement à froid identique à celui des huiles naphthéniques.

Compte tenu de la haute stabilité des huiles pures, les organismes internationaux recommandent cependant de n'utiliser les huiles inhibées qu'exceptionnellement pour des ouvrages particuliers.

## 2.3 Évacuation de chaleur

### 2.3.1 Principe

En charge, un transformateur est le siège de pertes thermiques dans son noyau (pertes fer) et dans ses enroulements (pertes Joule). Cette énergie qui se manifeste par une augmentation locale de la température, doit être évacuée à l'extérieur de la cuve.

Compte tenu de la quantité de chaleur entrant en jeu dans les transformateurs de puissance, la conception du système de refroidissement est particulièrement importante. Lorsque le transformateur fonctionne normalement, la circulation de l'huile assure son refroidissement ; la chaleur produite est véhiculée par circulation naturelle ou forcée du diélectrique liquide (huile), vers les dispositifs de refroidissement. Un transfert de chaleur bien organisé permet d'éviter la présence de zones dont la température locale est particulièrement élevée. Ceci est obtenu par une circulation d'huile abondante et bien répartie. L'efficacité est directement influencée par la viscosité et la chaleur spécifique du diélectrique utilisé.

Le transfert de chaleur a lieu en deux étapes : à l'intérieur de la cuve de la partie active (enroulements, noyau) à l'huile ; puis de l'huile au réfrigérant à travers un échangeur de chaleur externe.

### 2.3.2 Circulation de l'huile

Le mouvement des fluides peut être soit naturel par thermosiphon, soit forcé au moyen d'un équipement tel qu'une pompe ou un ventilateur. Plusieurs types de réfrigération ayant des caractéristiques propres leur correspondent. Leurs dénominations sont rassemblées dans les tableaux 2.1 et 2.2 [65].

TAB. 2.1 – *Désignation des transformateurs selon leur type.*

<b>Mode de refroidissement</b>	<b>Symbole</b>
Huile minérale	O
Askarel	L
Gaz	G
Eau	W
Air	A
Isolant solide	S
<b>Nature de la circulation</b>	<b>Symbole</b>
Naturelle	N
Forcée	F
Forcée et dirigée dans les enroulements	D

TAB. 2.2 – *Principaux types de transformateurs de puissance.*

<b>Type</b>	<b>Symbole</b>
Sec à refroidissement naturel dans l'air	AN
Sec à ventilation forcée d'air	AF
Circulation naturelle d'huile et d'air	ONAN
Circulation naturelle d'huile et ventilation forcée d'air	ONAF
Circulation forcée d'huile et d'air	OFAF
Ventilation forcée d'air et circulation dirigée d'huile	ODAF
Circulation forcée d'huile et d'eau	OFWF

Le coefficient de transmission de chaleur de l'huile dans l'air (réfrigérant) est essentiellement fonction des conditions côté air. Par la mise en mouvement de l'air au voisinage des surfaces d'échange thermique au moyen d'un ventilateur, la réfrigération est améliorée. Cependant, un refroidissement plus intense de l'huile entraîne une augmentation sensible de sa viscosité. Ceci freine sa circulation et limite l'augmentation de puissance.

Pour obtenir des performances élevées, il convient d'utiliser une pompe de circulation d'huile. Le coefficient de transmission de chaleur huile-air subit une nouvelle augmentation. Dans ce cas, la circulation d'huile dans les enroulements se fait toujours essentiellement par thermosiphon ; la plus grande partie de l'huile mise en mouvement par la pompe s'écoule dans la cuve à l'extérieur des enroulements. Ce dernier système est généralement amélioré en canalisant l'huile sortant du réfrigérant vers la base du circuit magnétique et des enroulements.

### 2.3.3 Point chaud

La zone la plus chaude de l'enroulement d'un transformateur en charge est appelée point chaud. Comme cela apparaît plus loin, la température à cet endroit constitue un des paramètres de fonctionnement les plus importants. Sa localisation au sein de l'enroulement dépend de la constitution du transformateur et il n'existe par conséquent aucune règle générale pour la déterminer a priori.

Un modèle simplifié de la distribution thermique au sein de l'enroulement a toutefois été établi par la CEI [154]. Il est représenté graphiquement à la figure 2.4 : la température de l'huile à l'intérieur des enroulements augmente linéairement depuis le bas jusqu'au sommet ; l'échauffement du conducteur le long de l'enroulement croît parallèlement à l'échauffement de l'huile avec une différence constante  $g$  entre les deux droites ; l'échauffement du point chaud est supérieur à l'échauffement du conducteur au sommet de l'enroulement (prise en compte de l'accroissement des pertes supplémentaires).

Pour tenir compte des non-linéarités, l'expression  $Hg$  matérialisant le rapport entre la température du point chaud et la température au sommet des enroulements est définie. Le facteur  $H$  est déterminé expérimentalement et peut varier entre 1.1 et 1.5 suivant la nature du

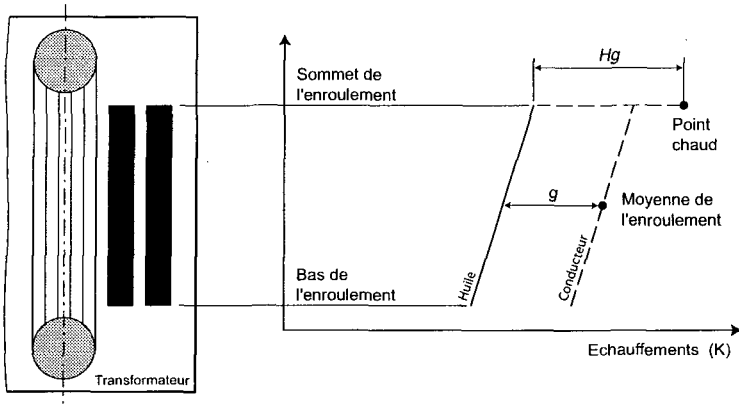


FIG. 2.4 – Diagramme thermique d'un enroulement selon [154].

transformateur. Les normes donnent des valeurs typiques de 1.1 pour les transformateurs de distribution et 1.3 pour les unités de moyenne et grande puissances.

## 2.4 Vieillessement

### 2.4.1 Définitions

Le rôle du transformateur est d'amener l'énergie électrique d'un niveau de tension à un autre niveau de tension. Tant qu'il le fait dans certaines conditions stratégiques et économiques, il est encore utile. Selon [110], le vieillissement d'un transformateur est défini par la "dégradation irréversible de la capacité de service d'un transformateur".

Cette capacité de service est en général altérée par l'augmentation de la probabilité conditionnelle de défaillance avec le temps (due par exemple à la dégradation des matériaux isolants). Dans certains cas, la capacité de service peut également diminuer en raison de la modification des contraintes économiques et stratégiques.

Par conséquent, un des paramètres déterminant pour qualifier le vieillissement d'un transformateur est le vieillissement des matériaux. La définition de cette notion tient compte non pas d'une fonction à rem-

plir, mais de la capacité à résister à certaines contraintes. Celle-ci varie en fonction de la propriété considérée et peut être énoncée de la manière suivante [110] : *"modification néfaste irréversible d'une des propriétés d'un matériau"*.

## 2.4.2 Dégradation des isolants

Le vieillissement des isolants est le résultat d'une détérioration graduelle due à des réactions physiques et chimiques affectant la tenue mécanique et diélectrique du système d'isolation. La détérioration du papier placé entre les spires des enroulements est normalement associée à la pyrolyse, à l'oxydation et à l'hydrolyse qui se produisent dans les matériaux [110]. Ces réactions entraînent la rupture des macromolécules de cellulose et réduisent le degré de polymérisation moyen. Par ailleurs, elles sont accélérées par toute augmentation de la température, de la concentration en oxygène et du taux d'humidité.

La dégradation chimique des papiers et cartons se manifeste par l'affaiblissement des propriétés mécaniques de la cellulose. Cependant, bien que les pertes diélectriques tendent à augmenter, la rigidité diélectrique du papier n'est pas touchée de manière significative par ces détériorations. L'huile isolante est également altérée par la détérioration chimique. Comme cela est expliqué plus loin (§ 2.5.5) elle s'oxyde, ce qui forme des composés acides et des boues. Ceux-ci réduisent la rigidité diélectrique et affectent la capacité d'évacuation de chaleur du transformateur.

Un certain nombre de facteurs dont l'action est généralement combinée viennent accélérer ces détériorations :

- la chaleur ;
- la présence d'une force mécanique ;
- l'application d'une tension ;
- la contamination.

Les vibrations, les efforts électromagnétiques causés par les courts-circuits et les courants d'appel, les chocs au cours du transport ainsi que les efforts résultant des variations des différents matériaux sont les principales contraintes mécaniques qui altèrent le papier et le carton. Si celles-ci sont suffisamment fortes, elles détruiront les matériaux isolants

et entraîneront une défaillance diélectrique. En outre, les efforts de compression entraînent le tassement des isolants solides et par conséquent une prise de jeu de l'ensemble de l'assemblage. Ceci rendra le transformateur plus vulnérable en cas de court-circuit.

Les contraintes occasionnées par l'application d'un champ électrique provoquent également un vieillissement de l'isolation électrique. Sous l'action de ce champ, le matériau est le siège de décharges partielles qui entraînent son érosion et affectent par conséquent sa rigidité diélectrique et sa tenue mécanique. Cependant, dans les transformateurs modernes bien imprégnés, cet effet est très faible : en fonctionnement normal, les isolants sont soumis à des champs dont le niveau est inférieur au seuil d'apparition des décharges partielles.

Enfin, le vieillissement peut être accéléré par la contamination de l'huile. Les particules métalliques, les fibres humides et les bulles de gaz présentes dans l'huile peuvent pénétrer dans des endroits où le champ électrique est élevé. De plus, la rigidité diélectrique de la cellulose peut diminuer en raison d'infiltration d'humidité.

### 2.4.3 Calcul du vieillissement

La durée de vie d'un transformateur dépend beaucoup d'événements exceptionnels, tels que les surtensions, les courts-circuits sur le réseau et les surcharges. En plus des paramètres de fonctionnement normal, ils représentent autant de grandeurs susceptibles d'influencer le vieillissement des isolants. Le phénomène de vieillissement est complexe et difficile à modéliser. À ce jour, seule l'influence des contraintes thermiques sur l'évolution des propriétés mécaniques et chimiques fait l'objet d'un modèle mathématique [152, 163].

En négligeant toutes les autres influences, sous l'effet de la température de l'huile, comme cela a été expliqué au paragraphe précédent, l'isolation subit une dégradation chimique. Après un certain temps d'exploitation, selon un certain critère retenu, l'isolation est affaiblie. La loi d'Arrhenius sur les réactions chimiques donne l'intervalle de temps pour atteindre ce point final :

$$\text{durée de vie} = \alpha \exp(\beta/T) \quad (2.1)$$

où  $\alpha$  et  $\beta$  sont des constantes et  $T$  la température absolue exprimée en Kelvin (K).

Empiriquement, sur un intervalle de température limité, Montsinger a établi une simplification de la loi d'Arrhenius. Celle-ci donne une évaluation pessimiste de la dégradation thermique utilisée comme base dans la norme CEI relative à la charge des transformateurs de puissance :

$$\text{durée de vie} = \exp(-p\theta) \quad (2.2)$$

où  $p$  est une constante et  $\theta$  est la température en degrés Celsius ( $^{\circ}\text{C}$ ).

Comme il est difficile de déterminer de manière univoque la fin de vie d'un transformateur, la vitesse de vieillissement a été définie sur la base de la durée de vie selon Montsinger :

$$\text{vitesse de vieillissement} = K \exp(p\theta). \quad (2.3)$$

La constante  $K$  est fonction entre autre de la quantité initiale des produits celluloseux ainsi que des paramètres d'environnement. En pratique, entre  $80^{\circ}\text{C}$  et  $140^{\circ}\text{C}$ , le coefficient  $p$  a une valeur telle que la vitesse de vieillissement double pour chaque incrément de  $6^{\circ}\text{C}$ .

#### 2.4.4 Éléments externes

Les considérations faites et les relations établies ici ne concernent que la partie principale du transformateur. L'état des éléments externes tels que le changeur de prises en charge ainsi que les bornes de traversées constituent également des paramètres importants pour évaluer la capacité de service d'un transformateur. Sans entrer dans les détails, au niveau des changeurs de prises en charge, l'usure est essentiellement mécanique et diélectrique (qualité de l'huile). Pour les bornes de traversée, qu'elles soient sèches ou remplies d'huile, seule la dégradation de l'isolation est prédominante [126].

Comparé à l'ensemble du transformateur, leur coût est modéré mais leur défaillance peut entraîner la destruction de tout le système. Ces éléments n'interviennent donc pas de manière déterminante dans l'évaluation globale du vieillissement du transformateur, mais du fait de leur importance stratégique, ils nécessitent une maintenance régulière.

## 2.5 Défaillances

### 2.5.1 Répartition typique

En considérant l'ensemble des composants d'un transformateur, les défaillances pouvant survenir sont multiples. La figure 2.5 recense la distribution de ces défauts pour les transformateurs de puissance munis de changeur de prises en charge et la figure 2.6 la durée d'indisponibilité des unités en fonction de la nature des défaillances [109].

Après les défauts sur les changeurs de prises, les problèmes les plus fréquents sont ceux survenant sur les enroulements et les isolants. Suivant la localisation du problème, la durée d'indisponibilité est variable. Alors que la mise hors service dure quelques jours en moyenne lorsque la défaillance a lieu sur le changeur de prises ou les traversées, elle est de plusieurs mois si elle affecte les enroulements ou le circuit magnétique.

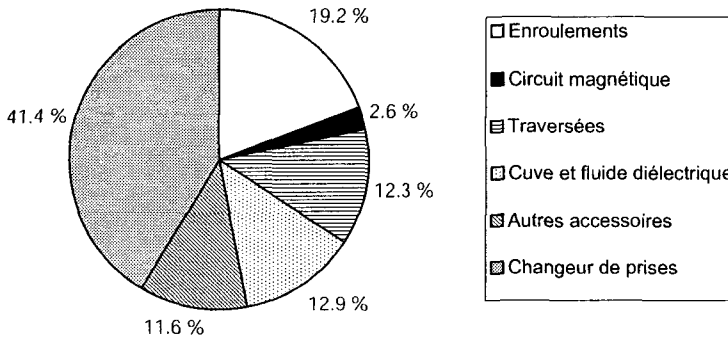


FIG. 2.5 – Distribution typique des défaillances pour les transformateurs de puissance pourvus de changeur de prises en charge (selon [109]).

Les défauts sont le plus souvent de nature mécanique ou diélectrique. Sur les changeurs de prises, ils apparaissent généralement lors de la mise en action du système. En revanche, concernant la partie principale des transformateurs, les avaries sont souvent conséquentes à des contraintes externes (coup de foudre, court-circuit, défaut de protection). Celles-ci sont parfois suffisamment sévères pour endommager un transformateur



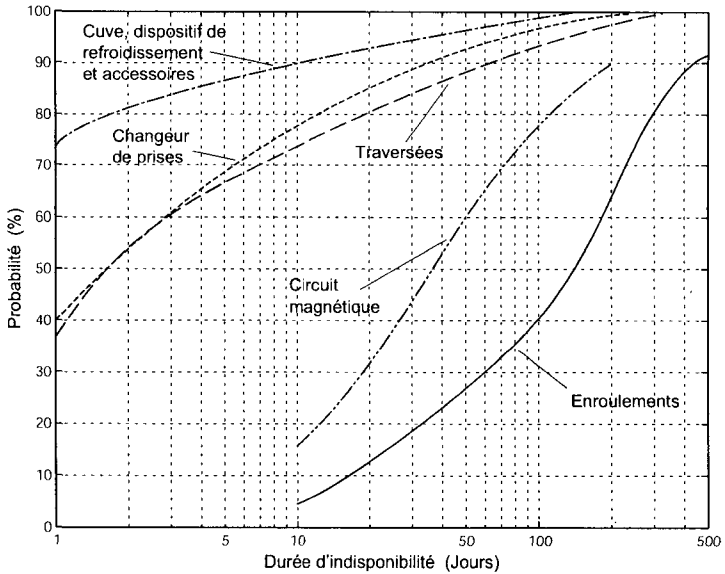


FIG. 2.6 – Probabilité cumulée de la durée d'indisponibilité après une défaillance (transformateurs de puissance pourvus de changeur de prises en charge) [109].

parfaitement sain. Dans d'autres cas, un stress même limité suffit à provoquer le claquage de l'isolation d'un transformateur vulnérable.

### 2.5.2 Causes

Quelle que soit la nature de l'événement qui entraîne la défaillance, celle-ci n'est souvent que le résultat de la détérioration graduelle normale (vieillesse) ou anormale (contraintes accidentelles répétées) du transformateur. Toutes circonstances de fonctionnement confondues, les principales causes d'altération des transformateurs sont [61] :

- la température ;
- l'humidité ;
- l'oxygène ;
- les bulles de gaz ;

- la contamination ;
- les surintensités (surcharges, courts-circuits) ;
- les surtensions (arcs électriques, décharges partielles).

La présence de particules (contamination) a un effet néfaste sur la rigidité diélectrique de l'huile. Elle est source de soucis principalement pour les transformateurs de tension nominale égale ou supérieure à 400 kV [3]. De manière générale, les quatre premiers éléments de cette liste sont prédominants et font l'objet d'une description détaillée.

### 2.5.3 Température

Comme cela est décrit à la section 2.3, le fonctionnement d'un transformateur de puissance est tel que lorsqu'il est chargé, sa température augmente. Sa capacité à dissiper cette chaleur est fonction de sa conception et de l'état de son huile. Les températures élevées sont dangereuses, essentiellement parce qu'elles provoquent une accélération de la dégradation chimique de la cellulose. Celle-ci est encore plus rapide en présence d'humidité et d'oxygène (§ 2.5.4 et 2.5.5). De plus, l'échauffement peut être responsable de l'apparition de bulles de gaz qui affaiblissent localement la tenue diélectrique de l'huile (§ 2.5.6).

Les défaillances liées à une élévation de température sont :

- des courts-circuits entre les tôles du noyau pouvant détruire l'isolation ou déformer le noyau ;
- un retrait et une fragilisation de l'isolation qui peut se détruire suite à une contrainte mécanique.

### 2.5.4 Humidité

L'humidité est toujours présente dans la cellulose dont sont constitués les isolants. Sa quantité varie entre 0.5 et 3 % de la masse totale. Suite à une élévation de température, elle est drainée en dehors du papier qui constitue l'isolant solide, vers l'huile du transformateur. Elle est aussi un des produits de décomposition de la cellulose et peut également rentrer dans le transformateur par d'éventuelles fuites dans la cuve du transformateur [184].

Il est reconnu depuis longtemps que la présence d'humidité dans l'isolation solide accélère son vieillissement. Par conséquent, elle contribue également à la dégradation de la qualité de l'huile ainsi qu'à celle de la structure du transformateur. L'humidité n'affecte pas directement la tenue diélectrique de la cellulose. Par contre, sous forme d'eau libre, lorsqu'elle condense, elle a une action particulièrement nocive sur l'ensemble de l'isolation.

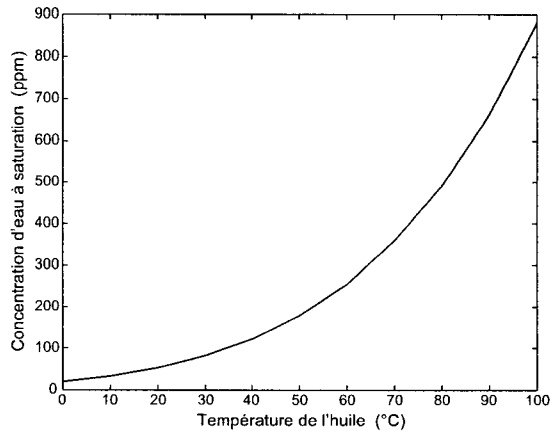


FIG. 2.7 – Évolution de la concentration d'eau à saturation dans l'huile en fonction de la température (valeurs tirées de [112]).

Le facteur déterminant est le point de saturation de l'humidité dans l'huile. Comme le montre la courbe de la figure 2.7, la concentration d'eau contenue dans l'huile à saturation augmente avec la température (cf. [33] et [112] pour plus de détails). La solubilité de l'eau dans des huiles de transformateurs âgés peut ainsi décupler entre 20 °C et 80 °C. Parallèlement, l'absorption de l'eau par la cellulose est fonction de l'inverse de la température. L'état du système résulte donc de l'équilibre entre ces deux phénomènes dont les constantes de temps diffèrent [114]. Si un transformateur relativement chargé est mis hors-service, la température de l'huile chute assez rapidement. Cette diminution peut être encore accélérée par la chute de la température extérieure si l'arrêt a lieu à la tombée du jour. La dynamique de l'absorption de l'eau par le papier

étant beaucoup plus lente que celle de sa libération, il en découle une augmentation passagère du taux relatif d'humidité dans l'huile. Dans certains cas, si le transformateur est exploité avec un taux d'humidité élevé, ce processus peut entraîner la condensation de l'eau à saturation et, si le transformateur est remis en service à ce moment-là, la destruction de l'appareil.

### 2.5.5 Oxygène

L'oxygène présent dans l'huile provient non seulement de l'extérieur de la cuve, mais aussi de la dégradation thermique de la cellulose. Il réagit avec des hydrocarbures pour former des oxydes qui constituent un des facteurs de vieillissement de l'huile (§ 2.4.2).

Le phénomène d'oxydation s'amplifie par la présence de catalyseurs et d'accélérateurs [61] dont les principaux éléments sont :

- l'humidité (catalyseur) ;
- le cuivre des enroulements (catalyseur) ;
- la chaleur (accélérateur) ;
- les vibrations (accélérateur) ;
- les efforts électriques (accélérateur).

Le résultat principal de ces réactions d'oxydation est la production de boues. Il s'agit de substances partiellement conductrices qui précipitent hors de l'huile sur les isolants, la cuve, les surfaces du circuit de refroidissement, etc. Ces dépôts peuvent constituer des barrières conductrices en surface, diminuer la tenue mécanique du transformateur, obstruer les conduits de circulation d'huile et affecter le coefficient de transfert de chaleur entre les enroulements et l'huile. Le système de refroidissement perd ainsi son efficacité et l'élévation de température en découlant accélère l'oxydation et la dégradation des isolants solides.

### 2.5.6 Bulles de gaz

La formation de gaz libres dans le système d'isolation est particulièrement préoccupante. Compte tenu de leur très faible rigidité diélectrique par rapport à l'huile ou la cellulose, leur accumulation sous forme de

bulles à des endroits où les contraintes diélectriques sont particulièrement élevées peut entraîner le claquage de l'isolation.

Les bulles de gaz peuvent se former suivant trois mécanismes [98] :

- la sursaturation de l'huile en présence d'une couverture gazeuse ;
- la décomposition thermique de la cellulose ;
- la vaporisation de l'humidité adsorbée par la cellulose.

Tous ces mécanismes découlent de la loi de Henry. À l'équilibre, elle lie la concentration  $C_i$  d'un gaz  $i$  dissous dans un liquide à la pression  $P_i$  du même gaz à la surface du liquide. La concentration est exprimée en part par million (ppm) et est définie par

$$C_i = K_i P_i \quad (2.4)$$

avec

$$K_i \sim \exp(A_i/T) \quad (2.5)$$

où  $T$  est la température absolue exprimée en Kelvin et  $A_i$  est une constante dépendant de la nature de la combinaison gaz-liquide.

La formation de bulles dans l'huile est la conséquence d'une différence de pression lorsque le système, suite à un changement brusque des conditions, vient s'établir d'un point d'équilibre à un autre point d'équilibre. En tenant compte du fait qu'il y a plusieurs gaz dissous, la condition pour qu'il y ait formation de bulles s'exprime par

$$\sum C_i/K_i > P_{stat} \quad (2.6)$$

où  $P_{stat}$  est la pression statique exercée sur la surface du liquide.

Le premier phénomène apparaît sur les transformateurs équipés d'un système d'isolation gazeuse de l'huile. Il est déclenché par une baisse soudaine de charge à partir d'un état fortement chargé. Lors de son fonctionnement à pleine charge, la chaleur dégagée provoque une dilatation de l'huile. Par conséquent, le volume des gaz diminue et leur pression augmente. Selon l'équation (2.4), avec une telle pression, à l'équilibre la concentration de l'huile en gaz est supérieure au cas où la température est basse. Lors d'une baisse de charge brusque, le système de refroidissement est tel que la température chute rapidement, entraînant une baisse de pression. Suivant la rapidité de celle-ci, les gaz dissous n'ont pas le temps de diffuser à la surface. La concentration en gaz de l'huile est alors

supérieure à la limite définie par la loi de Henry et des bulles se forment spontanément.

Le deuxième phénomène survient à la surface des conducteurs. Avec l'augmentation de la température, la décomposition des isolants augmente, influençant localement la concentration des gaz dans l'huile. Les principaux produits de la décomposition de la cellulose sont l'eau, le CO et le CO<sub>2</sub>. Si la température est suffisamment élevée, l'augmentation de ces gaz sera plus rapide que leur diffusion vers le reste du liquide. Selon le même principe que pour le premier phénomène, des bulles se formeront localement.

En pratique, ce phénomène n'est observé que pour des températures dépassant 150 °C. Au-dessous de ce seuil, lorsque des bulles de gaz sont formées, elles sont essentiellement constituées de vapeur d'eau (troisième phénomène). La cellulose contient une forte proportion (en masse) de l'eau contenue par le système. Elle agit comme un réservoir infini fournissant de l'eau dans l'huile lorsque la température augmente et adsorbant celle-ci lorsque la température diminue. Lors d'une augmentation rapide de la charge, la température locale du papier isolant les conducteurs s'élève, libérant de la vapeur d'eau. La pression supplémentaire peut ainsi entraîner localement un dépassement de la pression statique sur l'huile et conduire à la formation de bulles gazeuses libres.

## Chapitre 3

# Évaluation de l'état des transformateurs

### 3.1 Introduction

Afin de détecter la variation des divers paramètres de fonctionnement des transformateurs de puissance, un grand nombre de capteurs et de méthodes de mesure ont été développés [21]. Ces techniques peuvent être classifiées en deux grands groupes :

- mesures dites "off-line" nécessitant l'arrêt du transformateur ;
- mesures dites "on-line" mises en œuvre en continu alors que l'unité fonctionne normalement.

Qu'elles appartiennent à l'une ou l'autre des deux catégories, ces méthodes sont utilisées soit indépendamment les unes des autres, soit de manière combinée. Elles permettent aux opérateurs, dans un premier temps, d'effectuer la surveillance des unités et, le cas échéant si la situation le nécessite, de procéder à un diagnostic de panne (chap. 4).

### 3.2 Paramètres de l'huile

#### 3.2.1 Analyse visuelle (coloration)

Lors d'un prélèvement d'huile, la première analyse faite concerne la coloration du diélectrique. Cette mesure donne une indication quant à la détérioration relative de l'isolant. Lorsque l'huile est neuve, elle est

transparente. Avec le temps, elle passe graduellement du jaune à l'ambre indiquant un vieillissement avancé ou un défaut. Elle peut être encore brune ou brun foncé. Dans ces derniers cas, il est fortement recommandé de changer l'huile. Bien que cette technique soit essentiellement qualitative, une échelle quantifiant la coloration peut être établie [162]. Une augmentation anormalement rapide de la coloration vers le brun est un indicateur à la fois sensible et fiable d'un changement important au sein du transformateur (détérioration, contamination de l'huile ou les deux).

### 3.2.2 Analyses physico-chimiques

Outre les tests qui sont décrits en détail dans les paragraphes qui suivent, les analyses physico-chimiques de l'huile [150, 151, 159] utiles pour l'évaluation de l'état des transformateurs de puissance portent essentiellement sur les quatre propriétés décrites ci-après.

- a. La *densité relative* est le rapport de poids entre un volume donné d'huile et le même volume d'eau. De moins en moins effectuée, cette mesure permet de déceler une éventuelle contamination de l'huile.
- b. Le *taux d'acidité* est obtenu à l'aide d'un test de neutralisation consistant à déterminer la quantité d'hydroxyde de potassium nécessaire pour neutraliser 1 g d'huile. Le nombre de neutralisation est exprimé en KOH/g. Il permet de détecter la contamination et le vieillissement de l'huile.
- c. La *tension interfaciale* permet de déceler les contaminants solubles et les produits de dégradation de l'isolation. Elle donne une indication essentiellement qualitative mais très fine du début du processus de vieillissement.
- d. Le *taux de sédimentation* est une valeur indiquant l'état général de l'huile. Il est obtenu par décantation-centrifugation mais aussi par précipitation des boues solubles.

### 3.2.3 Rigidité et facteur de pertes diélectriques

Les tests de rigidité diélectrique permettent d'évaluer l'aptitude de l'huile à résister au stress diélectrique à la fréquence industrielle de 50 ou 60 Hz. Ils consistent à relever les tensions disruptives du milieu entre



diverses configurations d'électrodes dont les conditions expérimentales sont fixées par les normes internationales [146, 155].

La mesure du facteur de pertes  $\tan\delta$  est régie par les mêmes normes. Celui-ci représente la mesure du déphasage entre la tension appliquée aux bornes du diélectrique et le courant le traversant. En présence de boues et autres matériaux contaminants, sa valeur a tendance à augmenter.

### 3.2.4 Mesure de la température

Comme les considérations faites au chapitre précédent le montrent (sec. 2.3 et 2.5), la chaleur est un des principaux facteurs de défaillance et de vieillissement. Par conséquent, il est fortement souhaitable de connaître l'état thermique du transformateur. Les modèles de vieillissement (§ 2.4.3) ainsi que les limites de fonctionnement sont liés à la température du point chaud [152]. Malheureusement, à cause de la tension élevée présente sur les enroulements, l'utilisation de capteurs classiques en métal et de câbles est impossible. De plus, la localisation exacte du point chaud est difficile à prévoir à l'avance.

Afin de contrôler le dispositif de refroidissement et de bénéficier d'une indication de température, traditionnellement, des systèmes mécaniques spéciaux [61] sont placés au sommet des enroulements. Pour la détermination de la température du point chaud, une installation complémentaire appelée dispositif d'image thermique est souvent installée [61]. Il s'agit d'une résistance de chauffage alimentée par l'intermédiaire d'un transformateur de courant installé sur l'une des bornes primaire ou secondaire. L'ensemble est étalonné de telle sorte que l'écart de température produit corresponde à l'écart effectif entre la température du point chaud et celle où se trouve le capteur (sommet des enroulements).

Sur une cuve de transformateur, le nombre de points accessibles avec des thermomètres ou des thermocouples ordinaires est relativement grand. En pratique, la température de l'huile est mesurée soit indirectement en plaçant un capteur sur la cuve, soit directement en utilisant un capteur placé dans un doigt de gant<sup>1</sup>. Les points de mesure sont généralement situés dans la partie supérieure de la cuve au sommet d'un

---

1. Gaine de protection.

enroulement et au bas de la cuve à la sortie des circuits de refroidissement. La température du point chaud est ensuite estimée à l'aide de modèles. Bien que de nombreux organismes ou instituts de recherche proposent diverses variantes basées sur le même principe [89, 163], le modèle le plus communément utilisé est celui décrit au paragraphe 2.3.3 [154].

Afin de résoudre les problèmes liés à la localisation des points chauds, des techniques récentes basées sur l'utilisation de capteurs optiques ont été développées. Elles ont abouti à la mise au point d'un dispositif capable de donner en permanence une image thermique de tout l'enroulement [20]. Le principe est basé sur l'utilisation de fibres optiques insérées au sein même des conducteurs. Ce procédé offre l'avantage décisif de permettre de localiser exactement l'emplacement du point chaud : à partir des modifications des propriétés optiques locales, le profil exact de température le long de la fibre est calculé. Cependant, la mise en œuvre d'une telle technique nécessite l'intégration de la fibre au sein même des conducteurs. Par conséquent, cette méthode n'est adaptée que sur les nouvelles unités.

TAB. 3.1 – Paramètres thermiques de référence pour un vieillissement normal en exploitation continue (CEI 60354).

Paramètre		Transformateur		
		ON..	OF..	OD..
Température ambiante	(°C)	20	20	20
Échauffement du point chaud	(K)	78	78	78
Échauffement au sommet de l'enroulement	(K)	52	56	49
Échauffement au bas de l'enroulement	(K)	34	36	43

Quelle que soit la technique de mesure utilisée, en vertu des modèles de vieillissement énoncés à la section 2.4, les organismes de normalisation nationaux et internationaux ont adopté un certain nombre de recommandations quant aux limites d'excursion en température. Celles-ci dépendent du type du transformateur et de ses conditions d'exploitation : température ambiante, service continu ou cyclique, etc. Le tableau 3.1 donne les valeurs de référence selon [152] pour un fonctionnement en continu impliquant un vieillissement qualifié de normal ; vitesse de

vieillesse selon (2.3) égale à un. La température moyenne du point chaud constitue le paramètre déterminant. Dans tous les cas, elle est fixée à 98 °C.

### 3.2.5 Gaz dissous

Les réactions chimiques induites par le fonctionnement du transformateur provoquent la formation de composés gazeux dissous dans l'huile (§ 2.4.2). L'analyse de ces gaz (DGA) permet de déceler le développement lent d'anomalies et d'irrégularités dans le système d'isolation huile/papier, telles que des points chauds, des mauvais contacts, des arcs ou des décharges partielles. La figure 3.1 représente les types de gaz générés suivant la température de l'huile.

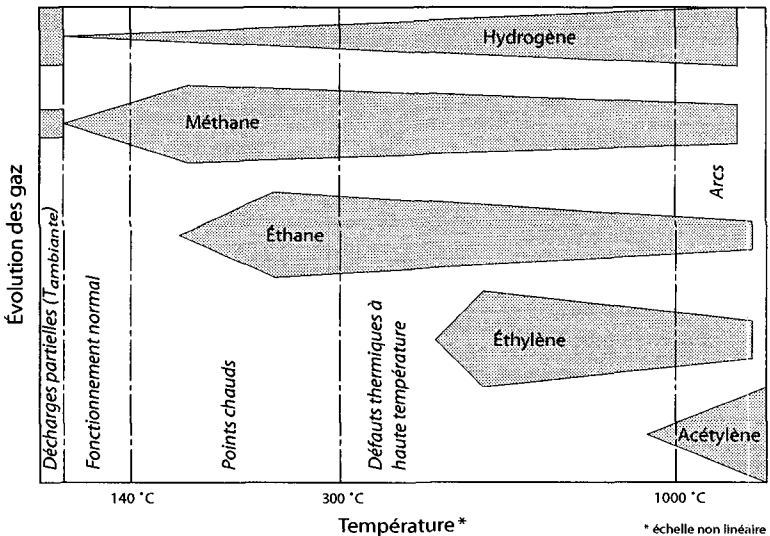


FIG. 3.1 - Évolution des gaz dissous en fonction de la température de l'huile [61].

La technique de mesure généralement utilisée est l'analyse par chromatographie en phase gazeuse d'un échantillon d'huile préalablement prélevé sur le transformateur [61]. Une fois les gaz collectés, l'interpré-

tation des résultats peut être faite de diverses façons. Il existe de nombreuses méthodes utilisées dans ce contexte. Elles peuvent être regroupées en trois catégories principales [100] et deux catégories d'application particulières :

- les méthodes des gaz clés ;
- les techniques des rapports ;
- les représentations graphiques ;
- les méthodes de diagnostic avancées ;
- les techniques de mesure continue.

### Gaz clés

Les méthodes des gaz clés, comme leur nom l'indiquent, fournissent pour chaque gaz détecté l'interprétation principale que l'on peut en donner (tab. 3.2). Les résultats sont principalement qualitatifs et servent de base pour des investigations approfondies [161].

TAB. 3.2 – *Interprétation principale des gaz dissous dans l'huile.*

Gaz détecté		Interprétation
Oxygène	$O_2$	Défaut d'étanchéité du transformateur
Mono- et dioxyde de carbone	$CO$ et $CO_2$	Décomposition de la cellulose (vieillessement ou défaut)
Hydrogène	$H_2$	Décharges électriques (effet couronne, faibles décharges partielles)
Acétylène	$C_2H_2$	Défaut électrique (arc, étincelle)
Éthylène	$C_2H_4$	Défaut thermique (surchauffe locale)
Éthane	$C_2H_6$	Indicateur secondaire d'un défaut thermique
Méthane	$CH_4$	Indicateur secondaire d'un arc ou d'une surchauffe importante

Suivant le même principe, une technique globale est parfois utilisée : seule la quantité totale des gaz combustibles dissous est prise en compte pour effectuer le diagnostic. Si la quantité totale des gaz combustibles est dans la plage de 0 à 500 ppm, le transformateur est considéré comme

fonctionnant normalement. À 1000 ppm, la décomposition est significative et à plus de 2500 ppm de gaz combustible, la décomposition est substantielle [81].

## Rapports

Les méthodes dites des rapports permettent une analyse plus fine de l'état du transformateur. Leur principe est le calcul de rapports des quantités mesurées de gaz. L'interprétation de ces rapports est donnée par de nombreux guides. Les plus utilisés sont la norme CEI 60599 [157], les rapports de Dörnenburg [160] et les rapports de Rogers [160]. Les rapports diffèrent d'une méthode à l'autre mais se basent tous sur l'utilisation des cinq gaz suivants :  $H_2$ ,  $CH_4$ ,  $C_2H_2$ ,  $C_2H_4$  et  $C_2H_6$ . En fonction des valeurs des rapports, une table (tab. 3.3 et 3.4 pour le guide CEI 60599) basée sur l'expérience établit le diagnostic de l'état de l'unité testée.

TAB. 3.3 – *Diagnostic selon la méthode des rapports (CEI 60599).*

Cas	Diagnostic	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
0	Pas de défaut	0	0	0
1	Décharges partielles (faible énergie)	0	1	0
2	Décharges partielles (haute énergie)	1	1	0
3	Décharges de faibles énergies	1/2	0	1/2
4	Décharges de hautes énergies	1	0	2
5	Défaut thermique ( $T < 150$ °C)	0	0	1
6	Défaut thermique ( $150$ °C $< T < 300$ °C)	0	2	0
7	Défaut thermique ( $300$ °C $< T < 700$ °C)	0	2	1
8	Défaut thermique ( $T > 700$ °C)	0	2	2
Rapport des gaz caractéristiques :		< 0.1	0	1
		0.1-1	1	0
		1-3	1	2
		> 3	2	2

Le principal désavantage de ces méthodes est une certaine ambiguïté dans l'analyse lors de l'apparition simultanée de plusieurs défauts. De plus, le niveau absolu des concentrations de gaz n'est pas du tout pris en compte. En revanche, leur intérêt réside dans l'indépendance des mesures par rapport au volume d'huile et aux unités de concentration [100].

TAB. 3.4 – Exemples typiques de défauts diagnostiqués à l'aide de la méthode des rapports (CEI 60599, tab. 3.3).

Cas	Exemples typiques
0	Vieillessement normal
1	Décharges dans des poches de gaz provenant d'une imprégnation incomplète, d'une sursaturation ou d'une teneur en eau élevée
2	Idem que cas 1, mais provoquant le marquage ou la perforation de l'isolation solide
3	Amorçage continu dans l'huile dû à de mauvaises connexions ou à des potentiels flottants
4	Décharge de puissance / Rupture diélectrique d'huile entre enroulements, entre spires, ou entre spire et masse / Courant de coupure dans le sélecteur
5	Échauffement généralisé d'un conducteur isolé
6-8	Surchauffe locale du circuit magnétique due à des concentrations de flux / Points chauds de températures croissantes : petits points chauds dans le circuit magnétique ; mauvais contacts ; points chauds dus à des courants de circulation entre circuit magnétique et cuve

### Représentation graphique

La représentation graphique des gaz permet de suivre visuellement l'évolution des défauts. Cette technique repose sur le calcul du pourcentage relatif de trois gaz : méthane, éthylène et acétylène. Ceux-ci sont représentés sur un triangle dont chaque coin représente 100 % d'un gaz et 0 % des autres gaz [35]. L'évolution vers un défaut est ainsi repérée suivant la forme du triangle. D'utilisation facile, cette méthode a le désavantage d'être un peu lourde et de ne tenir compte que de trois gaz.

### Diagnosics avancés

Malgré l'existence de normes et une utilisation très répandue de la DGA (depuis plus de 40 ans), une appréciation correcte du degré de gravité de l'anomalie et sa localisation à l'intérieur du transformateur sont souvent difficiles. En général, les valeurs limites des concentrations de gaz maximales admissibles pour un transformateur spécifique exploité dans des conditions spécifiques ne sont pas disponibles.

Par ailleurs, un groupe d'action de la CIGRÉ a publié récemment de nouvelles directives pour l'interprétation de l'analyse des gaz dissous [100]. Ces recommandations reprennent l'essentiel des points développés plus haut et ajoutent deux rapports supplémentaires à ceux définis par les normes CEI et IEEE ( $CO_2/CO$  pour l'identification de dégradation cellulosique et  $C_2H_2/H_2$  pour l'identification de défauts sur le changeur de prises).

L'exploitation des résultats fait encore l'objet de nombreuses recherches dans le domaine de l'application de méthodes basées sur l'intelligence artificielle notamment [39, 69, 168, 169, 182, 188].

### Mesure continue

La méthode DGA est largement utilisée non seulement pour identifier la cause d'un défaut, mais aussi pour détecter une anomalie naissante dans un transformateur de puissance. Sa sensibilité pour une évaluation de l'état de fonctionnement général d'un système isolé est limitée par sa nature intégrale et cumulative. À cause du grand volume d'huile, des petits défauts, même localisés sur une partie critique, entraînent un temps de réaction très long. En procédant à des mesures à intervalles fixes (typiquement 1 à 5 ans entre chaque analyse), il y a un risque que des défauts se développant rapidement et induisant des changements rapides dans le comportement du transformateur ne soient pas détectés suffisamment tôt.

Pour palier à cet inconvénient, des appareils capables d'extraire et d'analyser l'huile en continu ont été développés [11, 174]. Ce type de capteur se limite souvent à l'analyse d'un seul gaz (les versions multi-gaz [143] sont encore peu fiables et chères). En pratique, comme le montre le tableau 3.2, l'hydrogène, bien que principal indicateur de décharges électriques, est présent pour la plupart des types de défaut. Une surveillance constante de son taux est donc une bonne indication de l'état général de l'isolation du transformateur. Lorsque celui-ci augmente trop fortement ou atteint une valeur limite, la mesure peut être affinée avec une DGA classique.

### 3.2.6 Bulles de gaz

Les gaz non dissous sont collectés par un relais à pression de gaz, dit "relais de Buchholz". Celui-ci est monté sur une canalisation reliant le point le plus haut de la cuve au conservateur. Il s'agit d'une chambre comportant généralement deux flotteurs à contact. Le gaz produit vient s'accumuler lentement en faisant descendre progressivement le flotteur du haut jusqu'à ce qu'il actionne un dispositif d'alarme. Si le dégagement de gaz est important et rapide, un ordre de mise hors service instantané est déclenché par un second contact. Les gaz collectés dans la chambre sont ensuite récupérés et analysés selon des méthodes similaires à celles utilisées pour les gaz dissous.

### 3.2.7 Humidité

Les considérations faites sur le vieillissement du transformateur ont mis en évidence non seulement l'influence de l'humidité sur la dégradation des isolants, mais aussi le phénomène de condensation de l'eau lors de la variation rapide de la température de l'huile.

La technique classique de mesure d'humidité dans l'huile est une méthode connue sous le nom de "test de Karl Fischer" [156]. Spécialement conçue pour quantifier des faibles concentrations, elle est généralement mise en œuvre en laboratoire avec des échantillons d'huile prélevés sur site. Utilisant la titration colorimétrique, les résultats qu'elle fournit sont des quantités absolues mesurées en particules par million (ppm). Sa précision pour les huiles de transformateurs commence à être mise en doute et de nouvelles techniques sont en cours de développement [76]. Elle reste toutefois la technique de référence à ce jour.

Si les résultats ainsi obtenus sont importants pour caractériser le vieillissement de l'isolation, en charge, la grandeur importante est l'humidité relative (RH). C'est pourquoi des capteurs disponibles dans le commerce pour des mesures en environnement gazeux ont été adaptés à celui des huiles de transformateurs [113]. La mesure du taux relatif d'humidité est basée sur la variation des paramètres électriques des capteurs en fonction de l'humidité. Les plus appropriés à ce jour utilisent des films polymères haute température dont la constante diélectrique varie en fonction du taux d'humidité relative. L'intégration de ces détecteurs



à la cuve des transformateurs permet ainsi de surveiller en temps réel l'évolution générale du taux d'humidité mais aussi de prévenir tout problème lors de l'exploitation dans des conditions aussi sévères que celles décrites au paragraphe 2.5.4 [114].

### 3.2.8 Produits de décomposition

Lorsque l'évaluation des résultats de la DGA indique une dégradation thermique de l'isolation en cellulose (concentration inhabituellement élevée en oxyde de carbone), les produits de décomposition peuvent être quantifiés à l'aide d'une chromatographie en phase liquide à haute performance (HPLC). Cette technique permet de détecter les composés furaniques produits par la décomposition de la cellulose. Bien que les aspects analytiques de cette méthode soient bien définis dans [153], l'interprétation quantitative des résultats de la HPLC reste difficile car la corrélation entre les mesures de concentration en furane dans l'huile et les changements de propriétés de la cellulose (réduction de la tenue mécanique) n'est pas encore totalement comprise [64]. Un groupe d'action CIGRÉ a mis sur pied un programme visant à étudier ce phénomène afin d'établir des règles d'interprétation [116]. Actuellement, comme pour les gaz, il n'existe pas de valeur numérique limite caractérisant une dégradation trop avancée de la cellulose. La surveillance s'effectue donc principalement sur la base de l'évolution des grandeurs mesurées au fil du temps [135].

## 3.3 Paramètres électriques

### 3.3.1 Valeurs efficaces et instantanées

En fonctionnement normal, les transformateurs de puissance sont équipés de capteurs qui relèvent les paramètres électriques de fonctionnement (tensions et courants). Ceux-ci sont essentiels pour connaître le point de fonctionnement par rapport aux caractéristiques nominales et ainsi maintenir le transformateur dans des conditions pour lesquelles il a été conçu. Le traitement de ces mesures permet de mettre en œuvre des dispositifs de protection tels que les relais de surintensité, les relais différentiels et les relais de défauts de terre [61, 65].

Avec des capteurs possédant des caractéristiques dynamiques suffisamment élevées, la mesure des paramètres électriques permet également de déceler des phénomènes transitoires tels que les surtensions et les courts-circuits créés par les manœuvres, la foudre et les défauts internes [2]. Ces mesures sont généralement les signaux d'entrées de relais qui viennent s'ajouter aux éléments de protection passifs tels que les parafoudres.

### 3.3.2 Décharges partielles

Une décharge partielle est un phénomène de rupture diélectrique qui se produit sur une partie seulement de l'isolation des conducteurs [148]. Toute décharge partielle génère des ondes électromagnétiques, des ondes acoustiques, un échauffement local et des réactions chimiques. L'analyse des gaz dissous (§ 3.2.5) apporte une contribution à leur mesure en permettant de déceler la présence des produits chimiques. Cependant, les techniques de mesure des décharges partielles les plus efficaces sont les méthodes de localisation électrique et acoustique. Afin d'obtenir une meilleure précision dans la quantification et la localisation des décharges partielles, les deux méthodes sont parfois utilisées simultanément.

#### Méthodes de localisation électrique

Toutes les techniques de localisation électrique sont basées sur l'analyse des signaux temporels mesurés sur les traversées des transformateurs. Ceux-ci représentent simplement les réponses du circuit électrique à des excitations provenant de pics de courant (de l'ordre de quelques nanosecondes) générés par des décharges partielles. Elles sont largement utilisées en laboratoire et documentées par la norme CEI 60270 [148]. Théoriquement, il existe une réponse spécifique à chaque type et chaque localisation de décharge partielle. Malheureusement, à cause de contraintes pratiques évidentes qui limitent les points de mesure aux traversées et la difficulté de disposer d'un modèle électrique interne suffisamment précis, il est actuellement extrêmement difficile de traiter précisément les décharges qui ne sont pas localisées directement sur des éléments reliés aux bornes de mesure.

Les procédés classiques de localisation électriques sont la mesure de la charge apparente et celle dite de comparaison des profils. La première d'entre elles consiste à mesurer la décharge se produisant dans une capacité donnée de façon à avoir une idée de la charge libérée lors de la décharge partielle. Il convient donc de savoir dans quelle capacité a lieu la décharge. Pour les raisons évoquées plus haut, sa mise en œuvre dans le cas des transformateurs de puissance est difficile ; à moins de disposer d'un modèle précis du système enroulement-isolant (rarement le cas en pratique) il n'est ainsi possible de mesurer précisément que les décharges dans les traversées et les décharges entre la borne haute tension et la masse.

La technique dite de comparaison des profils est plus appropriée aux mesures sur les transformateurs. En disposant de plusieurs mesures simultanées, elle consiste à rechercher le profil de simulation qui se rapproche le plus du profil de la décharge. Une application typique de cette technique au cas des transformateurs de puissance est décrite dans [43]. En injectant un signal connu entre deux points accessibles sur le transformateur hors tension, et en recueillant les réponses sur tous les points accessibles, un profil représentant les coefficients de transmission est dressé. Un certain nombre de profils sont ainsi recueillis (autant que d'injections). Lorsque le transformateur est sous tension, les niveaux de décharges partielles sont mesurés aux différents points ayant servi à réaliser les profils d'étalonnage. La comparaison du nouveau profil à ces derniers permet de localiser du point de vue électrique la source de décharges partielles ; le cas le plus simple serait évidemment celui où le profil sous tension est analogue à celui d'une injection hors tension.

Sur site, la mise en œuvre telle quelle des techniques classiques décrites ci-dessus s'avère délicate voire impossible à cause des perturbations électromagnétiques externes provenant principalement des sous-stations et des lignes en service. Les données nécessitent un filtrage préalable. Les méthodes existantes sont basées sur la discrimination selon la polarité des décharges [17] et/ou le traitement numérique des signaux (filtrage fréquentiel et corrélations [167] ; analyse à résolution de phase [2]). Associées à des méthodes avancées d'analyse de profils [185], elles fournissent aujourd'hui d'excellents résultats.

### Méthode des ultrasons

Elle est basée sur la détection des ondes acoustiques (entre 70 et 150 kHz) émises par les décharges. Le lieu où ces dernières sont produites est déterminé géométriquement à partir de mesures extérieures au transformateur. Pratiquement, un dispositif de mesure est constitué de plusieurs capteurs qui sont répartis sur la paroi de la cuve du transformateur [36]. Pour chacun de ces détecteurs, le temps écoulé entre l'instant d'apparition de la décharge partielle et sa mesure permet de déterminer la distance source-capteur. Par triangulation, connaissant plusieurs de ces distances, il est possible de situer géométriquement la source. Malheureusement, l'appareillage détecte également les bruits provenant de l'extérieur du transformateur. Il faut donc effectuer un filtrage des mesures. Une seconde difficulté réside dans le fait que le transformateur n'est pas un milieu homogène. Les vitesses de propagation de l'onde émise, dépendent des matériaux traversés. Il faut donc en tenir compte lors du calcul des distances. Comme pour les mesures électriques, même si cette méthode s'avère peu précise pour déterminer l'énergie des décharges et nécessite un grand nombre de capteurs [167], les techniques de traitement du signal actuelles permettent d'obtenir des systèmes de localisation des décharges partielles performants comme celui décrit dans [121].

### 3.3.3 Fonction de transfert

D'un point de vue électrique, chaque transformateur peut être modélisé selon sa constitution. Compte tenu de la variété des types de transformateurs de puissance, chacun possède son modèle et donc respectivement sa propre réponse électrique. Ces paramètres s'avèrent être très sensibles aux moindres changements du système et donc potentiellement intéressants pour identifier des défauts (tab. 3.5).

La méthode de détection des défauts consiste à analyser la réponse fréquentielle du système. Les techniques classiques se basent sur l'utilisation d'un analyseur de réseau [2, 31]. Alors que le transformateur est déconnecté, l'analyseur génère des signaux sinusoïdaux à des fréquences précises. Le gain et le déphasage entre le signal d'entrée et celui qui parcourt l'enroulement du transformateur auquel il est appliqué

TAB. 3.5 – *Corrélations entre les défauts et la variation des paramètres électriques [144].*

Paramètre	Type de défaut
Inductance série	Déformation des enroulements
Capacité shunt	Mouvement des enroulements
Capacité série	Vieillessement de l'isolation
Résistance	Décharges partielles

sont ainsi évalués pour un domaine fréquentiel défini. Les variations des paramètres énumérés dans le tableau 3.5 en fonction des différents défauts s'illustrent par le déplacement des pôles et zéros, l'apparition de nouvelles fréquences d'oscillation ou l'affaiblissement de l'amplitude de certains pôles de la fonction de transfert ainsi obtenue. La comparaison de la réponse obtenue avec des résultats de simulation et/ou une base de données permet de détecter les défauts recherchés.

Parallèlement, depuis de nombreuses années, les transformateurs subissent des essais de choc normalisés [149, 158] pour éprouver leur tenue face à des contraintes diélectriques importantes. La référence [94] décrit une méthode permettant d'obtenir la fonction de transfert à partir des mesures de ces essais de chocs. Grâce à des résultats obtenus dans le domaine fréquentiel, cette méthode s'avère être indépendante de la forme et de l'amplitude de la tension de choc. En appliquant des impulsions d'entrée d'amplitude réduite (en anglais LVI pour "Low Voltage Impulse"), cette application particulière de l'essai de choc fournit une alternative à l'approche classique de détermination de la fonction de transfert.

Associée à l'utilisation de techniques avancées du traitement de l'information, cette méthode permet non seulement de déterminer les défauts mais elle a également le potentiel de les localiser précisément [173]. De plus, son indépendance par rapport à la forme et à l'amplitude du signal d'entrée lui permet d'être appliquée de manière "off-line" sans source de tension externe. Dans ce cas, seuls les disjoncteurs sont utilisés : ils contrôlent des impulsions qui correspondent à la mise sous tension du système. De la même manière, les surtensions de manœuvre et autres régimes transitoires apparaissant aux bornes du transformateur lorsqu'il fonctionne, permettent d'obtenir les fonctions de transfert alors

que l'unité est "on-line" [87]. Prometteuses, ces deux dernières techniques ne sont encore mises en œuvre que dans un cadre expérimental.

### 3.3.4 Réponse diélectrique

Les techniques classiques d'évaluation de l'état de l'isolation des composants électriques sont la mesure de la résistance en courant continu (DC) et la détermination du facteur de perte  $\tan\delta$ . La mesure de la résistance DC est difficilement applicable aux transformateurs de puissance. L'isolation composée de couches successives de papier, carton et huile forme un schéma équivalent relativement complexe. Lors de l'application d'une tension continue entre les bornes haute et basse tensions, un courant transitoire correspondant à la polarisation des différentes couches vient se superposer à la composante continue. Dans le cas d'un transformateur neuf, plusieurs heures sont nécessaires avant d'atteindre un régime continu stabilisé. En revanche, la mesure de  $\tan\delta$  (§ 3.2.3) fait partie des tests couramment effectués sur les transformateurs de puissance. Mises en œuvre pour la fréquence industrielle (50 ou 60 Hz), les mesures sont effectuées entre les enroulements primaire et secondaire respectivement court-circuités. Leurs résultats donnent de bonnes indications sur le vieillissement de l'isolation et plus particulièrement sur sa teneur en humidité [162]. Les grandeurs mesurées sont malheureusement fortement dépendantes de la composition de l'isolation. En toute rigueur, elles ne peuvent être utilisées que pour surveiller l'évolution de ces paramètres sur des transformateurs de même type.

Face aux limitations des méthodes traditionnelles, trois méthodes basées sur la dépendance temporelle ou fréquentielle des réponses diélectriques ont été développées. La première, connue sous le nom de spectroscopie diélectrique est en fait une généralisation de la mesure du facteur de pertes  $\tan\delta$  [49] pour une plage de fréquence donnée. La deuxième technique est appelée mesure de la tension de retour [16]. Elle est basée sur le phénomène suivant : l'application d'une tension continue pendant un temps donné, suivie d'une décharge en court-circuit de durée plus courte, engendre finalement une tension entre les enroulements primaire et secondaire si ceux-ci sont laissés en circuit ouvert. Enfin, la troisième méthode consiste à relever les réponses des courants de polarisation et dépolarisation, respectivement lors de l'application

d'une tension et après suppression de la tension et mise en court-circuit des enroulements vers la terre [67]. Pour chacune de ces techniques, en connaissant la géométrie de l'isolation et après avoir caractérisé les propriétés diélectriques de l'huile, il est possible de qualifier l'état, plus particulièrement la teneur en eau, du papier [49, 67].

## 3.4 Mesure des vibrations

### 3.4.1 Origine du bruit

Les transformateurs en service émettent un bruit caractéristique dont les causes essentielles sont énumérées ci-après [61, 65].

- a. La *magnétostriction des tôles* du circuit magnétique sous l'effet des variations de l'induction magnétique est la source de bruit prédominante. L'allongement et le raccourcissement périodique des tôles produit des vibrations dont la fréquence fondamentale est le double de celle du réseau, soit 100 Hz. Le phénomène n'étant pas linéaire, de nombreuses harmoniques s'y ajoutent.
- b. Les *forces de Lorentz* dues à l'interaction entre le flux commun et le courant dans les enroulements contribuent également de manière importante aux vibrations à basse fréquence.
- c. Enfin, le fonctionnement des *équipements auxiliaires* tels que les pompes et ventilateurs de réfrigération produisent des bruits caractéristiques qui viennent s'ajouter au spectre total (entre 500 et 2000 Hz).

Les vibrations sont transmises du circuit magnétique à la cuve soit par conduction directe aux points de liaison, soit au travers de l'huile et des isolants. La cuve et les radiateurs rayonnent ensuite l'énergie sonore vers l'extérieur. Des changements dans la structure des enroulements (modification de la position relative des enroulements HT et BT, diminution de la pression des enroulements après le serrage final en usine) ou du noyau ainsi qu'un changement de régime de fonctionnement de l'équipement auxiliaire se manifestent automatiquement par la modification de la signature acoustique.

### 3.4.2 Dispositif de mesure

Pour une plage de fréquence comprise entre 70 et 150 kHz, la signature acoustique du transformateur permet de localiser les décharges partielles (§ 3.3.2). Dans le cas présent, la technique généralement retenue consiste à mesurer directement les vibrations à l'aide d'accéléromètres fixés sur la cuve [99] ou à l'intérieur même du transformateur [54, 83, 136]. *Largement répandue pour la détection des défauts mécaniques sur les disjoncteurs [29, 111, 133]*, la méthode consiste à comparer les réponses temporelles ou fréquentielles relevées avec des courbes de référence. Dans le cas des transformateurs de puissance, il a été démontré expérimentalement que cette technique donne des indications intéressantes concernant le serrage des enroulements [99]. Toutefois, les résultats obtenus sont non seulement dépendants de la structure du transformateur mais aussi de la position des capteurs. L'interprétation des mesures est une question qui reste ouverte à ce jour.

## 3.5 Équipement auxiliaire

Comme l'ont montré les études commentées au paragraphe 2.5.1, les principales sources de défaillances proviennent de l'équipement auxiliaire. Celui-ci est composé de trois éléments ou groupes d'éléments.

Le premier d'entre eux est le dispositif de refroidissement. Sa surveillance se limite généralement aux signaux logiques de fonctionnement. Suivant les cas, des paramètres additionnels sont mesurés. Il s'agit des températures (huile et air) entre l'entrée et la sortie des échangeurs de chaleur ainsi que les vitesses des fluides [88, 166].

Les deux autres éléments, qui sont par ailleurs les parties les plus sensibles des transformateurs, font l'objet des deux prochains paragraphes.

### 3.5.1 Bornes de traversées

Les mesures effectuées sur les bornes de traversées isolées à l'huile sont les suivantes [52, 88, 126] :

- analyse visuelle du niveau et pression de l'huile ;
- mesure des paramètres diélectriques (capacité et facteur de pertes) ;
- analyse des gaz dissous (prélèvement d'huile) ;



- localisation de points chauds par thermographie.

Comme le montre cette liste, excepté la mesure de pression de l'huile, les techniques mises en œuvre ne sont pas de nature à être appliquées en continu dans le cadre d'une surveillance automatisée. Elle nécessitent toutes la planification de mesures ponctuelles rendant parfois difficile la détection de défauts à leur stade initial.

### 3.5.2 Changeurs de prises

Les méthodes les plus répandues pour la surveillance de l'état de l'insérteur en charge peuvent être réparties en quatre catégories [52] :

- mesure du profil de températures [19] ;
- analyse des gaz dissous (prélèvement d'huile) [88] ;
- enregistrement du couple [88], de la puissance [166] ou du courant moteur [19] ainsi que de la position du changeur de prise.
- relevé des vibrations ou de l'enveloppe acoustique lors des opérations de changement de prise [13, 78, 126].

Les deux premières méthodes permettent de détecter la surchauffe des contacts et les défaillances diélectriques. L'analyse des gaz ne peut malheureusement pas se faire "on-line" car les capteurs actuellement disponibles sont inopérants dans les huiles polluées du changeur de prises. Les mesures des paramètres des moteurs fournissent non seulement des indications relatives à leur état mais aussi à celui de tous les assemblages mécaniques du changeur de prises. La dernière technique citée permet quant à elle de prendre en compte à la fois les défaillances mécaniques et électriques. Sa mise en œuvre est simple, mais le traitement des résultats est relativement complexe.

## 3.6 Synthèse

Les méthodes de mesure passées en revue dans ce chapitre sont, pour la plupart d'entre elles, largement utilisées et permettent de prévenir de nombreux incidents. Le tableau 3.6 résume toutes les techniques recensées (équipements auxiliaires exclus) ainsi que leur champ d'application.

Ce tableau met en évidence une certaine redondance entre les différentes méthodes de mesure. Pourtant, seules les relations directes entre

TAB. 3.6 – Récapitulation des principales méthodes de surveillance.

Type de mesure	Défaillances et indicateurs de défaillances									
	Contamination de l'huile	Surcharge	Degradation de la cellulose	Decharges partielles	Bulles de gaz	Arc électrique	Defaut du noyau	Defaut des conducteurs et connexions	Deplacement/jeu des enroulements	
Température	-	×	-	-	-	×	×	×	-	-
Analyse des gaz (DGA)	×	×	×	×	×	×	×	-	-	-
Composés furaniques (HPLC) *	-	-	×	-	-	-	-	-	-	-
Humidité de l'huile	×	×	×	-	-	-	-	-	-	-
Mesure électrique des DP (CEI 60270) *	-	-	-	×	×	×	-	-	-	-
Signature acoustique HF	-	-	-	-	×	×	-	-	-	-
Vibrations	-	-	-	-	-	-	×	-	×	-
Fonction de transfert *	-	-	×	×	-	-	-	×	×	-
Réponse diélectrique *	×	-	×	-	-	-	-	×	-	-
Propriétés diélectriques de l'huile *	×	-	×	-	-	-	-	-	-	-
Paramètres physico-chimiques de l'huile *	×	-	×	-	-	-	-	-	-	-

\* Méthode off-line exclusivement

les mesures et les défaillances ou indices de défaillances sont prises en compte. La redondance serait encore plus grande si l'on tenait compte de toutes les relations de cause à effet. Par exemple, un défaut mécanique dû à une connexion qui a pris du jeu entraîne une augmentation locale de température. Celle-ci se manifestera par la production de gaz. Ainsi, indirectement, l'analyse des gaz permet de détecter des défauts de connexions.

Chacune de ces méthodes a cependant ses particularités (sensibilité, aptitude à être utilisée "on-line", etc.). Suivant la nature du cahier des charges, leur complémentarité permet d'accélérer la détection des défauts et d'améliorer la précision des diagnostics. En effet, ces techniques doivent encore être différenciées suivant leur précision et leur aptitude à être utilisée en continu ou non. Enfin, le tableau 3.6 montre l'importance de l'analyse des gaz dissous. Celle-ci constitue sans aucun doute l'élément clé de toute procédure de surveillance et de diagnostic de l'état des transformateurs de puissance.



# Chapitre 4

## Surveillance continue

### 4.1 Introduction

#### 4.1.1 Motivations

Le rôle de la surveillance continue ("monitoring on-line" en anglais) est de qualifier en permanence l'état de fonctionnement d'une machine ou plus généralement d'un processus industriel. Le but principal n'est pas de faire un diagnostic précis quant à la nature exacte d'un éventuel défaut mais de disposer de mesures fiables permettant de déceler rapidement toute anomalie.

Chaque transformateur de puissance constitue un dispositif complexe et coûteux conçu spécialement suivant son utilisation. En prenant en compte les contraintes économiques d'exploitation, l'intérêt d'un système de surveillance continue a été justifié à la section 1.1. Dans ce contexte, les objectifs visés sont les suivants :

- surveillance du système pour des conditions d'exploitation situées près ou au-delà des limites prescrites (réduction des facteurs de sécurité) ;
- détection de défauts graves (action automatique de protection) ;
- mise en évidence de défauts naissants (avant même qu'une défaillance ne survienne) ;
- qualification de l'état de vieillissement de l'unité afin d'aider à la planification du programme de maintenance.

La modélisation exacte des phénomènes physiques régissant le fonctionnement d'un transformateur de puissance est extrêmement complexe et fortement dépendante de sa structure. Toute unité possède ses caractéristiques propres et il n'existe pas de modèle générique global. Pour les mêmes raisons, il est n'est pas envisageable de faire des simulations fiables sur des modèles réduits. Les principales connaissances quant à l'évolution des différents paramètres de fonctionnement proviennent presque exclusivement des mesures expérimentales. Les informations relatives aux états défaillants sont donc relativement limitées et, comme cela est décrit ci-après, l'interprétation des mesures continues est particulièrement délicate. Concrètement, dans un contexte d'exploitation où les capacités des transformateurs sont utilisées au maximum et les périodes de maintenance limitées au minimum, même avec des règles d'interprétation simples, les techniques de mesure en continu permettent de limiter les risques d'endommagement et d'indisponibilité des unités. Le cas échéant, les résultats fournis doivent servir de base à la mise en œuvre de méthodes d'investigation plus poussées lors de l'arrêt de l'unité.

#### 4.1.2 Interprétation des mesures

Comme cela a été vu au chapitre précédent, les techniques permettant d'évaluer l'état des transformateurs de puissance sont nombreuses. Parmi elles, le tableau 3.6 en recense cinq susceptibles d'être utilisées en continu. Les premiers systèmes d'acquisition numériques sont apparus dans les années 1980 avec la généralisation de l'informatique [55, 123, 124]. Ceux-ci peuvent être dédiés à un capteur particulier ou intégrer les mesures fournies par plusieurs instruments ainsi que les alarmes logiques de fonctionnement. Aujourd'hui, un grand nombre de solutions industrielles sont disponibles sur le marché (cf. par exemple [48, 86, 91, 125, 166]). Cet afflux de nouvelles données soulève la question de leur interprétation et des techniques qui doivent être mises en œuvre. Avant d'élargir le propos aux méthodes d'analyse utilisées pour la surveillance de processus en général, les principaux travaux publiés dans le domaine des transformateurs de puissance seront passés en revue.

La première approche consiste à considérer l'acquisition et l'analyse des différents paramètres individuellement. Dans un grand nombre de cas, les modèles régissant les phénomènes considérés sont encore mal

connus. La méthode d'interprétation la plus simple consiste alors à définir un ou plusieurs seuils d'alarmes fixes. Comme cela sera exposé plus loin, cette manière de procéder comporte un certain nombre d'inconvénients (§ 4.4.1). En règle générale, ces niveaux d'alarmes sont complétés par des valeurs limites d'accroissement (tendance à court terme). Cette dernière solution (seuils fixes sur les valeurs absolues et les accroissements) permet de générer automatiquement des messages d'alarme à l'opérateur qui, le cas échéant, effectuera une analyse plus poussée des données [19, 48].

La température de l'huile, en tant que paramètre principal lié à la charge du transformateur, a été l'une des premières grandeurs surveillées en continu. La technologie liée aux capteurs est bien maîtrisée et l'expérience accumulée dans le domaine devient importante. La mesure de la température de l'huile permet non seulement d'identifier tout échauffement anormal (point chaud) mais également d'ajuster dynamiquement la limite de charge (§ 3.2.4). Comme cela a déjà été mentionné à plusieurs reprises, le paramètre déterminant est la température du point chaud. À quelques exceptions près, cette valeur n'est pas accessible directement : elle est généralement obtenue grâce à des modèles [145, 163] à partir de la température au sommet des enroulements. Cette dernière peut être mesurée directement ou calculée à partir du courant de charge et de la température ambiante (§ 4.4.3).

L'estimation indirecte de la température du point chaud est utilisée pour déceler une surchauffe ou calculer la vitesse de vieillissement du transformateur [51, 88]. Il est également possible de tirer parti du modèle établi pour la température de l'huile au sommet des enroulements ; dans le cas où cette mesure est disponible, une comparaison entre la sortie du modèle et la mesure permet de déceler un événement anormal [54]. Désignée par le terme de surveillance par modèle, cette technique a l'avantage de fournir un seuil dépendant des conditions de fonctionnement. Les modèles des guides de charges n'ont pas été développés pour cette application particulière : ils conviennent particulièrement bien lorsque les variations de température sont dues principalement au courant de charge. En pratique, la température de l'huile est fortement liée au système de refroidissement. Celui-ci est caractérisé par la variation journalière et annuelle de la température ainsi que par le régime

de fonctionnement des groupes de refroidissement (pour les unités disposant de ventilateurs). Partant de ce constat, les travaux [89] et respectivement [40, 41] décrivent des adaptations efficaces de ces modèles. Enfin, en conservant les mêmes entrées, le modèle a également été réalisé avec succès à l'aide de réseaux de neurones artificiels non supervisés [2, 60, 175].

Le taux de gaz dissous est certainement le paramètre le plus couramment mesuré après les températures. Les réactions entrant en considération sont malheureusement beaucoup plus compliquées. En continu seuls les détecteurs mono-gaz représentent aujourd'hui des solutions techniquement et économiquement intéressantes. Bien que des modèles mathématiques simples aient été développés [82], le traitement des mesures se limite généralement à la définition de seuils fixes sur les valeurs absolues et les tendances à court terme (augmentation journalière) [19, 48]. Lorsque les solutions techniques le permettront, les méthodes d'interprétation automatique développées pour les analyses périodiques pourront être adaptées au cas continu (§ 3.2.5). Celles-ci peuvent être subdivisées en trois classes : les systèmes experts [32, 69, 169, 188] ; les méthodes basées sur l'utilisation de réseaux de neurones artificiels supervisés [190] ; les techniques combinées [182]. Comme cela sera vu aux paragraphes 4.6.1 et 4.6.2, les systèmes experts sont des structures décisionnelles basées sur l'interprétation de règles logiques. De leur côté, les réseaux de neurones supervisés s'organisent tous seuls sur la base de couples entrées-sorties provenant de la fonction à réaliser (§ 4.6.3 et sec. 5.1). La combinaison des deux approches permet de prendre en compte à la fois les règles qui peuvent être formalisées et des relations plus subtiles parfois même inconnues des experts humains.

Plus ancienne, la représentation graphique dite du triangle de Duval [35] fournit une interface particulièrement conviviale. La subdivision de la surface d'affichage en zones correspondant à autant de diagnostics possibles permet à un opérateur d'évaluer la situation en un coup d'œil. Avec la même philosophie, [39] décrit une application mettant en œuvre les cartes auto-organisées dans ce contexte. Élément central de cette thèse, cet outil et son application au domaine sera décrit en détail aux chapitres suivants. Mentionnons simplement ici que ces cartes ont la propriété d'effectuer la projection d'un hyper-espace sur un espace géné-



ralement bidimensionnel tout en respectant la topologie de l'ensemble de départ. Ceci est réalisé grâce à un apprentissage non-supervisé durant lequel les poids des connexions s'adaptent itérativement (§ 5.4.3). Dans ce contexte particulier, l'apprentissage définit des zones qui pourront être interprétées comme le triangle de Duval. La différence fondamentale est que comme l'organisation n'a pas été dirigée, les résultats mettent en évidence des zones supplémentaires à celles connues initialement.

L'humidité dans l'huile, les décharges partielles (mesure acoustique) et les vibrations sont les trois dernières grandeurs couramment mesurées en continu. La littérature fait part d'applications mettant en œuvre l'utilisation de modèles linéaires et neuronaux respectivement pour l'humidité [130] et le spectre fréquentiel des vibrations [83]. Cependant, en pratique les mesures collectées sont généralement analysées avec des seuils fixes (valeur absolue et accroissements).

Comme cela a été mentionné plus haut, plusieurs paramètres parmi les cinq<sup>1</sup> qui viennent d'être passés en revue sont souvent mesurés simultanément. En prenant encore en compte les équipements auxiliaires, la quantité d'information à traiter devient très importante. Leur utilisation rationnelle et efficace passe inévitablement par un système intelligent capable de prendre automatiquement toutes ces données en compte. La généralisation de l'approche basée sur des modèles [2, 53, 125] à l'ensemble des capteurs fournit une série de signaux d'erreurs. Ceux-ci étant caractérisés par des seuils d'activation variables en fonction de l'état de fonctionnement (sec. 4.4), ils constituent une première étape vers un système intelligent. Sur site, ces signaux d'alarmes peuvent être interprétés selon un modèle logique simple prenant en compte les principales relations existant entre eux pour les modes de défaillance prédéfinis [26]. Selon la même approche que pour le traitement des gaz dissous, les méthodes les plus évoluées fournissent un diagnostic à l'aide de systèmes expert flous [5, 137]. Ceux-ci tiennent compte à la fois des paramètres de construction des unités, des mesures périodiques et des valeurs fournies en continu par les capteurs. Comme cela sera vu au chapitre 6, l'utilisation des cartes auto-organisées a également été étudiée dans ce contexte [80].

---

1. Pour mémoire, il s'agit des températures, du taux de gaz dissous, de l'humidité, des décharges partielles et des vibrations.

## 4.2 Analyses statistiques

### 4.2.1 Ensemble de données

Soit  $n$  le nombre de capteurs installés sur un transformateur de puissance. Suivant le type de surveillance voulu, ceux-ci peuvent mesurer l'évolution d'un éventail plus ou moins large de paramètres et présenter une certaine redondance (température mesurée en plusieurs points par exemple). Toutefois, tous fournissent des valeurs en continu au sens de la définition du chapitre 3. Les mesures sont collectées par un système d'acquisition et archivées pour un taux d'échantillonnage fixe. L'ensemble des données brutes peut être représenté sous forme de tableau (tab. 4.1).

TAB. 4.1 – Représentation des données brutes.

Observation	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	...	$\mathbf{v}_n$
1	$v_{11}$	$v_{12}$	$v_{13}$	...	$v_{1n}$
2	$v_{21}$	$v_{22}$	$v_{23}$	...	$v_{2n}$
⋮	⋮	⋮	⋮	⋮	⋮
m	$v_{m1}$	$v_{m2}$	$v_{m3}$	...	$v_{mn}$

Ces valeurs constituent ainsi une matrice  $\mathbf{V}$  de dimensions  $m \times n$  où chaque ligne se réfère à un instant d'échantillonnage  $k$  et chaque colonne à un capteur particulier  $j$ .

### 4.2.2 Concepts de base

Avant de décrire les méthodes statistiques les plus couramment utilisées pour la surveillance de processus, par soucis de clarté quelques définitions élémentaires seront rappelées ici (cf. [101] pour plus de détails).

Une variable aléatoire discrète  $X$  peut être caractérisée par son espérance mathématique  $E(X)$  et sa variance  $Var(X)$  :

$$E(X) = \sum_k x_k p_k \quad (4.1)$$

$$\text{Var}(X) = E((X - E(X))^2) = \sum_k (x_k - \mu)^2 p_k \quad (4.2)$$

où  $p_k$  est la probabilité  $P(X = x_k)$ . Si la variable aléatoire  $X$  a une distribution uniforme, l'espérance mathématique est alors égale à la moyenne. Celle-ci est généralement désignée par  $\mu$ . De manière équivalente, la variance est désignée par  $\sigma^2$ . On parle également d'écart-type  $\sigma$  pour la racine de la variance. Un ensemble multivariées de dimension  $n$ , sera caractérisé par une moyenne et une variance pour chacune des  $n$  variables :

$$\boldsymbol{\mu} = [\mu_1 \mu_2 \dots \mu_n] \quad (4.3)$$

$$\boldsymbol{\sigma}^2 = [\sigma_1^2 \sigma_2^2 \dots \sigma_n^2]. \quad (4.4)$$

En faisant l'hypothèse que la densité de probabilité de chacune des variables est uniforme, les équations (4.1) et (4.2) peuvent être remplacées par leurs estimateurs numériques définis ci-après. Pour les  $m$  observations  $n$ -dimensionnelles qui constituent la matrice  $\mathbf{V}$  l'espérance mathématique devient

$$\bar{v}_j = \frac{1}{m} \sum_{k=1}^m v_{jk} \quad (4.5)$$

et la variance

$$s_j^2 = \frac{1}{m-1} \sum_{k=1}^m (v_{jk} - \bar{v}_j)^2. \quad (4.6)$$

### 4.2.3 Mesures de corrélation

La covariance est une mesure de dépendance de deux variables aléatoires. Formellement, son expression est

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y). \quad (4.7)$$

En reprenant les données de la matrice  $\mathbf{V}$ , sous les mêmes conditions que précédemment, l'estimateur numérique  $s_{tu}$  est le suivant :

$$s_{tu} = \frac{1}{m-1} \sum_{k=1}^m (v_{tk} - \bar{v}_t)(v_{uk} - \bar{v}_u) \quad (4.8)$$

où  $j = t$  et  $j = u$  sont les indices de deux variables différentes.

Si deux variables aléatoires sont indépendantes, leur covariance est nulle. En dehors de ce cas particulier, l'interprétation de cet indicateur est difficile; elle dépend de l'amplitude des deux variables. Le coefficient de corrélation linéaire introduit une normalisation permettant de s'affranchir de ce problème :

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \quad (4.9)$$

et l'estimateur

$$r_{tu} = \frac{s_{tu}}{\sqrt{s_t^2 s_u^2}} = \frac{s_{tu}}{s_t s_u}. \quad (4.10)$$

Les valeurs de ces deux dernières expressions sont bornées dans l'intervalle  $[-1; 1]$ . Le coefficient de corrélation est unitaire pour deux variables se comportant de manière identique, négatif pour des variables de comportements opposés (par exemple  $X = -Y$ ) et nul pour deux variables indépendantes.

L'exploitation des expressions (4.9) et (4.10) permet de déceler certains liens existant entre les variables d'une base de données. Leur mise en œuvre peut être faite sur la totalité des échantillons (corrélation globale), sur plusieurs sous-ensembles (corrélation locale), en retranchant l'effet d'une variable par régression linéaire (corrélation partielle) ou en introduisant un décalage temporel afin de prendre en compte les constantes de temps (corrélogramme). Ces techniques sont détaillées et illustrées dans [101] et [119].

#### 4.2.4 Analyse en composantes principales

L'analyse en composantes principales (PCA en anglais) fait partie des méthodes multivariées; elle traite des cas où plusieurs caractéristiques d'un objet sont mesurées simultanément. L'idée directrice est la réduction de la dimension des données tout en conservant au maximum les variations des variables d'entrée. En d'autres termes, la PCA consiste à déterminer un ensemble réduit de variables maximisant la variance et dont la combinaison linéaire permet d'approcher les données originales.

Formellement, il s'agit d'une transformation linéaire de  $n$  variables corrélées  $v_1, v_2, \dots, v_n$  vers  $n$  nouvelles variables non corrélées  $c_1,$

$c_2, \dots, c_n$ . Elle est définie par la matrice ortho-normale  $\mathbf{U}$  de dimensions  $n \times n$  :

$$\mathbf{c} = \mathbf{U}^T(\mathbf{v} - \bar{\mathbf{v}}) \quad (4.11)$$

où  $\mathbf{v}$  et  $\bar{\mathbf{v}}$  sont respectivement les vecteurs des variables observées et leurs valeurs moyennes. La matrice de covariance  $\mathbf{S}_c$  des nouvelles variables peut être déduite par un simple changement de base

$$\mathbf{S}_c = \mathbf{U}^T \mathbf{S} \mathbf{U} \quad (4.12)$$

à partir de la matrice de covariance de l'ensemble original  $\mathbf{S}$ .

La condition de non corrélation des nouvelles variables est alors respectée si et seulement si la matrice  $\mathbf{S}_c$  est diagonale. Ses composantes non nulles sont donc les  $n$  valeurs propres  $\lambda_i$  de  $\mathbf{S}$  calculées à l'aide de l'équation caractéristique

$$|\mathbf{S} - \lambda_i \mathbf{I}| = 0 \quad (4.13)$$

et classées (indice  $i$ ) par ordre décroissant. Les colonnes de la matrice  $\mathbf{U}$  sont formées par les vecteurs propres associés aux valeurs propres. En toute généralité, la combinaison linéaire des variables centrées  $\mathbf{v} - \bar{\mathbf{v}}$  avec le vecteur propre  $\mathbf{u}_i$  selon l'expression (4.11) est dite  $i^{\text{ème}}$  composante principale. Elle est caractérisée par une moyenne nulle et une variance  $\lambda_i$  (maximale pour  $i = 1$ ).

Dans de nombreux cas, les dimensions des variables ne sont pas compatibles entre elles. Pour les rendre au moins statistiquement comparables, comme on le verra plus loin pour d'autres techniques (§ 5.5.2), elles sont souvent centrées et réduites. La matrice de covariance ainsi normalisée correspond alors à la matrice de corrélation  $\mathbf{R}$ . Ainsi les résultats peuvent être sensiblement différents suivant que la matrice  $\mathbf{S}$  ou  $\mathbf{R}$  est utilisée.

En plus des possibilités relatives à l'identification des liens entre les différentes variables, l'intérêt de l'analyse en composantes principales est multiple pour la surveillance de processus. La première caractéristique utile est la réduction de la complexité. Comme le montre la figure 4.1 pour les mesures provenant d'un transformateur de puissance, un nombre restreint de composantes suffit à représenter l'information fournie avec une précision satisfaisante : dans le cas considéré, l'essentiel de

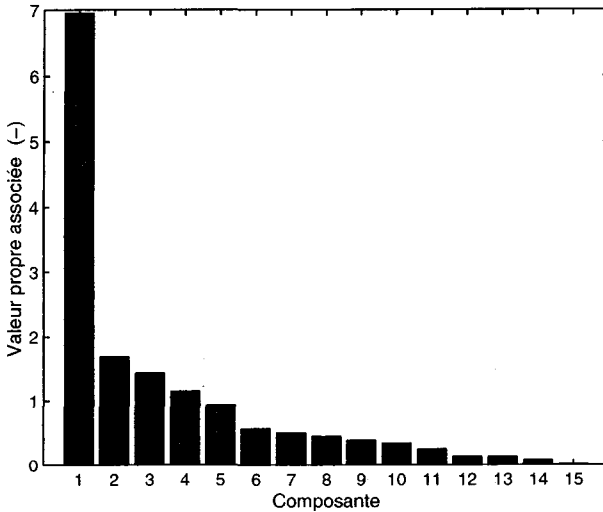


FIG. 4.1 – *Décomposition en composantes principales de 15 variables caractérisant l'état d'un transformateur.*

la variance peut être expliquée par les 5 premières composantes (la première d'entre elles représente même près de 50 % de la variance). Pour autant que le processus ait un comportement suffisamment stable, il est ainsi possible de déceler des anomalies suivant l'évolution des variables en fonction d'un nombre réduit de ces composantes [181].

La transformation opérée à partir de variables corrélées vers de nouvelles variables non corrélées est une propriété supplémentaire particulièrement intéressante pour la surveillance. Généralement, les variables d'un processus sont prises en compte séparément avec des limites d'acceptation basées sur les variances des mesures (processus considéré gaussien). Cette technique s'avère être insuffisante car elle ne prend pas en compte les corrélations existant entre les variables. L'utilisation des composantes principales, combinée avec une mesure de variabilité totale permettent d'affiner sensiblement la qualité des résultats (cf. [75] pour plus de détails).

### 4.2.5 Méthodes de coalescence

Les méthodes de coalescence ("clustering" en anglais) constituent un second groupe de méthodes permettant de rechercher l'information contenue dans des grandes bases de données. Le principe est de regrouper les  $m$  vecteurs dans  $\mathbb{R}^n$  en  $N$  groupes de sorte à ce que chaque vecteur soit plus proche de tout vecteur du même groupe que de tout vecteur d'un groupe différent. Le nombre de classes limité permet de réduire la complexité du problème et donc de tirer plus aisément les informations que ces vecteurs contiennent.

Ces techniques d'analyse de données sont basées sur l'application de méthodes de quantification vectorielle qui seront vues plus loin dans un cadre plus général (sec. 5.2). La détermination du nombre de classes nécessaires pour représenter de manière adéquate un processus est un problème complexe qui ne sera pas abordé ici. Le lecteur se référera à [34] pour une description plus complète dans le contexte du diagnostic de processus.

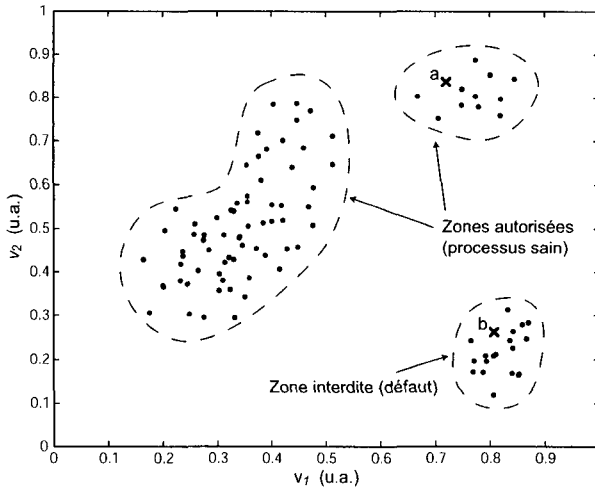


FIG. 4.2 – Détection de défaut par la méthode de coalescence. Processus caractérisé par deux variables de mesures: (a) état sain; (b) état défaillant.

Si l'ensemble de base comprend des données correspondant aux différents modes de fonctionnement admissibles ainsi qu'à toutes les défaillances envisageables, les méthodes de coalescence peuvent être utilisées pour identifier directement des défauts (fig. 4.2). L'ensemble des états issus de mesures préalables (base de données) sont subdivisés en classes distinctes correspondant à autant de modes de fonctionnement. À chaque instant de fonctionnement, le vecteur de mesure est associé à l'une des classes. L'état est ainsi identifié au mode de fonctionnement représenté par la classe sélectionnée. Cette manière de faire est malheureusement impossible à appliquer telle qu'elle pour la surveillance de transformateurs car les caractéristiques des modes défaillants ne sont pas connues avec précision.

## 4.3 Évaluation des tendances

### 4.3.1 Constantes de temps

Qu'il s'agisse de grandeurs originales, de composantes principales ou de grandeurs issues d'un quelconque traitement, la surveillance de ces valeurs s'opère sur deux niveaux bien distincts.

- a. L'observation de l'évolution à court terme permet d'identifier les défaillances graves résultant de changements rapides.
- b. La surveillance de l'évolution à long terme met en évidence un éventuel stade de vieillissement avancé ou des défauts naissants.

Le traitement de la première famille de phénomènes fait l'objet des sections 4.4 et 4.5. Le principe général est l'évaluation des valeurs instantanées par rapport à un seuil. On verra que celui-ci peut être fixe ou s'adapter en fonction des conditions. En revanche, l'évolution à long terme, doit être évaluée séparément.

### 4.3.2 Décomposition structurelle

Une série temporelle peut être décomposée en éléments distincts répartis suivant trois catégories fondamentales [58, 119] :

$$\text{Série} = \text{Tendance} + C_{\text{périodique}} + C_{\text{irrégulière}} \quad (4.14)$$



où *Tendance* est la variation à long terme (§ 4.3.3),  $C_{\text{périodique}}$ , respectivement  $C_{\text{irrégulière}}$  les composantes périodique et irrégulière. Les éléments appartenant à cette dernière catégorie sont totalement indépendants de toute évolution temporelle ; à l'exception du bruit (précision de la mesure, événement aléatoire, etc.) ils sont complètement déterminés par l'état du processus. Les variations périodiques sont dictées quant à elles, par l'évolution de phénomènes dont la période peut être variable. L'évolution saisonnière de la température ambiante moyenne en est une bonne illustration. Elle est déterminée uniquement en fonction du temps.

### 4.3.3 Évolution à long terme

Le calcul des tendances (en anglais "trends") des séries chronologiques constituées pas les mesures des variables caractérisant un processus, joue un rôle essentiel pour l'évaluation de son évolution à long terme. Ces valeurs peuvent aussi bien être déduites des mesures originales, des variables issues d'un traitement mathématique tel que l'analyse en composantes principales (§ 4.2.4) ou, comme on le verra plus loin (§ 4.4.2), des paramètres de modèles adaptés périodiquement.

En toute généralité, il n'existe aucune technique automatique infaillible permettant de déterminer une tendance. Celle-ci peut être de n'importe quelle nature (linéaire, exponentielle, croissante, décroissante, etc.). Les méthodes d'identification de tendance consistent généralement à trouver les paramètres de la fonction qui est la plus proche de la série temporelle au sens des moindres carrés [101]. Dans la plupart des cas, cette fonction est choisie linéaire et la période de régression est déterminée de telle sorte que cette hypothèse soit vérifiée. L'éventuelle composante alternative devra être préalablement supprimée si sa période est proche ou supérieure à l'intervalle de régression. La tendance ainsi obtenue est caractérisée par les coefficients  $\hat{\alpha}$  et  $\hat{\beta}$  du modèle linéaire d'une série temporelle  $y(t) = \alpha + \beta t + \varepsilon$ . Son influence est éliminée en retranchant la valeur  $\hat{u}(t)$  à la série

$$\hat{u}(t) = \hat{\alpha} + \hat{\beta}t. \quad (4.15)$$

La problématique de la surveillance de processus implique que la tendance d'une série temporelle peut non seulement suivre une fonction bien définie, mais aussi changer subitement de comportement suite à une

modification des conditions (défaut naissant, changement des conditions de fonctionnement, etc.). Dans ce contexte, la tendance doit être évaluée périodiquement. Les critères décisionnels sont alors la valeur de la dérive  $v_{max}$  et la vitesse  $\hat{\beta}$ . La situation générale est décrite par le premier critère pour lequel des limites absolues sont définies. Le second paramètre permet de pondérer l'observation faite en fonction de sa dynamique.

## 4.4 Méthode des résidus

### 4.4.1 Remarques préliminaires

La détection des changements rapides s'opère généralement en fixant un ou plusieurs seuils d'alarmes pour chaque grandeur mesurée. Dans la plupart des cas, deux seuils distincts sont définis : lorsque le premier d'entre eux est atteint, un simple signal d'avertissement est généré ; si le second seuil est dépassé, le processus est interrompu (arrêt d'urgence).

Pour un processus dont les paramètres mesurés varient fortement en fonction des conditions de fonctionnement, l'ajustement du niveau des seuils est très critique. Même si des techniques statistiques comme celles décrites au paragraphe 4.5.4 peuvent être appliquées dans ce contexte, il est souvent difficile voire impossible de trouver un compromis entre la robustesse et la sensibilité de détection. Cette problématique est bien illustrée par la figure 4.3. Si le niveau du seuil est bas, lorsque l'état de fonctionnement est caractérisé par des variables de valeurs élevées, l'erreur de type I (probabilité qu'une alarme soit déclenchée alors que le processus est bon) est grande. Ceci se manifeste par l'apparition fréquente d'alarmes intempestives ("fausses alarmes"). À l'opposé lorsque le niveau du seuil est haut, il se peut que pour toute une plage de fonctionnement le système de détection soit insensible à un problème (erreur de type II). En conséquence, la détection précoce des défauts peut s'avérer moins efficace voire inopérante.

Considérons le cas particulier des transformateurs de puissance. Leur état peut être surveillé simultanément par plusieurs capteurs mesurant des paramètres physiques différents. Du fait de sa simplicité, la technique des seuils fixes est couramment utilisée indépendamment pour chacun de ces dispositifs [48, 166]. Cependant, en plus des limitations

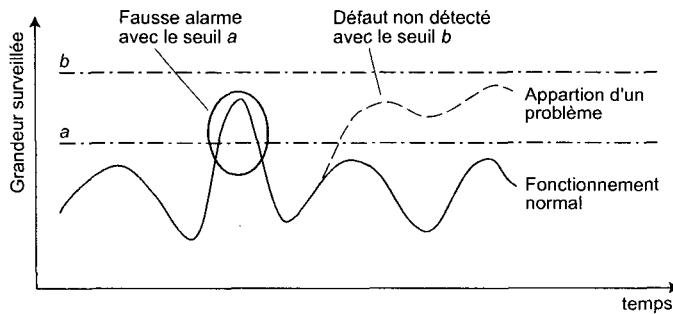


FIG. 4.3 – Génération d'un signal d'alarme avec un seuil fixe : le seuil ( $a$ ) est fixé trop bas, des fausses alarmes sont générées trop fréquemment ; le seuil ( $b$ ) est trop haut, la sensibilité n'est pas assez grande.

déjà évoquées, cette manière de procéder implique que les liens existants entre l'évolution des différents paramètres ne sont pas pris en compte. De plus, la nature unique de chaque transformateur ainsi que les considérations liées à leur coût ont pour conséquence qu'il n'existe que très peu de connaissances (théoriques ou expérimentales) permettant de fixer de manière infaillible les différents seuils [2].

#### 4.4.2 Principe

Les limitations énoncées ci-dessus ne sont pas spécifiques aux transformateurs de puissance. Les références [42, 47, 187] permettent de faire un tour d'horizon des solutions classiques qui ont été élaborées dans le domaine général de la surveillance de processus. Ces méthodes peuvent être subdivisées en deux catégories :

- la *redondance physique* qui consiste à multiplier le nombre de capteurs et à tester la consistance des informations qu'ils fournissent ;
- la *redondance analytique* qui est basée sur le calcul des résidus entre les valeurs réellement mesurées et celles fournies par des modèles mathématiques.

La première méthode s'avère onéreuse en raison du nombre élevé d'instruments de mesure qu'elle nécessite. En pratique son emploi se li-

mite aux cas où la sécurité est critique. La seconde méthode en revanche est beaucoup plus économique. Son principe est basé sur la modélisation des variables observées : à chaque instant, les valeurs calculées sont comparées aux valeurs mesurées. En fonction de l'amplitude des résidus obtenus, une alarme est alors déclenchée. Comme on le verra plus loin, étant donné la nature non linéaire des processus, les réseaux de neurones artificiels se prêtent bien à la construction de ces modèles [103, 122].

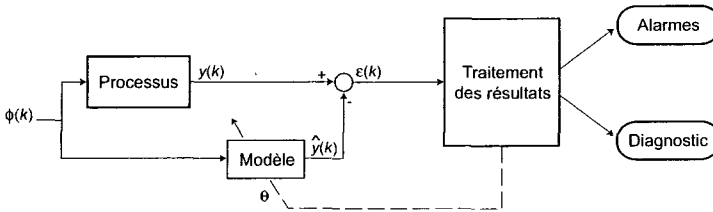


FIG. 4.4 – Schéma de principe de la surveillance suivant la méthode des résidus.

La première application de cette technique aux transformateurs de puissance a été réalisée au MIT en 1989 [54]. La figure 4.4 représente son schéma de principe. Pour chaque mesure, un modèle du comportement normal est construit selon l'expression générale

$$\hat{y}(k) = f(\phi(k), \theta) \quad (4.16)$$

où  $f(\phi(k), \theta)$  est définie selon le modèle choisi (§ 4.4.3 à 4.4.5),  $k$  est l'indice d'échantillonnage,  $\hat{y}$  l'estimateur de la valeur du capteur,  $\phi$  le vecteur des variables considérées et  $\theta$  le vecteur des paramètres du modèle.

Indépendamment de la modélisation utilisée, le vecteur  $\theta$  est déterminé par identification des modèles avec les observations réelles. Le comportement des transformateurs n'étant par ailleurs pas figé dans le temps, ces paramètres sont périodiquement recalculés. L'évolution de ces derniers est soumise aux calculs de tendance pour la surveillance des phénomènes lents.

### 4.4.3 Modélisation physique

Les techniques de génération de résidus les plus répandues sont basées sur la modélisation physique des phénomènes où chaque variable est caractérisée par son propre modèle [25, 53, 137]. Les connaissances sur les processus considérés permettent d'établir des expressions mathématiques qui sont pondérées par des facteurs qu'il s'agit de déterminer par identification.

L'exemple le plus classique est la modélisation de la température de l'huile au sommet des enroulements. Il existe plusieurs variantes de modèles [41, 89] basées sur celui établi par la norme CEI 60354 [152] :

$$\widehat{T}_h(k) = \theta_1(T_h(k-1) - T_a(k-1)) + \theta_2 I_{ch}(k)^{1.6} + T_a(k) \quad (4.17)$$

avec  $T_h$  la température de l'huile,  $T_a$  la température de l'air,  $I_{ch}$  la charge,  $\theta_1$  et  $\theta_2$  des constantes.

Le taux d'humidité relative  $RH_h$  constitue un autre exemple (modèle selon [25]) :

$$\widehat{RH}_h(k) = \alpha \exp\left(\frac{T_h(k)}{\beta}\right) \quad (4.18)$$

où  $\alpha$  et  $\beta$  sont deux constantes. En posant  $\alpha \equiv \exp\theta_1$  et  $\beta \equiv \frac{1}{\theta_2}$ , l'équation (4.18) devient

$$\ln\left(\widehat{RH}_h(k)\right) = \theta_1 + \theta_2 T_h(k). \quad (4.19)$$

Les relations (4.17) et (4.19) peuvent être mises sous la forme  $\widehat{y}(k) = \theta_1 \phi_1(k) + \theta_2 \phi_2(k) + \dots + \theta_j \phi_j(k)$  ou plus généralement  $\widehat{y}(k) = \boldsymbol{\theta}^T \boldsymbol{\phi}(k)$ . Une telle formulation peut être obtenue pour les modélisation physiques les plus courantes. Dans ce cas, les paramètres sont obtenus par minimisation de l'erreur quadratique (§ 4.4.4).

### 4.4.4 Modélisation linéaire

L'approche développée au paragraphe 4.4.3 conduit généralement à une modélisation linéaire de chaque variable. Une telle représentation peut également être décidée d'emblée si le comportement physique du processus n'est pas bien défini. Dans ce cas l'expression (4.16) devient

$$\widehat{y}_i(k) = \boldsymbol{\theta}_i^T \boldsymbol{\phi}(k) \quad (4.20)$$

pour la variable  $i$  et plus généralement

$$\hat{\mathbf{y}}(k) = \Theta^T \phi(k) \quad (4.21)$$

pour l'ensemble des variables, où  $\Theta$  est la matrice des coefficients définissant la fonction linéaire et  $\phi$  le vecteur des variables d'entrée (variables d'état et conditions initiales).

Soit  $\mathbf{Y}$  une matrice composée de  $m$  observations consécutives de  $n$  variables  $\mathbf{y}(k)$ .  $\Phi$  est alors la matrice  $m \times n$  des valeurs des variables d'entrées associées. Les éléments de  $\Theta$  sont déterminés par la minimisation de l'erreur quadratique  $E(\Theta)$  (moindres carrés [101, 119]) :

$$E(\Theta) = \frac{1}{2} (\mathbf{Y} - \Theta^T \Phi)^T (\mathbf{Y} - \Theta^T \Phi). \quad (4.22)$$

La solution de  $\min(E(\Theta))$  est donnée par

$$\Theta = (\Phi^T \Phi)^{-1} \Phi \mathbf{Y} \quad (4.23)$$

qui n'est définie que lorsque la matrice  $\Phi^T \Phi$  n'est pas singulière. Dans le cas contraire, la pseudo-inversion est utilisée :

$$\Theta = \lim_{\lambda \rightarrow 0} (\Phi^T \Phi + \lambda \mathbf{I})^{-1} \Phi \mathbf{Y}. \quad (4.24)$$

Cette technique est simple à mettre en œuvre et les coefficients sont faciles à calculer numériquement (opération peu gourmande en temps de calcul). Cependant si le processus étudié n'est pas linéaire, la qualité des modèles ne sera pas bonne.

#### 4.4.5 Modélisation neuronale

L'approche neuronale est une généralisation de la méthode linéaire. Dans ce cas, aucune connaissance du modèle physique n'est nécessaire a priori. Le modèle est constitué d'une série d'opérateurs mathématiques élémentaires interconnectés au sein d'une structure appelée réseau de neurones artificiels. Sous certaines conditions, certains types de réseaux de neurones permettent de représenter n'importe quelle fonction non-linéaire (sec. 5.1). Dans le cas particulier, ils ont pour tâche de réaliser la fonction

$$\hat{\mathbf{y}}(k) = f(\phi(k), \theta). \quad (4.25)$$

Sans entrer dans les détails, la détermination des paramètres  $\theta$  est un problème d'optimisation non linéaire. Celui-ci doit être résolu numériquement à partir de la matrice des observations  $\mathbf{Y}$ . La technique la plus courante, appelée rétropropagation du gradient d'erreur, est basée sur une descente stochastique du gradient [59]. Cette phase est appelée apprentissage.

La souplesse de la structure est telle qu'elle permet de modéliser les différentes variables composant le vecteur  $\mathbf{y}$  globalement ou individuellement en utilisant respectivement un seul ou plusieurs réseaux. Malgré un coût numérique supérieur à celui des méthodes classiques, les techniques neuronales présentent l'intérêt de prendre en compte les caractéristiques non linéaires des processus modélisés. Elles ont été appliquées avec succès à des processus industriels [27, 103, 104, 118] et en particulier au cas des transformateurs de puissance [2].

## 4.5 Détection des changements

### 4.5.1 États de sortie

Quelle que soit la modélisation retenue, la technique des résidus permet de générer pour chaque variable  $i$  une mesure  $\varepsilon_i$  à chaque instant d'échantillonnage  $k$  :

$$\varepsilon_i(k) = |y_i(k) - \hat{y}_i(k)|. \quad (4.26)$$

Ces valeurs sont nulles tant que les modèles établis sur la base du comportement normal correspondent au processus réel. Si une valeur  $\varepsilon_i$  dépasse un certain seuil défini en fonction de la précision du modèle associé, celui-ci n'est plus en accord avec les observations et une action doit être entreprise.

L'ensemble forme un système de  $n$  indicateurs numériques mesurant la déviation du processus par rapport aux états rencontrés dans le passé. Le dépassement d'un seuil pour l'une des variable n'implique pas obligatoirement l'apparition d'un problème ; il peut s'agir simplement d'un état de fonctionnement normal non modélisé. Par mesure de précaution, cet événement sera tout de même traité comme s'il s'agissait d'une avarie. On parle alors plus généralement de détection de changement.

### 4.5.2 Seuil d'alarme intelligent

Chaque signal  $\varepsilon_i$  peut s'apparenter à la sortie d'un système de surveillance à seuil adaptatif. En effet, son niveau est fixé par rapport à l'estimation de la variable  $i$  donnée par le modèle. La valeur obtenue dépend donc de l'état du processus et le seuil peut ainsi être ajusté avec plus de précision que dans le cas représenté à la figure 4.3.

D'un point de vue statistique, les problèmes liés à la valeur de ce seuil subsistent tout de même (erreurs de type I et II). Compte tenu de l'incertitude sur les modèles, un léger dépassement du seuil de l'une des valeurs pour un intervalle limité dans le temps n'est pas forcément synonyme de défaut (fig. 4.5). Dans ce cas, au contraire de ce qu'indique l'interprétation directe des résultats, aucune alarme ne devrait être déclenchée. Le dépassement du seuil, même adaptatif n'est donc pas un critère satisfaisant en soit. L'amplitude et de la durée du dépassement doivent être prises simultanément en compte pour générer un signal d'erreur suffisamment robuste.

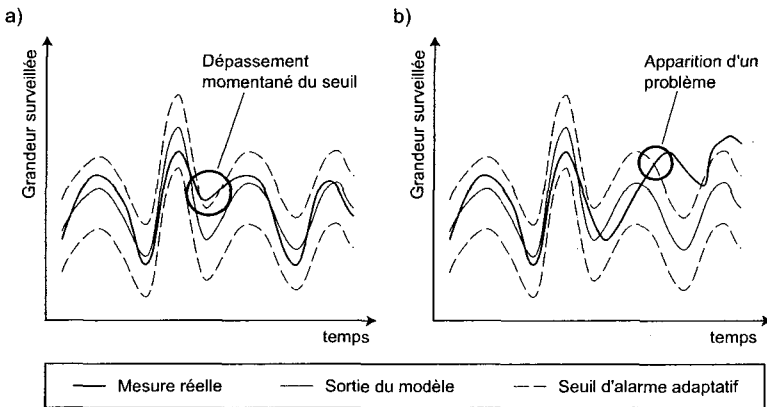


FIG. 4.5 – Génération d'un signal d'alarme avec seuil adaptatif: (a) dépassement momentané de faible amplitude du seuil ne devant pas entraîner d'alarme; (b) détection d'un problème par dépassement du seuil.



### 4.5.3 Rapport de vraisemblance

Les méthodes de détection de changement sont relatives à la théorie de la décision et plus particulièrement aux tests d'hypothèses [6, 7, 96]. Elles sont basées sur des tests statistiques permettant de prendre une décision à partir de résultats expérimentaux par rapport à un certain nombre d'hypothèses à contrôler.

Soit  $E_i$  un processus stochastique représentant la valeur du résidu  $\varepsilon_i(t)$  caractérisé par une densité de probabilité  $p_{\vartheta}(\varepsilon_i)$ . La valeur du paramètre  $\vartheta$  est déterminée d'après les deux hypothèses suivantes :

$$\begin{cases} \mathbf{H}_0 : \vartheta = \vartheta_0 & \text{si le résidu est nul,} \\ \mathbf{H}_1 : \vartheta = \vartheta_1 & \text{si le résidu atteint un certain seuil.} \end{cases} \quad (4.27)$$

Le logarithme du rapport de vraisemblance  $s_i$  est alors défini par l'expression

$$s_i = \ln \left( \frac{p_{\vartheta_1}(\varepsilon_i)}{p_{\vartheta_0}(\varepsilon_i)} \right) \quad (4.28)$$

où  $p_{\vartheta_0}$  et  $p_{\vartheta_1}$  sont les densités de probabilités associées respectivement aux hypothèses  $\mathbf{H}_0$  et  $\mathbf{H}_1$ . La valeur  $s_i$  est définie négative tant que la sortie du processus est assimilée à l'hypothèse  $\mathbf{H}_0$  ( $p_{\vartheta_0} > p_{\vartheta_1}$ ). Dans le cas contraire, l'hypothèse  $\mathbf{H}_1$  est vérifiée et le logarithme du rapport de vraisemblance est positif. Cette valeur constitue le signal d'entrée du dispositif robuste de détection de changement d'état qui est décrit ci-après.

### 4.5.4 Mesure robuste du changement

Comme cela a été mentionné au paragraphe précédent, lorsque le processus est sain, le comportement typique du rapport de vraisemblance logarithmique est négatif. Tout défaut sera donc identifié par une valeur positive de  $s_i$ . Par ailleurs, en introduisant un traitement cumulatif, on montre qu'il est possible de s'affranchir des variations statistiques des variables de sortie et de disposer d'un signal d'alarme robuste [117].

Dans ce cas, pour un instant d'échantillonnage  $k$ , la somme de tous les  $s_i$  intervient. Un changement est alors détecté lorsque la fonction de décision  $g_k$  remplit les conditions suivantes :

$$g_k = S_k - m_k \geq h \quad (4.29)$$

avec

$$S_k = \sum_{i=1}^k s_i \quad (4.30)$$

$$m_k = \min_{1 \leq j \leq k} S_j \quad (4.31)$$

et  $h$  est le seuil d'alarme. La figure 4.6 montre le comportement typique des grandeurs  $S_k$  et  $g_k$  lors de l'apparition d'un changement. L'instant du déclenchement de l'alarme est défini par

$$t_a = \min \{k | g_k \geq h\}. \quad (4.32)$$

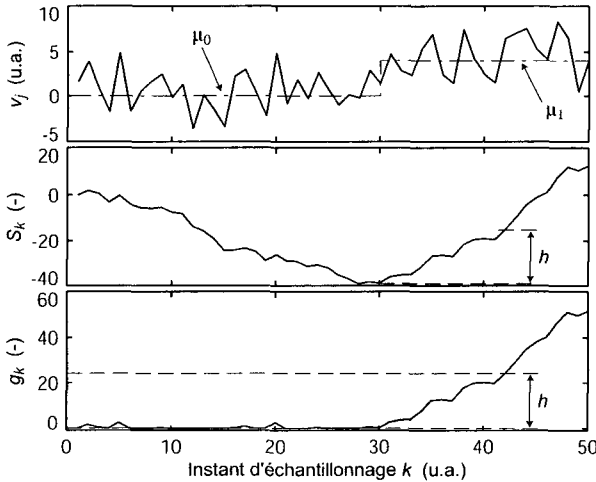


FIG. 4.6 – Principe de détection de changement : variations de la somme des rapports de vraisemblance logarithmique et de la fonction de décision pour un signal d'entrée gaussien dont la valeur moyenne  $\mu$  change.

Afin de fournir une mesure plus adaptée à la mise en œuvre de la méthode pour un suivi continu du processus, l'expression de  $g_k$  peut être écrite récursivement :

$$g_k = \begin{cases} g_{k-1} + s_k & \text{si } g_{k-1} + s_k > 0 \\ 0 & \text{si } g_{k-1} + s_k \leq 0 \end{cases} \quad (4.33)$$

avec  $g_0 = 0$ . Plus succinctement, avec la définition  $(x)^+ = \sup(0, x)$ , il vient :

$$g_k = (g_{k-1} + s_k)^+. \quad (4.34)$$

Faisons l'hypothèse que le comportement de la variable observée est gaussien. Il est caractérisé par une valeur moyenne  $\mu$  et une variance  $\sigma^2$ . Par ailleurs,  $\mu$  constitue le paramètre de changement  $\vartheta$  :

$$p_\vartheta = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right). \quad (4.35)$$

Pour un changement de moyenne brusque de  $\mu_0$  à  $\mu_1$ , le rapport  $s_i$  s'exprime alors par

$$s_i = \frac{\mu_1 - \mu_0}{\sigma^2} \left( y_i - \frac{\mu_0 + \mu_1}{2} \right) \quad (4.36)$$

et la somme devient

$$S_k = \frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=1}^k \left( y_i - \frac{\mu_0 + \mu_1}{2} \right). \quad (4.37)$$

Dès lors, la fonction de décision prend la forme suivante :

$$g_k = \left[ g_{k-1} + \frac{\mu_1 - \mu_0}{\sigma^2} \left( y_k - \frac{\mu_0 + \mu_1}{2} \right) \right]^+. \quad (4.38)$$

La fonction (4.38) ne permet de détecter des changements que si  $\mu_1 > \mu_0$ . Pour un saut de valeur connu  $\nu$ , la valeur moyenne après changement peut être  $\mu_1^+ = \mu_0 + \nu$  ou  $\mu_1^- = \mu_0 - \nu$ . Il est alors nécessaire d'utiliser deux fonctions de décision en parallèle [117]. L'expression déterminant l'instant du déclenchement d'alarme devient

$$t_a = \min \{k | (g_k^+ \geq h) \cup (g_k^- \geq h)\} \quad (4.39)$$

avec

$$g_k^+ = \left( g_{k-1}^+ + y_k - \mu_0 - \frac{\nu}{2} \right)^+$$

$$g_k^- = \left( g_{k-1}^- - y_k + \mu_0 - \frac{\nu}{2} \right)^+.$$

Par rapport à l'expression (4.38), le terme multiplicateur  $\frac{\mu_1 - \mu_0}{\sigma^2}$  a disparu : il est pris en compte dans le niveau du seuil  $h$ .

Le paramètre initial  $\vartheta_0$  est connu et vaut zéro (résidu). En revanche, l'amplitude du changement est généralement inconnue. Le rapport (4.36) est donc fonction de deux paramètres indépendants : l'instant de changement  $t$  et le paramètre  $\vartheta_1$ . L'approche la plus courante est de fixer l'amplitude du changement  $\nu$  a priori en fonction d'un seuil de détection minimum (voir [7] pour plus de détails).

## 4.6 Surveillance intelligente

Comme cela a été vu pour le cas des transformateurs (§ 4.1.2), les dispositifs de surveillance fournissent des alarmes plus ou moins robustes pour chacune des variables enregistrées. La quantité d'information ainsi fournie peut être relativement importante. Afin de faire un diagnostic de bonne qualité, il s'agit de sélectionner les points les plus importants et de les mettre en relation les uns avec les autres. Bien qu'il ne s'agisse plus à proprement parler de surveillance au sens strict, les lignes qui suivent décrivent brièvement les méthodes disponibles pour traiter ce problème de façon automatisée.

### 4.6.1 Systèmes experts

Les outils informatiques utilisés pour ce type de diagnostic sont appelés systèmes experts [183]. Sans entrer dans les détails, il s'agit d'algorithmes interprétant les informations fournies par le système de monitoring à l'aide de règles logiques. Celles-ci sont établies sur la base des connaissances d'experts humains. Elle donnent non seulement un diagnostic mais conseillent également les opérateurs quant aux mesures à prendre. Le formalisme utilisé ainsi que l'accumulation des informations constituent les points essentiels pour l'élaboration d'un système expert efficace.

La littérature faisant état de l'application de ces méthodes à la surveillance des transformateurs est nombreuse. À cet égard, [32] décrit l'un des premiers travaux dans le domaine des transformateurs de puissance. Les informations constituant les paramètres d'entrées des systèmes experts sont non seulement les valeurs des différents capteurs, les résidus générés et les signaux d'alarmes mais aussi toutes les informations rela-

tives à l'exploitation (informations relatives à la maintenance, tendances, etc.) [137]. Étant donné l'importance de la mesure des gaz dissous, les règles d'interprétation les plus nombreuses concernent ces dernières valeurs (cf. par exemple [182]).

#### 4.6.2 Logique floue

Classiquement, les règles de logique reposent sur la possibilité pour une variable de prendre deux valeurs : "vrai" ou "faux" (0 ou 1). Les variables sont combinées dans des propositions exprimées selon une syntaxe établie (fonctions logiques) dont le résultat n'est autre qu'une variable logique.

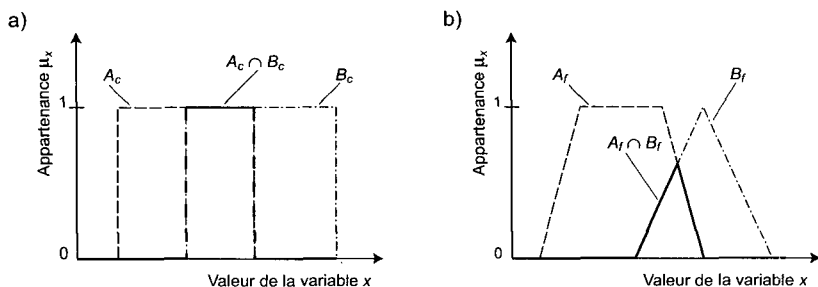


FIG. 4.7 – Exemples de définition d'un ensemble : (a) définition selon la logique formelle classique ; (b) définition selon la logique floue.

Établie sur une base théorique solide, cette représentation formelle s'avère souvent trop rigide pour prendre en compte toutes les nuances nécessaires. La logique floue ou logique à plusieurs valeurs permet de mieux représenter le raisonnement nuancé d'un expert humain. Comme le montre la figure 4.7, en logique classique la notion d'appartenance à un ensemble est caractérisée par une variable booléenne. Le concept de logique floue permet de définir des zones pour lesquelles cette fonction d'appartenance se situe entre ces deux valeurs. La notion d'incertitude (au sens de la modélisation de l'expertise humaine) est ainsi prise en compte. La combinaison de ces variables au sein d'une proposition est définie par des règles spécifiques [189] fournissant une valeur continue

comprise dans l'intervalle  $[0; 1]$  et respectant les axiomes de la logique classique.

De telles techniques ont été utilisées avec succès essentiellement pour l'interprétation des gaz dissous [69, 90, 188], mais aussi dans un contexte plus large pour plusieurs types de variables d'entrées [5, 169].

### 4.6.3 Réseaux de neurones

Les réseaux de neurones artificiels constituent une alternative aux systèmes experts pour le traitement des alarmes. En 1993 déjà, l'étude bibliographique [106] faisait état de nombreuses applications concernant la détection de défauts dans le domaine de l'énergie électrique. Leur principal avantage par rapport aux systèmes experts est de pouvoir modéliser des phénomènes complexes de manière implicite lorsque les relations intervenant entre les différents paramètres d'un processus sont mal connues. Les réseaux supervisés (§ 4.4.5 et sec. 5.1) sont conçus de sorte à réaliser une fonction logique correspondant aux paramètres d'entrée (caractéristiques techniques, valeurs des capteurs de mesure, etc.). Le traitement des résultats de l'analyse des gaz décrit dans [190] illustre particulièrement bien l'utilisation de cette approche dans le cas du diagnostic de transformateur : les entrées sont constituées des taux de gaz mesurés et les sorties correspondent aux variables logiques des différents défauts à identifier ; une valeur de 1 sur l'une des sorties indique que le diagnostic associé doit être pris en compte.

Comme cela a été mentionné à la section 4.1.2, les réseaux non supervisés (les cartes auto-organisées en particulier) présentent un potentiel d'application particulièrement intéressant pour la surveillance de processus. Nous verrons en détail aux chapitres suivants comment mettre en œuvre cette technique dans le contexte de la surveillance continue des transformateurs.

# Chapitre 5

## Cartes auto-organisées

### 5.1 Introduction

Les cartes auto-organisées font partie de la famille des modèles mathématiques connus sous le nom de modèles connexionistes ou réseaux de neurones artificiels. Comme on le verra plus loin pour le cas particulier des cartes auto-organisées (sec. 5.3), ils s'inspirent largement des connaissances acquises en neurobiologie.

Le principe de base est la construction de modèles complexes à partir de cellules appelées neurones, exécutant des opérations mathématiques élémentaires. Traditionnellement, celles-ci sont des fonctions linéaires, sigmoïdes ou gaussiennes. Elles fournissent pour chaque neurone une sortie dépendant des entrées et des facteurs de pondération des connexions les reliant. Ces dernières valeurs, aussi appelées poids synaptiques, prennent en compte le niveau d'interaction inter-neurone au sein de l'architecture du réseau qui peut être directe ou récursive. La figure 5.1 regroupe les représentations des architectures les plus classiques [59, 62, 131]. Le perceptron multicouches (a) et le réseau à fonctions radiales de base (b) sont sans aucun doute les modèles les plus répandus. Pour autant que le nombre de neurones soit suffisant, ils permettent d'approcher n'importe quelle fonction continue définie dans un espace multidimensionnel avec une précision choisie [50, 66]. Le réseau de Hopfield (c) avec son architecture récursive est un cas un peu particulier. Pour un ensemble de poids initiaux donnés, il converge itérativement

vers un état stable. Ses propriétés sont exploitées essentiellement dans le domaine des mémoires associatives.

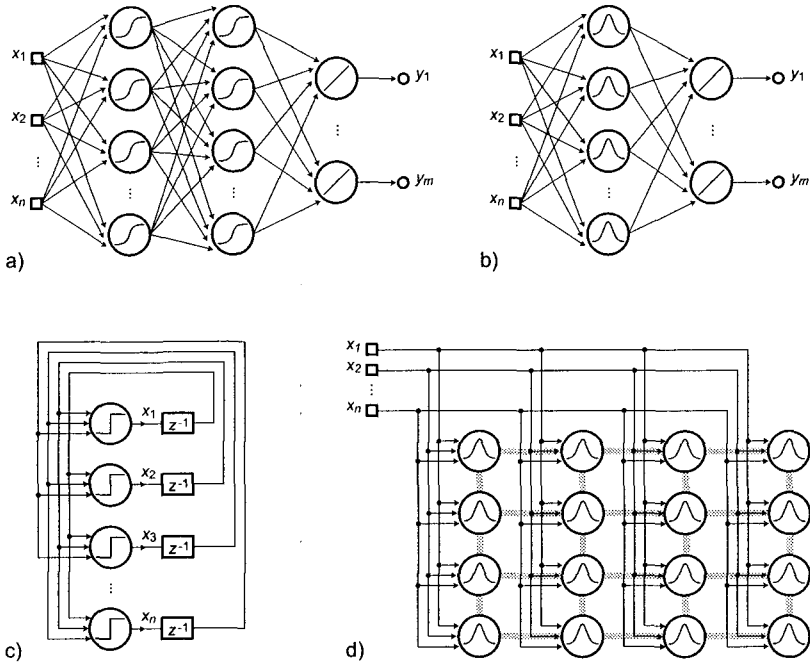


FIG. 5.1 – Architectures typiques de réseaux de neurones : (a) perceptron multicouche à deux couches cachées ; (b) réseau à fonctions radiales de base ; (c) réseau de Hopfield ; (d) carte auto-organisée de  $4 \times 4$  neurones.

Les poids des connexions sont déterminés lors d'un processus appelé apprentissage. Durant celui-ci, il s'agit de calculer itérativement les paramètres qui permettent au réseau de neurones de modéliser la fonction ou les états représentés par les vecteurs issus des ensembles d'entrée et de sortie. Les modèles énumérés jusqu'ici sont caractérisés par le mode de détermination dit *supervisé* de leurs paramètres. Ceux-ci sont adaptés de telle sorte que pour un vecteur d'entrée donné, le réseau fournisse la sortie la plus proche possible du vecteur de sortie associé.



Le réseau de la figure 5.1d représente quant à lui une carte auto-organisée. Il appartient à une catégorie supplémentaire dont l'apprentissage est qualifié de *non supervisé*. Comme cela est expliqué plus loin, les traits grisés mettent en évidence l'organisation topologique des neurones entre eux. La notion d'ensemble d'entrée et de sortie est remplacée par celle d'un ensemble unique d'états. Au cours de l'apprentissage, les vecteurs de cet ensemble sont classés automatiquement en groupes de vecteurs similaires.

## 5.2 Quantification vectorielle

Soit une distribution  $V$  définie par un ensemble de vecteurs de dimension  $n$ . Le nombre de vecteurs pouvant être important, il est souvent souhaitable de réduire la quantité de données. Ceci peut être fait en appliquant des techniques de réduction de dimension telle que l'analyse en composantes principales vue à la section 4.2. Une autre approche est la réduction du nombre de données en ne sélectionnant que quelques échantillons pertinents représentant tout l'ensemble. Comme les relations entre les différentes composantes sont prises en compte, on parle de quantification vectorielle.

Cette technique est essentiellement utilisée en traitement du signal (compression) et en analyse de données (coalescence ou "clustering" en anglais, § 4.2.5). Formellement, la quantification vectorielle consiste à représenter une distribution  $V$  dans un espace  $\mathbb{R}^n$  à l'aide d'un sous-ensemble fini de  $N$  éléments  $\mathbf{w}_i$  dans ce même espace. Ce sous-espace est appelé dictionnaire ou "codebook" en anglais. Il est représenté par la matrice  $\mathbf{W}$  de dimension  $N \times n$  dont les lignes sont les vecteurs-codes  $\mathbf{w}_i = [w_{i1}, \dots, w_{in}]$ :

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_N \end{bmatrix}. \quad (5.1)$$

La réduction du nombre de données est obtenue en associant un vecteur-code à chaque élément  $\mathbf{v}$  de la distribution. La qualité de la représentation subit en revanche une distorsion moyenne  $D(\mathbf{W})$ :

$$D(\mathbf{W}) = \mathbb{E}[d(\mathbf{v}, \mathbf{w}_{i^*})] \quad (5.2)$$

où  $d(\mathbf{v}, \mathbf{w}_{i^*})$  est la distorsion entre  $\mathbf{v}$  et le prototype qui le représente (par exemple la distance euclidienne).

Pour un nombre  $N$  de vecteurs-codes et une distribution fixée, le quantifieur vectoriel est optimal si les deux conditions suivantes, appelées conditions de Lloyd-Max sont remplies [92, 97] :

- le vecteur  $\mathbf{w}_{i^*}$  qui représente un  $\mathbf{v}$  doit être le vecteur-code  $\mathbf{w}_i$  le plus proche au sens la mesure  $d(\mathbf{v}, \mathbf{w}_i)$  choisie (condition du plus proche voisin) ;
- chaque vecteur-code  $\mathbf{w}_i$  doit être le centre de gravité de tous les vecteurs  $\mathbf{v}$  qu'il représente, c'est-à-dire  $\mathbf{w}_i = \mathbf{E}[\mathbf{v} | \mathbf{w}_{i^*}(\mathbf{v}) = \mathbf{w}_i]$  (condition des centroïdes).

Parmi les nombreuses versions d'algorithmes permettant de générer un dictionnaire répondant à ces critères, le plus connu est certainement celui des k-moyens (ou "k-means" en anglais) aussi appelé nuées dynamiques. Pour plus de détails, l'ouvrage [46] passe en revue et décrit toutes ces méthodes précisément.

Répondant également aux critères énoncés, les cartes auto-organisées forment une classe particulière de quantifieurs vectoriels. Grâce à la structure de leurs neurones, elles permettent de conserver la notion de voisinage entre les éléments de l'espace des données et ceux de l'espace de représentation (neurones).

### 5.3 Fondements biologiques

Le néocortex cérébral que l'on ne trouve que chez les mammifères est subdivisé en aires cytoarchitectoniques. Chacune de ces zones joue un rôle différent et peut être classée suivant l'une de ces trois catégories principales : les aires sensorielles primaires et secondaires qui reçoivent et traitent les informations ; les aires motrices associées au contrôle du mouvement volontaire ; les aires associatives impliquées dans des fonctions cognitives complexes telles que le langage ou la reconnaissance des formes.

Ces zones possèdent entre elles des liens plus ou moins forts suivant le degré d'interdépendance des informations qu'elles représentent. Parallèlement il existe, à l'intérieur de ces mêmes zones, une organisation locale hautement structurée. La représentation des grandeurs physiques

est effectuée, des cellules réceptrices aux aires sensorielles concernées, en conservant leur topologie [73, 77].

Le traitement de la vision en est un bon exemple. Les zones corticales concernées sont constituées de neurones connexes dont l'arrangement correspond, avec une certaine déformation, à l'ensemble du champ visuel. On parle de rétinitopie (fig. 5.2). Les cellules du cortex visuel primaire sont excitées sélectivement suivant la position de l'entrée par rapport au champ visuel global selon, en quelque sorte, une projection du champ sur la zone corticale. De la même manière, on parle de somatotopie et de tonotopie pour la représentation respective de la surface du corps et des sons.

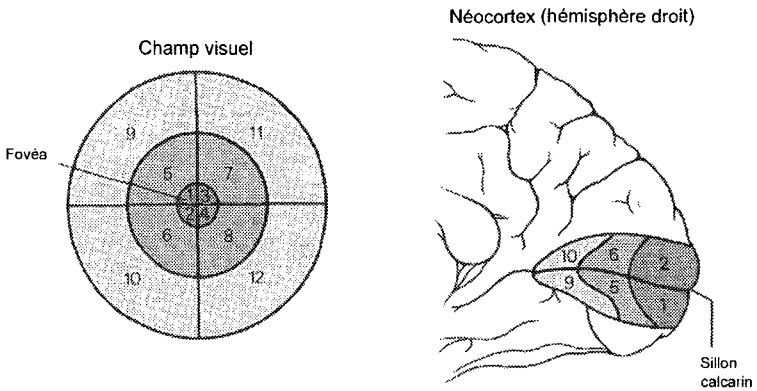


FIG. 5.2 – Représentation du champ visuel dans le néocortex [77].

Les fondements de cette organisation sont sans aucun doute d'ordre génétique. Actuellement, les mécanismes ne sont pas encore parfaitement connus. Il est néanmoins établi qu'au stade embryonnaire le développement des tissus récepteurs est lié à celui des cellules du cerveau. Par conséquent, les connexions tiennent compte de l'organisation topographique générale. En revanche, étant donné le grand nombre de connexions entrant en jeu (et par conséquent le nombre de poids synaptiques), il est évident que l'organisation initiale ne peut être que générale; la dernière phase du développement peut être expliquée par un

processus d'auto-organisation réalisé à l'aide des informations perçues par le système sensoriel.

Le premier modèle mathématique décrivant la formation des cartes rétinitopiques (et projections topiques en général) par auto-organisation est celui de von der Malsburg et Willshaw [186]. Il est constitué de deux couches formées par deux grilles bidimensionnelles de tailles différentes (fig. 5.3). L'une représente les capteurs et l'autre le réseau de neurones. Les deux couches sont entièrement interconnectées. En outre, les neurones possèdent entre eux des connexions latérales qui sont excitatrices pour un proche voisinage et inhibitrices à longue distance. Sans entrer ici dans les détails, ce modèle est capable d'apprendre, d'une façon non supervisée, à projeter un stimulus localisé à un endroit du plan de capteur vers un endroit correspondant dans la carte des neurones.

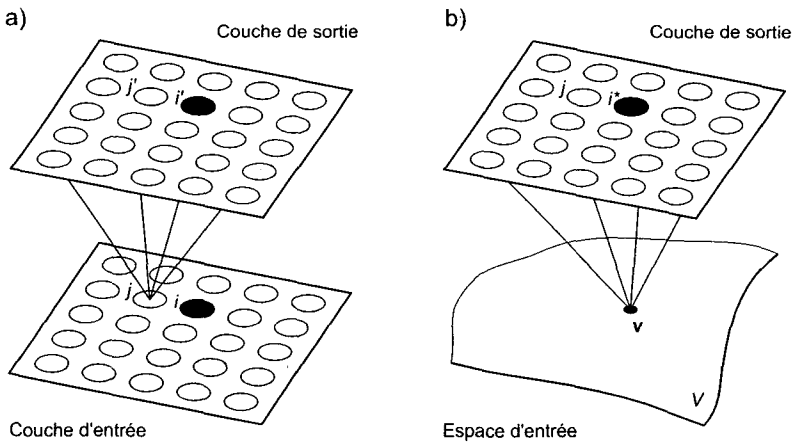


FIG. 5.3 – Architecture des cartes auto-organisées : (a) modélisation de la rétinitopie selon Von der Malsburg ; (b) simplification apportée par le modèle de Kohonen.

## 5.4 Réseau de Kohonen

### 5.4.1 Hypothèses de base

La modélisation connue sous le nom de cartes auto-organisées de Kohonen (en anglais KSOM) [84] apporte les simplifications suivantes par rapport à la méthode de von der Malsburg et Willshaw :

- le motif d'activité des capteurs de l'espace d'entrée est remplacé par un vecteur à composantes continues (fig. 5.3) ;
- les connexions inter-neurones sont supprimées au profit d'un apprentissage dont les poids sont adaptés en fonction de leur voisinage par rapport au neurone le plus actif (neurone gagnant) ;

Ce modèle, numériquement plus efficace, s'éloigne de la représentation rigoureuse des phénomènes biologiques. Il en conserve néanmoins l'idée de base, à savoir l'organisation topologique des données d'un ensemble donné sur une carte. À ce titre, cette méthode peut être vue comme une technique de quantification vectorielle sous contrainte d'une structure de voisinage entre les unités. On verra qu'elle a la propriété d'effectuer la projection non linéaire d'un espace multidimensionnel sur un espace généralement bidimensionnel tout en conservant une trace de la topologie des éléments de l'ensemble projeté.

### 5.4.2 Architecture

Soit  $\mathbf{v}$  un vecteur à  $n$  composantes de l'espace  $V$ . La carte auto-organisée de Kohonen qui lui correspond forme un réseau de  $N$  neurones à  $n$  entrées, toutes connectées aux neurones du réseau par  $n \cdot N$  connexions. Autrement dit, chaque neurone est un vecteur  $\mathbf{w}_i$  de dimension  $n$  dont les poids se fixent lors de la phase d'apprentissage. Ceux-ci sont placés dans un espace qui comporte généralement une ou deux dimensions. Ils possèdent un nombre défini de voisins auxquels ils sont connectés par une règle précise de voisinage qui dépend de la topologie du réseau. Habituellement, cette structure est hexagonale ou carrée (fig. 5.4).

Une mesure typique du voisinage entre deux neurones est la distance euclidienne entre leurs vecteurs coordonnées définis dans l'espace des neurones. Dans le cas bidimensionnel, cet espace forme une grille. Si les

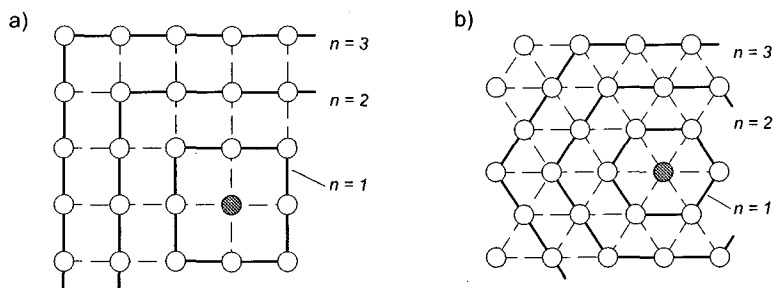


FIG. 5.4 – Définition de l'ordre de voisinage selon la distance euclidienne : (a) structure carrée; (b) structure hexagonale.

coordonnées de deux neurones sur cette grille sont les vecteurs  $\mathbf{x}_1$  et  $\mathbf{x}_2$ , la valeur de la distance euclidienne  $d$  est

$$d = \|\mathbf{x}_1 - \mathbf{x}_2\| = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2}. \quad (5.3)$$

### 5.4.3 Apprentissage

Formellement, l'algorithme d'apprentissage des cartes est un processus itératif dont chaque étape peut être décomposée en deux phases. Il y a tout d'abord une phase de compétition permettant de sélectionner le meilleur neurone. Ensuite, durant la phase coopérative, les poids de tous les neurones sont adaptés en tenant compte d'un facteur de pondération relatif au degré de voisinage par rapport au neurone gagnant.

#### Phase de compétition

Dans le modèle de von der Malsburg et Willshaw, la classification d'un point du plan des capteurs par le réseau de neurones est définie par le niveau d'interactivation de chacun des neurones. Selon la deuxième propriété énoncée (§ 5.4.1) les connexions inter-neuronaux n'existent pas dans le cas du réseau de Kohonen. L'unité la plus active doit donc être déterminée différemment. Une technique proche du fonctionnement des systèmes biologiques consiste à déterminer le neurone gagnant  $i^*$  par le

calcul des produits scalaires :

$$\mathbf{w}_{i^*} \cdot \mathbf{v} = \max_i (\mathbf{w}_i \cdot \mathbf{v}), \quad \text{aussi écrit} \quad i^* = \arg \max_i (\mathbf{w}_i \cdot \mathbf{v}). \quad (5.4)$$

Cet algorithme nécessite une normalisation des vecteurs-codes  $\mathbf{w}_i$ . Lors de l'adaptation du réseau, ces opérations engendrent des calculs relativement lourds. C'est pourquoi, la détermination à l'aide de la distance euclidienne lui est préférée :

$$i^* = \arg \min_i \|\mathbf{w}_i - \mathbf{v}\|. \quad (5.5)$$

La normalisation des vecteurs-codes n'étant pas une condition nécessaire à l'organisation dans le cas de ce second algorithme [84], la formulation mathématique du problème s'en trouve allégée. Néanmoins, dans le cas particulier où les vecteurs-codes sont normés, on peut montrer que les expressions (5.4) et (5.5) sont équivalentes.

### Phase de coopération

La règle de sélection d'après la distance euclidienne minimale est utilisée conjointement à une fonction de voisinage pour l'adaptation des poids des neurones :

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \eta(t)\Lambda_{i,i^*}(t)(\mathbf{v} - \mathbf{w}_i(t)) \quad (5.6)$$

où  $\eta(t)$  est une fonction décroissante dépendant du temps et  $\Lambda_{i,i^*}(t)$  une fonction dépendant à la fois du voisinage du neurone  $i$  par rapport au neurone gagnant  $i^*$  ainsi que du temps  $t$ .

Ces fonctions doivent être choisies de sorte à ce que les poids des vecteurs-codes convergent après un nombre suffisamment grand d'itérations. La fonction  $\eta(t)$ , exclusivement dépendante du temps est typiquement

$$\eta(t) = \eta_0 \left( \frac{\eta_f}{\eta_0} \right)^{\frac{t}{t_{max}}} \quad (5.7)$$

avec  $t_{max}$  le nombre maximal d'itération,  $\eta_0$  et  $\eta_f$  la valeur initiale respectivement finale de  $\eta(t)$ .

La fonction  $\Lambda_{i,i^*}(t)$  peut être définie discrètement selon l'ordre de voisinage des neurones [84] ou de manière continue à l'aide d'une fonction

de voisinage gaussien :

$$\Lambda_{i,i^*}(t) = \exp\left(-\frac{(\mathbf{x}_i - \mathbf{x}_{i^*})^2}{2\sigma_\Lambda(t)^2}\right) \quad (5.8)$$

avec

$$\sigma_\Lambda(t) = \sigma_{\Lambda 0} \exp\left(-2\sigma_{\Lambda 0} \frac{t}{t_{max}}\right) \quad (5.9)$$

où le vecteur  $\mathbf{x}_i$  contient les coordonnées du neurone  $i$  sur la carte et  $\sigma_{\Lambda 0}$  est la valeur initiale de  $\sigma_\Lambda(t)$ .

### Algorithme

Le tableau 5.1 contient l'algorithme d'apprentissage des cartes auto-organisées de Kohonen selon les définitions des paragraphes précédents. À l'issue du processus itératif, pour autant que  $t_{max}$  ait été défini suffisamment grand par rapport au nombre d'éléments d'apprentissage, le réseau est organisé. Comme pour les techniques classiques de quantification vectorielle, chaque vecteur-code  $\mathbf{w}_i$  représente un certain nombre de vecteurs d'entrée. L'espace des entrées est quantifié en classes et les neurones sont les éléments représentatifs de ces classes. Compte tenu de la conservation topologique, on dit alors que le réseau réalise la projection des vecteurs d'entrée à  $n$  dimensions dans l'espace défini par la carte.

TAB. 5.1 – *Algorithme d'apprentissage KSOM (incrémental).*

---

```

initialisation aléatoire des poids  $w_{ij}$ 
pour  $t \leftarrow 1, \dots, t_{max}$ 
  sélection d'un vecteur  $\mathbf{v}$  de l'ensemble  $V$  au hasard
  détermination du neurone gagnant  $i^*$  d'après la distance euclidienne :
     $i^* \leftarrow \underset{i}{\operatorname{argmin}} \|\mathbf{w}_i - \mathbf{v}\|$ 
  adaptation des poids selon l'expression :
     $\mathbf{w}_i \leftarrow \mathbf{w}_i + \eta(t) \Lambda_{i,i^*}(t) (\mathbf{v} - \mathbf{w}_i)$  pour  $i = 1, \dots, N$ 
fin

```

---

Dans le cas particulier où l'espace vectoriel  $V$  est bidimensionnel, la représentation du réseau de Kohonen dans l'espace des poids permet de



visualiser la dynamique de l'apprentissage<sup>1</sup> (fig. 5.5). Après une initialisation totalement aléatoire des poids, alors que les paramètres d'adaptation sont relativement grands, l'organisation générale du réseau est obtenue en tenant compte d'un voisinage relativement grand. Ensuite, au fur et à mesure que les paramètres décroissent, le réseau tend à faire diminuer l'erreur totale  $\sum_{i=1}^N \sum_{k=1}^M \|\mathbf{w}_i - \mathbf{v}_k\|$ . Les modifications sont plus fines et surtout, le périmètre d'influence beaucoup plus local (pour  $t \rightarrow t_{max}$  il se limite au neurone gagnant).

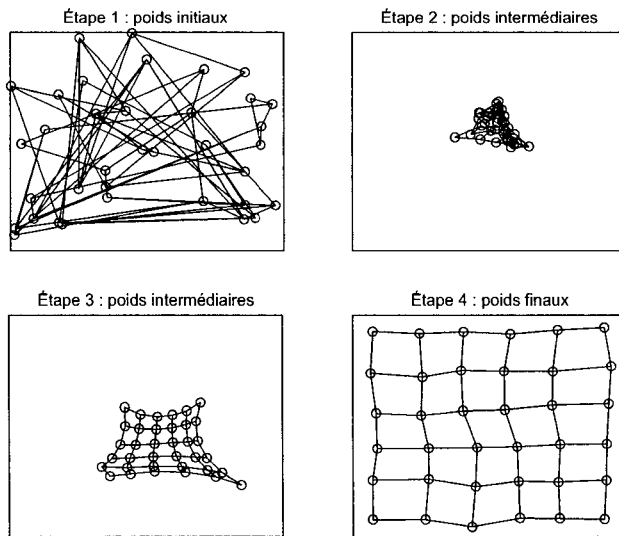


FIG. 5.5 – *Dynamique de l'algorithme d'apprentissage illustrée pour une distribution d'entrée représentant une surface carrée.*

1. Lorsque l'entrée et le réseau de neurones sont tous deux des espaces bidimensionnels, chaque vecteur-code contient la position du neurone dans l'espace d'entrée. Il est alors possible de projeter le réseau de Kohonen dans l'espace des poids et de visualiser la topologie en traçant des lignes entre les neurones adjacents de la grille.

## 5.5 Considérations statistiques

### 5.5.1 Préservation de la topologie

Formellement, pour le cas particulier unidimensionnel, un réseau est dit topologiquement organisé si pour trois neurones  $i$ ,  $j$  et  $k$ , leurs poids respectifs  $w_i$ ,  $w_j$  et  $w_k$  répondent aux conditions suivantes :

$$|w_i - w_j| < |w_i - w_k| \quad \text{et} \quad |i - j| < |i - k| \quad (5.10)$$

pour  $i, j, k = 1, \dots, m$ .

Plus généralement, la topologie est conservée si l'ordre de voisinage dans l'espace des neurones correspond à celui des vecteurs qu'ils représentent dans l'espace des entrées. Lorsque les dimensions de ces deux espaces ne sont pas identiques, cette condition ne peut être remplie totalement. Le degré de conservation topologique peut alors être évalué numériquement (cf. par exemple [8] et [180]).

La convergence de l'algorithme d'apprentissage de Kohonen vers un état topologiquement organisé n'a pu être démontrée rigoureusement que pour le cas de cartes monodimensionnelles [22, 37, 38, 129]. Les principaux résultats théoriques sont réunis dans [23]. Dans le cas général, pour une distribution discrète, on sait toutefois que la convergence impose une vitesse limite de décroissance du pas d'adaptation  $\eta(t)$  et du facteur de voisinage  $\Lambda(t)$ . Si ceux-ci sont modifiés trop rapidement, des défauts d'ordonnement peuvent apparaître dans le réseau (fig. 5.6). Les mesures de conservation topologique citées plus haut peuvent alors servir à déterminer la vitesse de diminution des paramètres (§ 5.7.5).

### 5.5.2 Normalisation

L'analyse de l'organisation a mis en évidence le fait que la localisation des neurones d'une carte ne se fait pas uniformément dans l'espace décrit par toutes les variables. L'organisation se fait préférentiellement par rapport aux variables dont la variance (en valeur absolue) est maximale [84, 129].

Cette propriété peut être utilisée avantageusement pour déterminer l'orientation préférentielle d'un espace. À ce titre, la carte de Kohonen

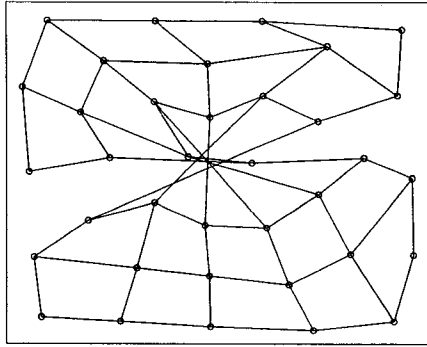


FIG. 5.6 – Défaut d'organisation topologique dû à une décroissance trop rapide des paramètres d'apprentissage (distribution 2D).

peut être vue comme une extension non-linéaire de l'analyse en composantes principales [128]. En revanche, lorsque les variables entrant en considération représentent des grandeurs qui ne sont par nature pas comparables, il convient de s'affranchir de cet effet de classification. Comme cela sera vu plus loin, la surveillance continue de transformateur entre dans cette catégorie.

Il existe de nombreuses solutions consistant à normaliser indépendamment chaque variable composant l'ensemble de départ : normalisation dans les intervalles  $[0; 1]$ ,  $[-1; 1]$ , etc. Afin de prendre en compte également la variance, la solution retenue consiste à normaliser l'ensemble de départ de sorte à ce que pour chaque variable, la valeur moyenne soit nulle et la variance unitaire. On parle de données centrées-réduites. Formellement, chaque élément constituant le vecteur  $\mathbf{v}$  sera normalisé selon

$$\tilde{v}_k = \frac{v_k - \mu_{v_k}}{\sigma_{v_k}}, \quad k = 1, \dots, n \quad (5.11)$$

où  $\mu_{v_k}$  et  $\sigma_{v_k}$  sont respectivement la valeur moyenne et l'écart type de la variable  $k$  sur l'ensemble d'entrée  $V$ .

### 5.5.3 Distribution de la probabilité des poids

On attribue généralement aux cartes de Kohonen la propriété de modéliser sous forme discrète la densité de probabilité générale de l'ensemble d'entrée [84]. Une analyse plus précise de la structure met en évidence une distorsion de cette densité entre l'espace d'entrée et celui des poids. En effet, lorsque la fonction de voisinage est telle que seul le neurone gagnant est adapté, l'algorithme d'apprentissage converge vers un arrangement qui tend à minimiser en moyenne la fonction utilisée pour quantifier l'espace  $V$  en  $N$  régions disjointes. Il en résulte que la densité des poids ne sera pas une fonction linéaire de celle de l'ensemble de départ.

Soit  $p(\mathbf{v})$  la fonction de densité de probabilité multidimensionnelle d'un vecteur  $\mathbf{v}$ . Cette fonction intégrée sur l'intégralité de l'ensemble de départ  $V$  doit être par définition :

$$\int_{-\infty}^{\infty} p(\mathbf{v}) d\mathbf{v} = 1. \quad (5.12)$$

Le facteur d'amplification  $m(\mathbf{v})$  est défini comme étant le nombre de neurones représentant un petit volume  $d\mathbf{v}$  de l'espace  $V$ . Intégré sur tout l'espace, le facteur d'amplification doit contenir le nombre total de neurones  $N$  :

$$\int_{-\infty}^{\infty} m(\mathbf{v}) d\mathbf{v} = N. \quad (5.13)$$

La représentation exacte de la densité de probabilité de l'espace d'entrée est obtenue lorsque  $m(\mathbf{v}) \propto p(\mathbf{v})$ . Cette propriété implique que les régions de l'espace d'entrée les plus souvent activées occupent proportionnellement plus de surface sur la carte topographique que les autres zones. En d'autres termes, pour ces régions, le nombre de neurones alloués sera plus grand.

En toute généralité, il n'est pas possible d'établir une relation directe entre le facteur d'amplification  $m(\mathbf{v})$  et la fonction de densité de probabilité  $p(\mathbf{v})$  pour une carte de Kohonen bidimensionnelle. Par contre, une telle relation a pu être établie pour une carte monodimensionnelle [127] :

$$m(\mathbf{v}) \propto p^\alpha(\mathbf{v}) \quad (5.14)$$

avec

$$\alpha = \frac{2}{3} - \frac{1}{3(n^2 + (n+1)^2)} \quad (5.15)$$

où  $n$  est le nombre de neurones voisins de part et d'autre du neurone gagnant. Ce résultat explique qualitativement l'observation faite expérimentalement qui montre que les cartes de Kohonen tendent à sur-représenter les régions dont la densité d'entrée est faible et sous-représenter les zones dont cette densité est grande.

L'algorithme d'apprentissage conduit parfois à un second défaut : pour certaines distributions, il arrive que des neurones ne soient jamais actifs pour l'ensemble d'apprentissage. On parle alors d'unités "mortes". Certains auteurs ont proposé des modifications de l'algorithme d'apprentissage afin de remédier à ce problème et avoir des neurones dont la probabilité d'activation est uniforme. L'idée est d'ajouter un facteur de pénalité aux neurones gagnant fréquemment. L'activation des autres cellules est ainsi favorisée.

## 5.6 Autres algorithmes

De manière générale, la modélisation correcte d'un processus nécessite que la condition d'équiprobabilité d'activation des neurones soit remplie. Comme cela apparaîtra plus loin, cette condition est essentielle pour l'application qui fait l'objet de ce travail. Un grand nombre d'alternatives à l'algorithme de Kohonen ont été proposées pour la formation de cartes topologiques. L'ouvrage [73] recense les principales d'entre elles. On en retrouve ci-dessous quelques catégories (liste non exhaustive).

- a. Les méthodes heuristiques basées sur les règles de *surveillance de l'activation* ajoutent un facteur de pénalité appelé conscience [30] (cf. [73] pour plus de détails sur les différentes versions). L'activation des neurones gagnant le moins fréquemment est ainsi favorisée et l'apprentissage est guidé vers une organisation équiprobable des neurones. Cependant, aucune garantie formelle ne peut être faite quant à la convergence de l'organisation vers cet état [74].
- b. Les méthodes de *définition dynamique de la topologie* permettent de représenter fidèlement l'ensemble de départ. Ceci se fait malheureu-

sement au prix de la suppression d'une structure bien définie (cf. [84] pour un aperçu des techniques).

- c. La *modification de la structure du réseau* permet d'améliorer la classification. Parmi les solutions développées, on peut citer le réseau VQP ("Vector Quantization and Projection") [28] et les cartes topographiques génératives (GTM) [15]. Du point de vue structurel, respectivement statistique, ces deux techniques améliorent la modélisation de l'espace d'entrée. Cependant, la notion de quantificateurs répartis sur une grille uniforme a disparu.

## 5.7 Réseau kMER

### 5.7.1 Maximisation de l'entropie

Les méthodes énoncées jusqu'à présent sont fondées sur la minimisation d'un critère de distorsion. Une approche plus globale est basée sur l'optimisation directe d'un critère d'information théorique. Dans ce cas, la carte s'organise de telle sorte à contenir le maximum d'information disponible. La densité des poids est alors proportionnelle à la densité des entrées. En d'autres termes, les neurones ont tous la même probabilité d'être activés.

Pour mémoire, quelques définitions sont données ci-après. Soit  $Z$  une variable aléatoire discrète<sup>2</sup>

$$Z = \{z_k | k = 0, \pm 1, \dots, \pm K\} \quad (5.16)$$

où  $z_k$  est une valeur numérique et  $(2K + 1)$  est le nombre de total de grandeurs discrètes. Pour compléter le modèle, la probabilité  $p_k = P(Z = z_k)$  est également définie avec la condition

$$0 \leq p_k \leq 1, \quad \text{et} \quad \sum_{k=-K}^K p_k = 1. \quad (5.17)$$

Suivant la probabilité d'apparition de l'événement  $Z = z_k$ , sa réalisation sera plus ou moins surprenante. Si  $p_k = 1$  alors  $p_i = 0, \forall i \neq k$

---

2. Selon les conventions admises en théorie des probabilités, une variable aléatoire est désignée par une lettre majuscule et la valeur que prend cette variable par une lettre minuscule.

et l'événement  $Z = z_k$  est sûr d'arriver. On dit alors que l'information transmise par l'événement  $Z = z_k$  est nulle. Formellement la quantité d'information transmise est définie par

$$I(z_k) = \log \left( \frac{1}{p_k} \right) = -\log p_k \quad (5.18)$$

où les logarithmes sont généralement naturels ou exprimés en base 2. Les unités sont alors respectivement les nats ou les bits.

L'entropie  $H(Z)$  de la variable aléatoire  $Z$  est l'espérance mathématique de la somme d'information livrée par message :

$$H(Z) = E[I(z_k)] = - \sum_{k=-K}^K p_k \log p_k. \quad (5.19)$$

Tout le raisonnement aboutissant à cette équation peut être appliqué de la même façon au cas continu par le passage à la limite  $K \rightarrow \infty$ . L'équation (5.19) montre que la quantité d'information est maximale lorsque  $p_k = \text{constante} = \frac{1}{K}$ . En associant les neurones aux valeurs de la variable aléatoire  $Z$ , l'utilisation du critère de maximisation de l'entropie revient à imposer une probabilité d'activation constante.

## 5.7.2 Quantificateur équiprobable

Pour construire un quantificateur équiprobable selon une règle de calcul itérative, Van Hulle a proposé une méthode basée sur l'adaptation des limites des classes [74]. Soit un espace scalaire uniforme et  $N$  quantificateurs distincts. Ceux-ci sont positionnés sur un axe réel par ordre croissant tel que  $w_1 < w_2 \dots < w_N$ . Comme le montre la figure 5.7, les poids  $w_i$  ne représentent plus dans ce cas la moyenne de chacun des intervalles  $L_i$  mais leurs limites. La probabilité que chacun de ces intervalles soient actifs est directement proportionnelle à leur longueur. Si  $P(L_i)$  est trop grand, la longueur de l'intervalle doit être diminuée en déplaçant ses délimitations et inversement si  $P(L_i)$  est trop petit.

Formellement, pour une organisation initiale quelconque des neurones formant un espace  $W$ , l'adaptation des poids peut être exprimée par [73]

$$w_i(t+1) = w_i(t) + \eta \sum_{\{j|L_j \in S_i\}} \mathbf{1}_j(v) \operatorname{sgn}(v - w_i(t)), \quad \forall i \in W \quad (5.20)$$

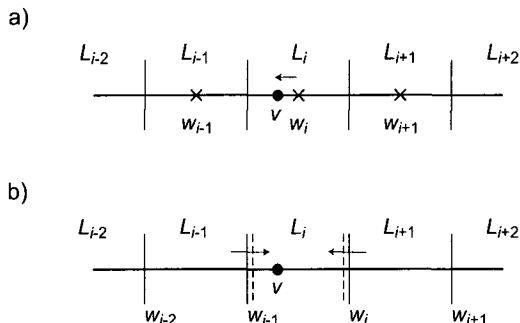


FIG. 5.7 – Adaptation des poids de deux types de quantificateurs scalaires : (a) apprentissage compétitif non-supervisé classique ; (b) apprentissage par adaptation des limites.

avec  $v \in V$ ,  $\eta$  un facteur d'apprentissage constant,  $S_i$  l'ensemble des intervalles ayant le point de séparation  $i$  en commun,  $\mathbb{1}_i(v)$  la fonction d'appartenance

$$\mathbb{1}_i(v) = \begin{cases} 1 & \text{si } v \in L_i \\ 0 & \text{si } v \notin L_i \end{cases} \quad (5.21)$$

et  $\text{sgn}(x)$  la fonction signe définie par

$$\text{sgn}(x) = \begin{cases} +1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -1 & \text{si } x < 0. \end{cases} \quad (5.22)$$

Cette méthode peut être étendue sans aucun problème à la quantification de vecteurs. En outre, lorsque le nombre  $N$  de neurones est grand, il est démontré qu'après convergence, la densité des poids  $p(\mathbf{w}_i)$  est proportionnelle à la densité des entrées  $p(\mathbf{v})$  [70]. Le principal défaut de cet algorithme est relatif à la dimension de l'espace des neurones : il doit nécessairement être égal à celui de l'ensemble d'entrée.



### 5.7.3 Structure du réseau kMER

L'approche dite kMER<sup>3</sup> est un développement de la méthode vue ci-dessus. Elle permet d'obtenir des cartes satisfaisant les deux conditions suivantes [71] :

- organisation topologique par un ensemble de représentation dont la dimension est indépendante de l'ensemble modélisé ;
- probabilité d'activation uniforme des neurones.

La structure du réseau est obtenue en reprenant l'architecture des cartes de Kohonen et en ajoutant un seuil  $\tau_i$  à chaque vecteur-code  $\mathbf{w}_i$ . Celui-ci définit une zone d'activation  $S_i$  centrée en  $\mathbf{w}_i$ . Par soucis de simplicité, comme le montre la figure 5.8, la fonction  $g$  définissant la zone d'activation est gaussienne. L'intersection de cette fonction avec le seuil  $\tau_i$  définit ainsi la région d'activation  $S_i$  (circulaire dans le cas bidimensionnel ou hypersphérique en toute généralité) dont  $\sigma_i$  est le rayon dans l'ensemble des variables  $V$ .

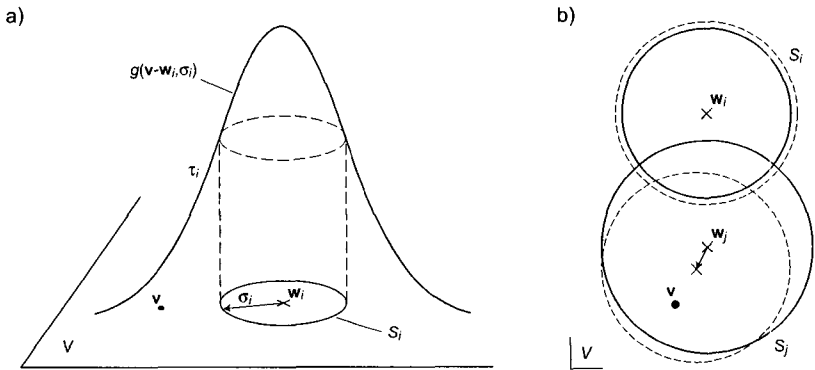


FIG. 5.8 – Réseau kMER : (a) définition du seuil et de la zone d'activation ; (b) adaptation des zones d'activation [73].

<sup>3</sup>. kMER est l'abréviation de l'expression anglaise "kernel-based Maximum Entropy learning Rule" (règle d'apprentissage selon la maximisation de l'entropie de fonctions nodales).

### 5.7.4 Apprentissage

Durant la phase d'apprentissage, contrairement aux cartes KSOM, plusieurs unités peuvent être désignées gagnantes : si le vecteur courant  $\mathbf{v}$  tombe dans la zone d'activation  $S_i$  alors le neurone correspondant est activé. La fonction d'appartenance (5.21) devient pour le cas considéré

$$\mathbb{1}_i(\mathbf{v}) = \begin{cases} 1 & \text{si } \mathbf{v} \in S_i \\ 0 & \text{si } \mathbf{v} \notin S_i. \end{cases} \quad (5.23)$$

Pour tenir compte de l'activation simultanée de plusieurs neurones, une fonction d'appartenance floue  $\Xi_i$  est introduite [189] :

$$\Xi_i(\mathbf{v}) = \frac{\mathbb{1}_i(\mathbf{v})}{\sum_{k=1}^N \mathbb{1}_k(\mathbf{v})} \quad (5.24)$$

définie telle que  $0 \leq \Xi_i(\mathbf{v}) \leq 1$  et  $\sum_i \Xi_i(\mathbf{v}) = 1$  pour  $i = 1, \dots, N$ .

### Adaptation

Les centres  $\mathbf{w}_i$  mais aussi les rayons  $\sigma_i$  sont adaptés itérativement. L'idée est de les ajuster de telle sorte qu'après convergence la probabilité d'un neurone  $i$  d'être actif est

$$P(\mathbb{1}_i(\mathbf{v}) \neq 0) = \frac{\rho}{N} \quad (5.25)$$

avec  $\rho$  une constante gérant le degré d'interaction des neurones (§ 5.7.5). Afin de répartir les neurones sur l'intégralité de l'ensemble d'entrée, les centres des zones d'activation sont modifiés proportionnellement à  $\Xi_i$  dans la direction de  $\mathbf{v}$  :

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \eta \sum_{j=1}^N \Lambda_{i,j}(t) \Xi_j(\mathbf{v}) \text{sgn}(\mathbf{v} - \mathbf{w}_i(t)) \quad (5.26)$$

où la fonction signe est appliquée indépendamment pour chaque composante. Les rayons sont quant à eux adaptés suivant

$$\sigma_i(t+1) = \sigma_i(t) + \eta \left( \frac{\rho_r}{N} (1 - \mathbb{1}_i(\mathbf{v})) - \mathbb{1}_i(\mathbf{v}) \right) \quad (5.27)$$

avec  $\rho_r \triangleq \frac{\rho N}{N - \rho}$ .

### Algorithme optimisé

Soit un ensemble d'apprentissage  $A = \{\mathbf{v}_\mu\}$  de  $M$  vecteurs défini dans l'ensemble d'entrée  $V$ . L'apprentissage par paquets ("batch" en anglais) fournit une alternative à la technique incrémentale ; il déclenche l'adaptation des paramètres uniquement après que la carte ait été interrogée pour tous les vecteurs constituant l'ensemble  $A$  [72]. Pour être appliquée, la technique par paquets nécessite que tous ces vecteurs soient disponibles au début de l'apprentissage. Le cas échéant, elle présente l'avantage de pouvoir être mise en œuvre avec une fonction de contrôle de convergence (§ 5.7.5). Comme on le verra au chapitre suivant, cette condition est parfaitement remplie dans le contexte de surveillance continue considéré. Désormais, il ne sera donc plus fait mention que de l'apprentissage par paquets. Les expressions (5.26) et (5.27) deviennent alors

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \eta \sum_{\mu=1}^M \sum_{j=1}^N \Lambda_{i,j}(t) \Xi_j(\mathbf{v}_\mu) \text{sgn}(\mathbf{v}_\mu - \mathbf{w}_i(t)) \quad (5.28)$$

et

$$\sigma_i(t+1) = \sigma_i(t) + \eta \sum_{\mu=1}^M \left( \frac{\rho_r}{N} (1 - \mathbb{1}_i(\mathbf{v}_\mu)) - \mathbb{1}_i(\mathbf{v}_\mu) \right). \quad (5.29)$$

Pratiquement, la formule (5.28) implique à chaque pas une adaptation de tous les poids des neurones par rapport à tous les neurones activés. L'algorithme optimisé tel qu'il apparaît dans le tableau 5.2 apporte deux simplifications. Tout d'abord, les calculs peuvent être accélérés de manière significative en adaptant les poids des neurones non activés  $i_n$  uniquement par rapport au neurone actif  $i^*$  correspondant à la distance euclidienne  $\|\mathbf{v}_\mu - \mathbf{w}_i\|$  minimale :

$$\mathbf{w}_{i_n}(t+1) = \mathbf{w}_{i_n}(t) + \eta \sum_{\mu=1}^M \Lambda_{i_n, i^*}(t) \Xi_{i^*}(\mathbf{v}_\mu) \text{sgn}(\mathbf{v}_\mu - \mathbf{w}_{i_n}(t)). \quad (5.30)$$

Les neurones actifs  $i_a$  sont quant à eux adaptés avec une contribution de voisinage maximale :

$$\mathbf{w}_{i_a}(t+1) = \mathbf{w}_{i_a}(t) + \eta \sum_{\mu=1}^M \Xi_{i_a}(\mathbf{v}_\mu) \text{sgn}(\mathbf{v}_\mu - \mathbf{w}_{i_a}(t)). \quad (5.31)$$

Lorsque la fonction de voisinage est importante, la dynamique de modification des poids est légèrement différente de celle définie pour l'algorithme original. Cependant, à la fin de l'apprentissage, lorsque l'influence de cette fonction aura disparu, les effets seront les mêmes pour les cas complet et simplifié.

Une deuxième mesure permettant d'accélérer l'apprentissage est l'adaptation systématique à chaque pas des poids par rapport au neurone  $i^*$ , indépendamment du fait qu'il soit activé ou non. En revanche, l'adaptation des rayons  $\sigma_i$  dépend toujours de l'état d'activité. De ce fait, lorsque les poids sont initialisés hors de l'espace défini par les éléments de l'ensemble d'apprentissage, ou que cette distribution est couverte incomplètement par les vecteurs-codes, l'adaptation de ces derniers dans la bonne direction est plus rapide.

### 5.7.5 Choix des paramètres

Pour que l'algorithme d'apprentissage converge, il convient de déterminer correctement les différents paramètres. Le premier d'entre eux est le taux d'apprentissage  $\eta$ . Contrairement à ce qui a été vu pour le réseau de Kohonen, celui-ci est constant dans le cas des cartes kMER. Van Hulle a montré que sa valeur doit être inférieure d'un ordre de grandeur par rapport au rayon de la plus petite zone d'activation [73].

Le deuxième paramètre à régler est le facteur d'échelle  $\rho$ . Lorsque  $\rho = 1$ , les zones d'activation sont mutuellement exclusives. Si  $\rho > 1$ , un certain recouvrement des zones d'activation, et par conséquent un certain lissage, est imposé (la fonction d'approximation est plus grossière). À l'inverse pour  $\rho < 1$ , on obtiendra un surplus de signaux d'entrée qui n'activeront pas les neurones [71]. Par conséquent, pour autant que le nombre d'éléments de l'ensemble d'apprentissage soit suffisamment grand, il est souhaitable d'imposer une valeur unitaire à ce facteur. La limite peut être fixée selon la règle expérimentale suivante [73] :

$$\rho \approx \max \left( 1, \frac{N30}{M} \right). \quad (5.32)$$

Le dernier paramètre est la valeur du facteur de voisinage  $\sigma_\Lambda$  intervenant dans le calcul de  $\Lambda_{i,j}(t)$ . La dynamique de cette fonction est donnée par la même expression que pour les cartes de Kohonen (5.8).

TAB. 5.2 – *Algorithme d'apprentissage des cartes kMER ("batch").*


---

```

initialisation aléatoire des poids  $w_{ij}$  et des seuils  $\sigma_i$ 
pour  $t \leftarrow 1, \dots, t_{max}$ 
  initialisation  $\Delta w_{ij} \leftarrow 0$  et  $\Delta \sigma_i \leftarrow 0$  pour  $i = 1, \dots, N$  et  $j = 1, \dots, n$ 
  pour  $\mu \leftarrow 1, \dots, M$ 
    détermination du neurone gagnant selon  $i^* \leftarrow \operatorname{argmin}_i \|\mathbf{v}_\mu - \mathbf{w}_i\|$ 
    pour  $i \leftarrow 1, \dots, N$ 
      initialisation des paramètres de contrôle  $cpt \leftarrow 0$  et  $CM_i \leftarrow 0$ 
      si  $\|\mathbf{v}_\mu - \mathbf{w}_i\| < \sigma_i$ 
        activation de l'index  $CM_i \leftarrow 1$ 
        incrémentation du compteur  $cpt \leftarrow cpt + 1$ 
      fin
    fin
    si  $cpt = 0$  (aucun neurone actif)
      compteur  $cpt \leftarrow 1$ 
      activation de l'index  $CM_{i^*} \leftarrow 1$ 
       $\Delta \sigma_{i^*} \leftarrow \Delta \sigma_{i^*} + \frac{\rho}{N} + 1$ 
    fin
    pour  $i \leftarrow 1, \dots, N$ 
      si  $CM_i = 0$  (neurones non-activés)
         $\Delta \mathbf{w}_i \leftarrow \Delta \mathbf{w}_i + \Lambda_{i,i^*}(t) \operatorname{sgn}(\mathbf{v}_\mu - \mathbf{w}_i)$ 
         $\Delta \sigma_i \leftarrow \Delta \sigma_i + \frac{\rho}{N}$ 
      sinon (neurones activés)
         $\Delta \mathbf{w}_i \leftarrow \Delta \mathbf{w}_i + \frac{\operatorname{sgn}(\mathbf{v}_\mu - \mathbf{w}_i)}{cpt}$ 
         $\Delta \sigma_i \leftarrow \Delta \sigma_i - 1$ 
      fin
    fin
  fin
  adaptation des poids selon  $\mathbf{w}_i \leftarrow \mathbf{w}_i + \eta \Delta \mathbf{w}_i$  pour  $i = 1, \dots, N$ 
  adaptation des seuils selon  $\sigma_i \leftarrow \sigma_i + \eta \Delta \sigma_i$  pour  $i = 1, \dots, N$ 
fin

```

---

La dépendance temporelle du facteur de voisinage doit par contre être reconsidérée. Comme dans le cas de l'algorithme de Kohonen, lorsque ce facteur est modifié trop rapidement, des défauts topologiques peuvent apparaître (fig. 5.6). Afin de régler automatiquement son adaptation, Van Hulle [73] propose une méthode basée sur l'évaluation de la topolo-

gie du réseau lors de l'apprentissage. Partant de la constatation que le réseau a plus de chance d'être déployé correctement si le nombre de neurones activés par une entrée est constant pour tout l'ensemble d'apprentissage, il introduit une mesure appelée la variabilité de recouvrement (ou *OV* pour "Overlap Variability").

Dans le cas de l'algorithme par paquets, pour l'itération  $t$ , le nombre de neurones actifs  $\mathcal{N}$  peut être défini par

$$\mathcal{N}_\mu(t) = \sum_{i=1}^N \mathbb{1}_i(\mathbf{v}_\mu). \quad (5.33)$$

Les expressions pour la valeur moyenne et la déviation standard sur l'ensemble d'apprentissage  $A$  sont alors

$$\text{moy}_{\mathcal{N}}(t) = \frac{1}{M} \sum_{\mu=1}^M \mathcal{N}_\mu(t) \quad (5.34)$$

et

$$\text{std}_{\mathcal{N}}(t) = \left( \frac{1}{M-1} \sum_{\mu} (\mathcal{N}_\mu(t) - \text{moy}_{\mathcal{N}}(t))^2 \right)^{\frac{1}{2}}. \quad (5.35)$$

La variabilité de recouvrement est finalement définie comme le quotient de ces deux valeurs :

$$OV = \frac{\text{std}_{\mathcal{N}}(t)}{\text{moy}_{\mathcal{N}}(t)}. \quad (5.36)$$

L'expression de décroissance temporelle du facteur de voisinage  $\sigma_\Lambda$  définie dans le cas des cartes de Kohonen (5.9) est alors modifiée :

$$\sigma_\Lambda(t) = \sigma_{\Lambda 0} \exp \left( -2\sigma_{\Lambda 0} \frac{t}{t_{max}} \gamma_{OV} \right) \quad (5.37)$$

où  $\sigma_{\Lambda 0}$  est la valeur initiale de  $\sigma_\Lambda(t)$  et  $\gamma_{OV}$  le gain de la fonction de décroissance. La mise en œuvre de cette technique d'adaptation contrôlée du facteur de voisinage est décrite par le tableau 5.3.

### 5.7.6 Densité de probabilité

Le réseau obtenu après convergence de l'apprentissage peut être utilisé pour construire la fonction de densité de probabilité  $p(\mathbf{v})$  des entrées.

TAB. 5.3 – *Algorithme d'adaptation contrôlée du facteur de voisinage.*


---

apprentissage initial avec $\sigma_\Lambda(t) = cte \Rightarrow$ état initial $\mathbf{W}_0$
initialisation $j \leftarrow 1$
apprentissage complet depuis $\mathbf{W}_0$ , $\sigma_\Lambda(t)$ calculé selon (5.37) et $\gamma_{OV_1} = 1$
identification du nombre d'itérations $t_1$ selon $t_1 = \min_i(OV)$
mise en mémoire des variable $OV_1$ et $\sigma_{\Lambda_1}$ associées
<b>répéter</b>
incrémentatation $j \leftarrow j + 1$
modification $t_{maxj} \leftarrow 2t_{j-1}$
adaptation du taux du gain de la fonction de décroissance selon :
$\gamma_{OV_j} = - \frac{\ln \frac{0.9\sigma_{\Lambda_j-1}}{\sigma_{\Lambda_0}}}{2\sigma_{\Lambda_0}}$
apprentissage complet depuis $\mathbf{W}_0$ (avec $t_{maxj}$ et $\gamma_{OV_j}$ )
identification du nombre d'itérations $t_j$ selon $t_j = \min_i(OV_j)$
mise en mémoire des variable $OV_j$ et $\sigma_{\Lambda_j}$ associées
<b>jusqu'à</b> ( $OV_j < OV_{j-1}$ ou $j > max_j$ )

---

Les deux approches généralement mises en œuvre pour les méthodes de quantification vectorielle (fenêtre de Parzen et noyau variable) ainsi que leur application dans les différents cas de figure sont décrites dans [73, 141].

La nature même des cartes kMER permet de considérer une troisième alternative dont l'exploitation des propriétés est particulièrement adaptée au problème de surveillance qui sera traité aux chapitres suivants. La densité de probabilité peut être considérée comme la combinaison linéaire des densités associées à chaque neurone :

$$p(\mathbf{v}) = \sum_{i=1}^N p(\mathbf{v}|i)P(i) \quad (5.38)$$

où  $P(i)$  est la probabilité que le neurone  $i$  soit activé par le vecteur  $\mathbf{v}$  et  $p(\mathbf{v}|i)$  la densité de probabilité conditionnelle du neurone  $i$ . Ces deux grandeurs répondent encore aux conditions suivantes :

$$\sum_{i=1}^N P(i) = 1, \quad 0 \leq P(i) \leq 1 \quad (5.39)$$

et

$$\int_{\mathbf{V}} p(\mathbf{v}|i) d\mathbf{v} = 1. \quad (5.40)$$

Compte tenu des propriétés de l'apprentissage kMER, les probabilités  $P(i)$  valent  $\frac{1}{N}$  quelle que soit la valeur de  $i$ . Quant à la densité  $p(\mathbf{v}|i)$ , elle est donnée par la forme de la fonction  $g$  définissant la zone d'activation (fig. 5.8). S'agissant d'une fonction gaussienne avec une matricé de covariance diagonale dont tous les termes sont identiques (noyau symétrique), la densité de probabilité de chaque composante s'exprime par

$$p(\mathbf{v}|i) = \frac{1}{(2\pi\sigma_i^2)^{\frac{d}{2}}} \exp\left(-\frac{\|\mathbf{v} - \mathbf{w}_i\|^2}{2\sigma_i^2}\right). \quad (5.41)$$

### 5.7.7 Détection de dépassement

La fonction de densité de probabilité du vecteur est une valeur qui s'avère insuffisante pour évaluer la qualité de modélisation locale du réseau de neurones. L'activation du neurone gagnant  $i^*$  pour un vecteur  $\mathbf{v}$  peut être caractérisée qualitativement suivant la valeur de la fonction d'activation  $g$  associée. Celle-ci dépend à la fois de la distance  $\|\mathbf{v} - \mathbf{w}_{i^*}\|$  et du rayon  $\sigma_{i^*}$ . Pour autant que la valeur maximale soit indépendante du rayon, un seuil de qualité tenant compte de la précision locale du modèle est obtenu. Dans ce but, l'équation (5.41) est modifiée de la manière suivante :

$$IC(\mathbf{v}, i^*) = \exp\left(-\frac{\|\mathbf{v} - \mathbf{w}_{i^*}\|^2}{2\sigma_{i^*}^2}\right) \quad (5.42)$$

où  $IC(\mathbf{v}, i^*)$  est l'indicateur de confiance quantifiant la qualité de la représentation du vecteur  $\mathbf{v}$  par le neurone  $i^*$ . Si  $IC(\mathbf{v}, i^*)$  est inférieur à un certain seuil, il faut considérer qu'il n'y a plus adéquation entre l'entrée et l'espace modélisé. En d'autres termes, cela signifie que le vecteur  $\mathbf{v}$  n'appartient pas à l'espace décrit par l'ensemble d'apprentissage.



## Chapitre 6

# Surveillance par les cartes auto-organisées

### 6.1 Principe

#### 6.1.1 Application des cartes auto-organisées

Les propriétés de représentation synthétique des données sur une surface de dimension réduite ont ouvert un vaste champ d'application aux cartes auto-organisées. Celui-ci couvre des domaines tels que l'analyse et le contrôle de procédés ainsi que le traitement de l'image et de l'information. Teuvo Kohonen dresse lui-même dans [85] un inventaire relativement complet des travaux conduits dans ces domaines. De manière plus générale encore, un document établi en 1998 fait part de 3043 publications concernant les cartes auto-organisées [79]. Parmi elles, une majorité est constituée d'applications originales.

Dans le domaine de l'analyse et le contrôle de procédés, bien que chaque application nécessite des adaptations particulières, les cartes sont toujours exploitées selon l'un des axes de la méthodologie générale suivante [1, 85, 142] :

- le processus est caractérisé par un certain nombre de mesures constituant une base de données ;
- une carte auto-organisée est générée sur la base de ces exemples ;
- la structure générale du réseau organisé met en évidence les relations entre les différentes grandeurs et les états du système ;

- lors de son utilisation, l'ensemble des mesures effectuées à un moment donné correspond à un neurone de la carte (neurone gagnant) ;
- la localisation ainsi que l'évolution temporelle du neurone gagnant permet de tirer des conclusions quant à l'état du système (fig. 6.1).

Comme cela sera développé plus loin dans ce chapitre, après l'apprentissage, la carte représente l'ensemble des états de la base de données initiale. Compte tenu de l'organisation topologique, différentes zones regroupant plusieurs neurones peuvent généralement être identifiées. Elles correspondent alors à autant de régimes particuliers. Dès que ces zones sont identifiées et repérées sur la carte, il est possible d'évaluer l'état d'un processus en fonction de la position ou de l'évolution du neurone gagnant (§ 6.6.1).

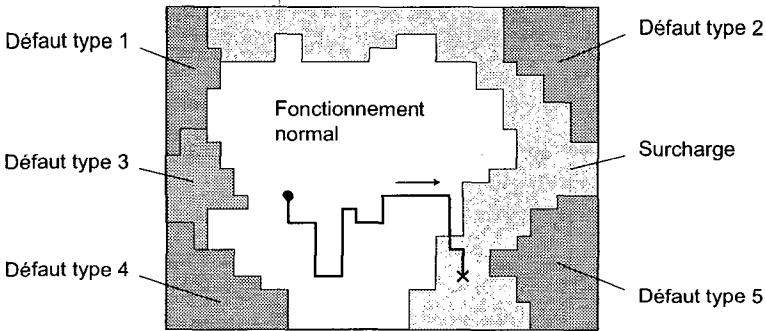


FIG. 6.1 – Principe de l'utilisation des cartes auto-organisées pour la surveillance en général.

Typiquement, cette approche a été appliquée à la surveillance de processus chimiques [171, 176], médicaux [164, 178] et informatiques [63]. Dans le domaine de l'énergie électrique, les problèmes liés à la prévision de la charge ont été les premiers à être traités avec une approche connexionniste [106]. Les chercheurs ont ensuite étendu rapidement l'application des réseaux de neurones artificiels à l'ensemble du domaine. En particulier, les cartes auto-organisées ont été appliquées à la prévision de charge [9, 24, 68, 93], la gestion des alarmes [4], l'analyse de sécurité [107, 108], la stabilité dynamique [102], la stabilité de tension [56]

ainsi que la surveillance des centrales [44, 105, 115] et de l'équipement [78, 111, 140, 165].

Plus spécifiquement, les transformateurs de puissance ou plutôt l'évaluation de leurs conditions de fonctionnement ont également fait l'objet d'applications. La littérature fait état de développements dans le cadre de l'interprétation d'essais de chocs [10, 173] où les cartes sont utilisées pour la reconnaissance de la forme de la fonction de transfert et des conditions qui lui sont associées (§ 3.3.3). Plus récemment, comme cela a déjà été mentionné à la section 3.2.5, les cartes de Kohonen ont été mises en œuvre dans le cadre de l'exploitation des résultats d'analyse des gaz dissous [39, 168]. La complexité du problème associée au grand nombre de données collectées depuis de nombreuses années rend cet outil particulièrement intéressant dans ce contexte.

L'idée d'utiliser les informations fournies par les capteurs installés sur la cuve en continu telle qu'elle est développée plus loin a été émise en 1991 déjà [80]. Cette application met en évidence la puissance de l'approche et a été citée à de nombreuses reprises pour illustrer le potentiel d'utilisation des cartes de Kohonen pour la surveillance de processus [84, 85]. Les travaux décrits dans [18] valident expérimentalement cette approche sur un modèle réduit (transformateur de distribution). En plus des paramètres thermiques et électriques, les vibrations sont enregistrées par plusieurs accéléromètres. Un dispositif ad-hoc simulant un défaut mécanique permet de modifier la réponse vibratoire du système. Après un apprentissage réalisé à partir de données obtenues dans toutes les configurations, la carte auto-organisée permet de différencier les états sains des états défaillants.

La limitation principale de la méthode vient du fait que seuls les états constituant l'ensemble d'apprentissage sont pris en compte par le modèle. Or, dans le cas de nombreux processus industriels, seules les données relatives au fonctionnement normal sont disponibles. De plus, s'il n'existe aucun modèle (numérique ou expérimental) pour simuler les états défaillants, seules les configurations de fonctionnement normales seront représentées par la carte. Cette limitation s'applique malheureusement aux transformateurs de puissance : chaque exemplaire est quasiment unique ; on ne dispose d'aucun modèle numérique complet ; les possibilités d'expérimentation sur des modèles réduits sont extrê-

vement limitées. Pour cette raison essentiellement, aucune application concrète n'a vu le jour dans le domaine. En développant une idée exposée dans [57], on verra au paragraphe 6.6.2 comment contourner ce problème en exploitant l'erreur de quantification de la carte (distance mesure-gagnant).

### 6.1.2 Cadre d'application

S'agissant de l'état d'un processus entier ou d'une machine évalué au moyen de plusieurs paramètres, les cartes auto-organisées viennent s'intégrer dans un dispositif de surveillance et de diagnostic qui peut être relativement complexe. Elles en constituent néanmoins l'un des maillons essentiels. Le schéma de la figure 6.2 représente le principe général de surveillance utilisant ces cartes. Celui-ci se compose de trois parties distinctes : le système d'acquisition automatique, la gestion de la base de données et le traitement de l'information.

Les capteurs permettant de mesurer automatiquement les paramètres de fonctionnement des transformateurs sont aujourd'hui largement répandus (chap. 3). Intégrés au système SCADA<sup>1</sup> des exploitants ou gérés de façon autonome par l'intermédiaire d'un système d'acquisition indépendant, ils délivrent des valeurs accessibles à distance de n'importe où et en tout temps.

Le problème de la gestion de la base de données est assez délicat : il s'agit de rassembler toutes les mesures disponibles sous une forme qui pourra être utilisée lors de la mise au point et de l'utilisation des modèles. Les conditions d'exploitation réelles sont caractérisées par la présence d'événements aléatoires tels que des mesures erronées, des données manquantes, des erreurs de transmission, etc. Un filtrage des données est donc nécessaire. De plus, comme cela apparaîtra plus loin, les contraintes du système de surveillance peuvent nécessiter un prétraitement numérique. De ce point de vue, la solution qui sera retenue est relativement élémentaire.

La contribution essentielle de la méthodologie développée réside dans le traitement de l'information : le schéma de la figure 6.2 fait la distinction entre le traitement des valeurs instantanées et celui des tendances.

---

1. SCADA est l'abréviation anglaise de "Supervisory Control And Data Acquisition".

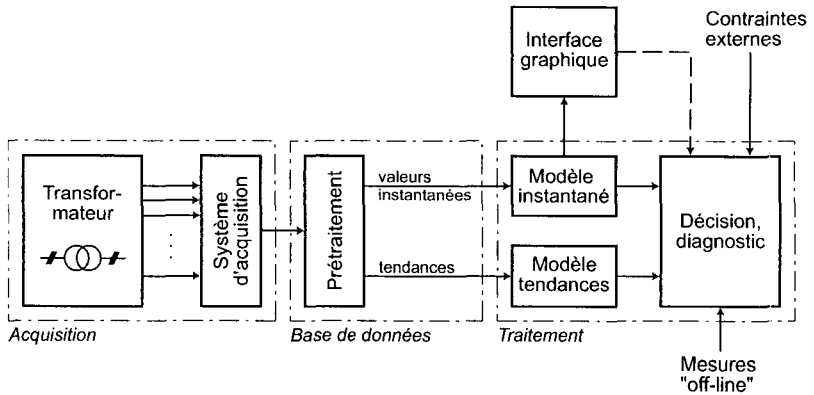


FIG. 6.2 – Principe du dispositif de surveillance continue intégré pour les transformateurs de puissance.

Bien qu'il existe quelques applications particulières où ces deux fonctions sont rassemblées, les travaux réalisés à ce jour dans le domaine des transformateurs montrent qu'il est judicieux de découpler les changements rapides des phénomènes se développant lentement (sec. 4.6).

Comme cela est expliqué à la section 6.3.2, les contraintes liées à la nature même des cartes nécessitent un réapprentissage périodique. Ceci implique que le modèle neuronal ne peut pas prendre en compte les changements dont la constante de temps est supérieure à l'intervalle de temps défini entre deux apprentissages successifs. Dès lors, les cartes interviennent dans l'analyse des grandeurs instantanées et des variations à court terme. La surveillance de l'évolution à long terme est bien découplée et son traitement ne sera pas abordé ici. Le schéma de principe de la figure 6.3 représente sous la même forme que celle établie au paragraphe 4.4.2 l'approche développée avec les cartes auto-organisées. Le réseau de neurones peut être assimilé à une structure réalisant à la fois la modélisation du processus ainsi que le calcul de la mesure de déviation des grandeurs mesurées par rapport à celles modélisées.

Pour être complet, les résultats des deux modules qui viennent d'être décrits doivent être interprétés par un système de diagnostic. Idéalement, il doit non seulement être capable de prendre en compte leurs

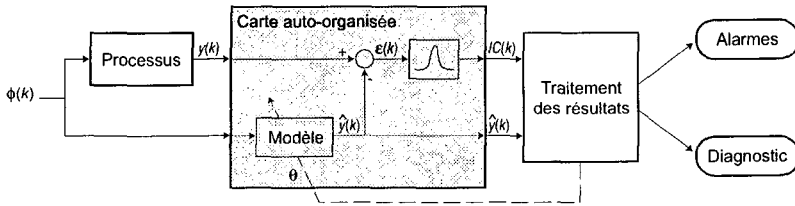


FIG. 6.3 – Principe de l'analyse des grandeurs instantanées à l'aide des cartes auto-organisées.

sorties mais aussi les pondérer en fonction de paramètres d'exploitation supplémentaires. Ceux-ci proviennent essentiellement des résultats fournis par les mesures périodiques classiques et les contraintes externes liées aux paramètres économiques.

## 6.2 Mise en forme des mesures

### 6.2.1 Base de données

Une fois récoltées, les mesures subissent un prétraitement et sont rassemblées dans une base de données unique. Le prétraitement peut être décomposé en trois étapes successives :

- identification et modification des données manquantes ou aberrantes ;
- calcul de variables découlant directement des mesures ;
- suppression éventuelle des tendances et composantes périodiques.

L'identification des données manifestement fausses et des valeurs manquantes est nécessaire au bon fonctionnement du système de surveillance situé en aval. Le but de cette opération est simplement de trouver une convention permettant au système de ne pas prendre en compte la donnée concernée ; elle sera soit reconstruite à l'aide des observations précédentes, soit ignorée.

Les valeurs prises en compte pour caractériser l'état du processus ou de la machine à surveiller ne sont pas strictement limitées aux grandeurs directement délivrées par les différents capteurs. Par exemple, suivant les

corrélations et les constantes de temps caractérisant le système étudié, il peut être judicieux de faire intervenir un retard, de ne tenir compte que des accroissements voire même de ne considérer qu'une valeur issue de la combinaison des sorties de plusieurs capteurs. D'autre part, il est également envisageable de caractériser la dynamique de l'évolution d'une variable par la prise en compte de coefficients tels que ceux issus d'une transformée de Fourier ou d'un calcul d'ondelettes.

Enfin, comme cela apparaîtra plus loin, l'efficacité des cartes auto-organisées est influencée par la stationnarité des grandeurs prises en compte. En d'autres termes, une grandeur entachée d'une tendance ne pourra pas être modélisée correctement. Pour autant que cela soit possible, une suppression de cette tendance avant le traitement par le réseau de neurones est souhaitable.

### 6.2.2 Vecteur d'entrée

Comme cela a été vu au paragraphe 4.2.1, après le prétraitement, la base de données peut être représentée par un tableau ou une matrice  $\mathbf{V}$  de  $m$  échantillons comprenant chacun les valeurs fournies par les  $n$  valeurs numériques prises en compte pour la modélisation. Chaque échantillon correspond alors à une ligne qui peut être identifiée par une référence temporelle (la date et l'heure par exemple).

Le vecteur d'entrée  $\mathbf{v}$  du réseau de neurones n'est alors rien d'autre que l'expression générale de l'une des lignes de la base de données :

$$\mathbf{v}(k) = [v_1(k), v_2(k), \dots, v_n(k)] \quad (6.1)$$

où  $v_i$  est la  $i^{\text{ème}}$  valeur numérique considérée et  $k$  est l'instant d'échantillonnage.

### 6.2.3 Normalisation

La nature des grandeurs rassemblées dans la base de données est par définition très variée. Aussi, l'amplitude moyenne et la variation relative des différentes variables peuvent différer dans une large mesure. Conformément aux propriétés exposées au paragraphe 5.5.2, l'organisation des cartes sera principalement influencée par les grandeurs possédant la plus grande variance.

Dans le cas de la surveillance de processus (et celui des transformateurs en particulier), il est essentiel de prendre en compte tous les signaux d'entrée avec la même pondération. A priori, toutes les grandeurs doivent être considérées comme étant indépendantes. C'est sur cette base seulement que les différentes corrélations locales pourront être révélées à l'issue de l'apprentissage.

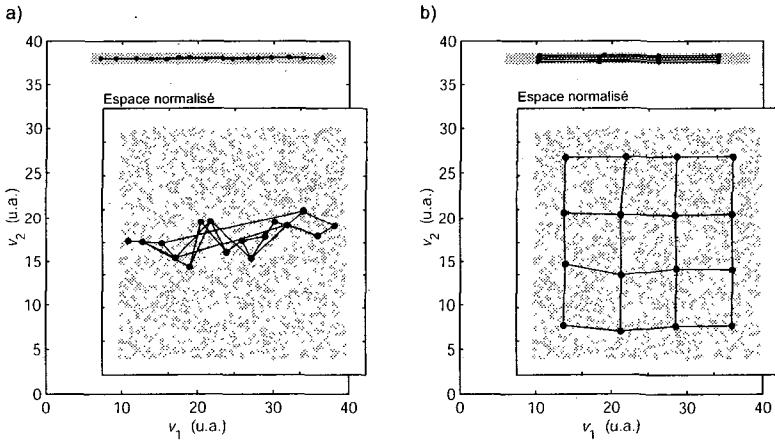


FIG. 6.4 – *Effet de la normalisation sur l'apprentissage du réseau auto-organisé : organisation d'un réseau kMER de 16 neurones pour un espace défini par les mesures de deux détecteurs dont les plages de mesure différent dans un rapport de 1/32 : (a) apprentissage sans normalisation ; (b) apprentissage avec normalisation.*

La figure 6.4 illustre le cas particulier où l'état de fonctionnement d'un transformateur est modélisé d'après deux paramètres dont les variances sont sensiblement différentes (typiquement du même rapport que celui mesuré entre la température et le taux d'humidité relative de l'huile) : les couples  $\{v_1(k); v_2(k)\}$  représentés en gris ont été appris par une carte auto-organisée kMER de 16 neurones ; un apprentissage a été réalisé avec les valeurs non-normalisées et l'autre avec les mêmes valeurs mais normalisées. Les graphiques montrent clairement que dans le premier cas, l'effet de la variation sur  $v_2$  n'est pour ainsi dire pas pris en compte. Par conséquent, afin de couvrir la plage de mesures de



chacun des capteurs, les valeurs de la base de données seront centrées et réduites indépendamment par colonne (5.11) avant d'être introduites dans les cartes auto-organisées.

## 6.3 Apprentissage et adaptation du modèle

### 6.3.1 Validation de l'algorithme d'apprentissage

Les considérations du chapitre 5 ont montré que les cartes kMER diffèrent des cartes de Kohonen dans la mesure où leur algorithme d'apprentissage converge vers une répartition équiprobable des neurones ; leur répartition dans l'espace d'entrée tient mieux compte des variations de la densité des points. L'utilisation de cette propriété d'équiprobabilité ouvre la voie à des méthodes de coalescence basées sur la discrimination en fonction de la densité de probabilité [72]. Pour la surveillance de processus, l'approche proposée exploite les propriétés de cet algorithme pour améliorer les performances de la méthodologie classique adoptée avec les cartes de Kohonen dans ce contexte [85].

Pour un processus caractérisé par des paramètres mesurés sur une période donnée, les propriétés énoncées plus haut impliquent que les états rencontrés fréquemment seront modélisés avec plus de précision pour un réseau kMER que pour un réseau KSOM. Inversement, les états les plus rares seront moins bien modélisés. Contrairement aux cartes de Kohonen, l'algorithme fournit une valeur  $\sigma_s$  caractérisant la zone d'activation de chaque neurone. Elle permet de calculer un indicateur de confiance  $IC$  (5.42) qui, cela sera décrit plus loin, servira à pondérer la réponse du réseau en fonction de sa précision locale. La figure 6.5 illustre l'organisation d'un réseau kMER dans le cas d'un ensemble d'apprentissage défini par deux détecteurs  $v_1$  et  $v_2$  dont la répartition des mesures n'est pas uniforme sur la plage de fonctionnement considérée. Les cercles représentent les zones d'activation des neurones ; elles croissent avec la diminution de la densité de points.

La figure 6.6 représente les résultats de six apprentissages effectués avec les réseaux KSOM et kMER à titre comparatif pour trois espaces d'entrée différents. Comme précédemment, les surfaces définies par les points peuvent être considérées comme les plages de fonctionnement nor-

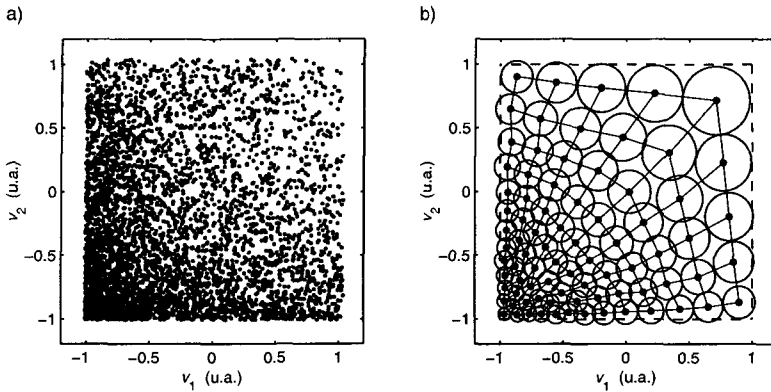


FIG. 6.5 – Mise en évidence de l'organisation équiprobable du réseau *kMER* pour un ensemble d'entrée bidimensionnel dont la répartition des points n'est pas uniforme : espace d'entrée (a) ; réseau organisé avec les zones d'activation (b).

mal d'un processus caractérisé par la mesure des deux paramètres  $v_1$  et  $v_2$ . Alors que pour une surface carrée de densité uniforme, les résultats sont similaires (cas non représenté), les trois domaines choisis illustrent parfaitement les différences entre les deux approches. Les représentations (a) et (b) montrent la façon dont les deux méthodes prennent en compte les variations de densité de probabilité. Il apparaît clairement que l'algorithme *kMER* organise les poids selon une répartition de densité de probabilité plus fidèle à celle de l'espace d'entrée que la méthode de Kohonen. Nous verrons au paragraphe 6.6.2 comment cette propriété permet d'améliorer la sensibilité du système de détection pour les zones les plus utilisées.

Les parties (c) à (f) mettent en évidence le pouvoir d'organisation du réseau *kMER* pour des espaces de géométries particulières. Les cartes KSOM viennent placer des neurones hors des plages définies par les points des ensembles d'apprentissage. Ces éléments sont appelés neurones

"morts" car ils ne représentent aucun état. Dans l'optique de la surveillance de transformateurs, l'ordre de l'espace défini par les paramètres

mesurés est bien supérieur à deux. Cet espace ne peut généralement plus être représenté graphiquement dans un repère défini par les coordonnées des variables. En toute généralité il est fort probable que ses limites ne soient pas régulières. Il peut également être composé de deux espaces disjoints comme dans les cas (e) et (f). La présence de neurones représentant des états inconnus va alors altérer la qualité du système de détection. Il est possible de repérer ces unités mortes et de les rendre inactives. La carte présentera alors des trous qui, en utilisant la méthode kMER peuvent être évités (contraintes liées à l'équiprobabilité).

La bonne organisation du réseau dépend, du choix des paramètres d'apprentissage  $\rho$ ,  $\eta$  et  $\sigma_{\Lambda 0}$  ainsi que de la dynamique de  $\sigma_{\Lambda}$  (§ 5.7.5). La détermination du facteur d'échelle optimal  $\rho$  est faite directement suivant la relation (5.32). Le facteur de voisinage initial  $\sigma_{\Lambda 0}$  doit être fixé en fonction de la taille du réseau de neurones. Typiquement, il doit être fixé à une valeur du même ordre de grandeur que la dimension de la carte. La pratique a montré que pour un réseau carré de  $N$  neurones,  $\sigma_{\Lambda 0} = \sqrt{N}/2$  donne de bons résultats. Sa décroissance est gérée de façon optimale grâce à la procédure décrite dans le tableau 5.3 (p. 97). Quant au taux d'apprentissage  $\eta$ , on a vu qu'il doit être inférieur d'un ordre de grandeur au rayon de la plus petite zone d'activation. A priori, il n'est pas possible de le déterminer exactement avant même d'avoir effectué l'apprentissage. En revanche, sa valeur minimale peut être calculée : elle correspond à la distance minimale entre deux vecteurs de l'ensemble d'entrée. À partir de ce seuil inférieur, la valeur optimale peut être déterminée heuristiquement en observant la convergence de la première phase du processus.

La figure 6.7 montre l'évolution du taux de recouvrement. L'algorithme de contrôle de convergence est appliqué pour l'espace d'entrée défini par les valeurs formant une surface en "U" dont la densité des points est uniforme. Les évolutions de la variabilité de recouvrement  $OV$  et du coefficient de voisinage  $\sigma_{\Lambda}/\sigma_{\Lambda 0}$  sont représentées pour chaque itération considérée sur les graphiques du haut. Au-dessous, l'organisation du réseau est représentée à l'issue de chaque période considérée. Lors de l'itération 0, le coefficient de voisinage ne varie pas au cours de l'apprentissage. Comme le graphique du bas le montre, les neurones se répartissent suivant la topologie du réseau au centre de l'espace d'en-

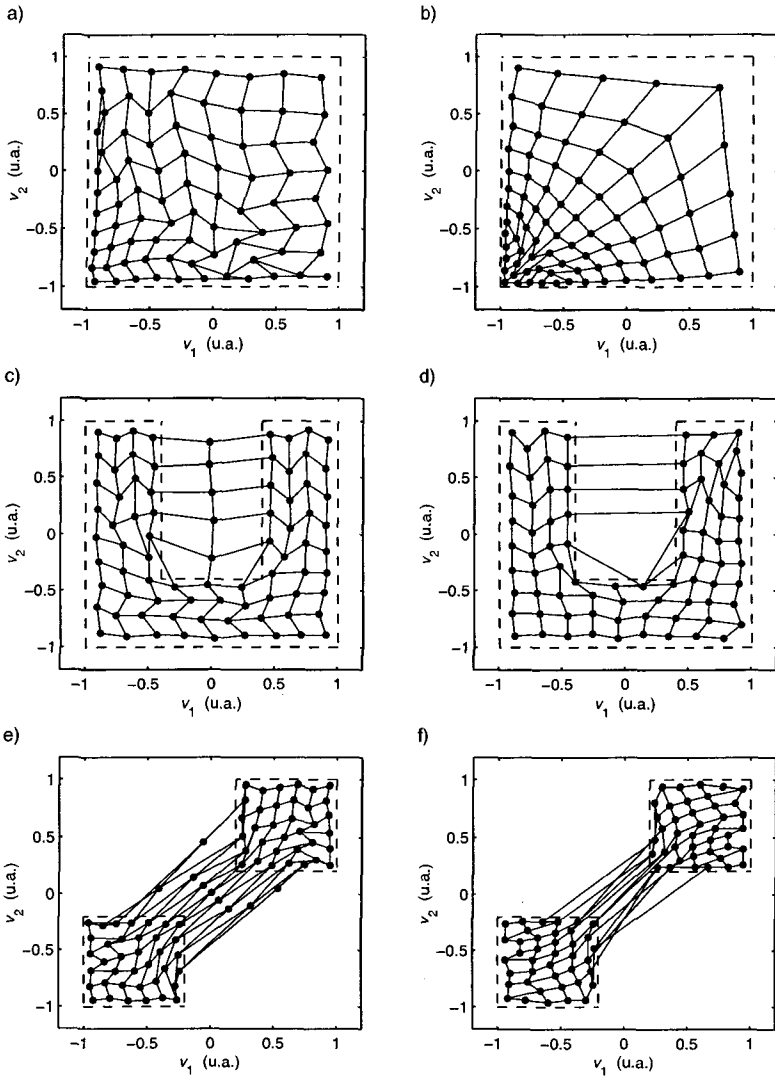


FIG. 6.6 – Comparaison des résultats obtenus pour l'algorithme KSOM (à gauche) et kMER (à droite) : surface carrée avec densité de points variable (a) et (b) ; domaine en "U" avec densité de points uniforme (c) et (d) ; carrés disjoints avec densité de points uniforme (e) et (f).

trée, mais la carte n'est pas déployée. Cette itération permet également d'ajuster la valeur du paramètre  $\eta$ . Si l'indicateur  $OV$  converge vers une valeur, le taux d'apprentissage est suffisamment faible. La vitesse de convergence peut être optimisée en augmentant  $\eta$ . En pratique, l'itération 0 est répétée pour des valeurs croissantes de  $\eta$  à partir de la valeur minimale rappelée plus haut. La limite est atteinte lorsque la valeur de  $OV$  n'arrive plus à se stabiliser. Les itérations suivantes sont caractérisées par la recherche d'un optimum entre le nombre d'époques et la valeur finale de  $\sigma_\Lambda$ . Concrètement, les résultats peuvent être vus sur les représentations des résultats d'apprentissage (fig. 6.7). Au bout du premier cycle, l'organisation générale est satisfaisante mais comporte quelques défauts. À l'issue de la septième itération, l'adaptation optimale de  $\sigma_\Lambda$  a permis de les éliminer.

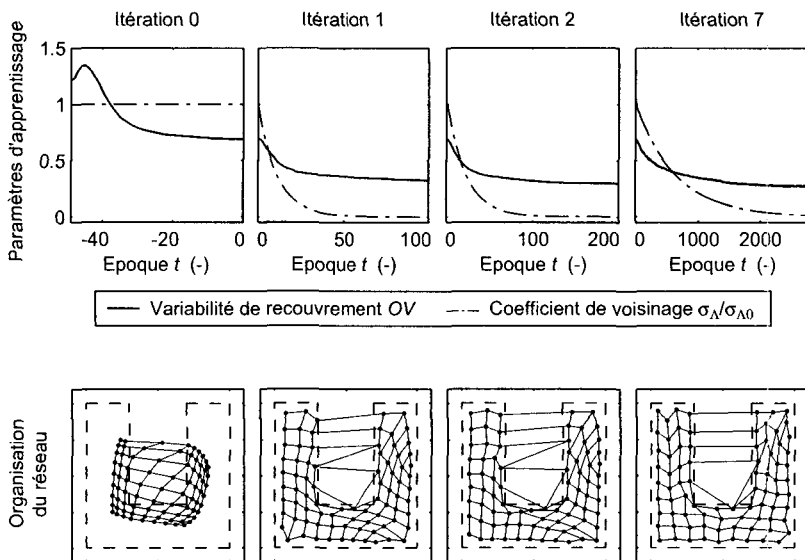


FIG. 6.7 – Illustration du contrôle de l'apprentissage au moyen de la surveillance de la variabilité de recouvrement  $OV$  : apprentissage pour une surface bidimensionnelles en "U" avec répartition uniforme des points.

### 6.3.2 Adaptation périodique

Les cartes auto-organisées, comme l'ensemble des réseaux de neurones, permettent de modéliser un processus sur la base d'observations passées en apprenant ces données et en les généralisant. Si les modèles ont la propriété d'effectuer une interpolation des valeurs dans un espace à  $n$  dimensions, ils ne permettent en aucun cas de faire des extrapolations. Ceci revient à dire que si l'apprentissage a été réalisé pour un certain mode de fonctionnement, le modèle ne sera valable que dans des conditions d'exploitation similaires.

Les études des phénomènes de vieillissement ont mis en évidence le fait que les paramètres mesurés sur les transformateurs évoluent avec le temps. Ceci est particulièrement vrai pour le taux d'humidité qui est un indicateur de l'état du papier (sec. 2.4). De plus, le régime de fonctionnement peut évoluer. Par exemple, dans le cas d'un transformateur, si après des années d'exploitation dans des conditions où la charge est faible en hiver, le courant moyen augmente de manière significative durant la période hivernale, le modèle établi sur la base des années précédentes n'est plus adéquat. Afin de remédier à ce problème, la base de données sera périodiquement adaptée en fonction des mesures les plus récentes. La fréquence de réapprentissage du réseau de neurones dépendra alors de la vitesse à laquelle les changements interviennent.

L'état initial de tout apprentissage est déterminé aléatoirement. Pour deux apprentissages successifs réalisés avec le même ensemble de données, leur convergence n'est pas une garantie quant à l'obtention de la même organisation des poids. La carte étant carrée, sa représentation peut subir une rotation multiple de  $\pi/4$  ou une symétrie axiale. L'affichage graphique des informations contenues dans la carte est un des résultats essentiels pour la surveillance. Afin d'assurer une continuité permettant une interprétation aisée, il faut à tout prix éviter que ce genre de transformation apparaisse entre deux apprentissages successifs.

Partant du principe qu'entre deux adaptations, les états modélisés sont proches, deux solutions sont envisageables pour remédier à ce problème. La première consiste, à l'issue de l'apprentissage, à identifier les éventuelles rotations et/ou symétries intervenues et à les supprimer. La deuxième solution intervient au début du processus d'apprentissage. Les considérations faites au paragraphe 6.3.1 montrent que durant la pre-

mière phase (itération 0), la carte vient s'organiser au centre du domaine de manière régulière (sans croisement pour le cas bidimensionnel). Pour les itérations suivantes la notion d'état initial aléatoire a disparu ; il sont tous obtenus à partir de ce cas de base. Par conséquent, l'idée consiste à reprendre cet état de base pour chacun des apprentissages successifs. L'orientation de la carte est ainsi assurée et l'évolution du système prise en compte. De plus, afin d'accélérer les apprentissages, les paramètres optimaux déterminés pour le premier apprentissage peuvent être repris pour les réapprentissage successifs. Il est donc possible d'éviter de passer systématiquement par l'application de l'algorithme du tableau 5.3 (p. 97). En revanche, en cas d'évolution sensible, il y a quand même lieu de vérifier périodiquement que ces paramètres sont toujours valables.

## 6.4 Visualisation

### 6.4.1 Généralités

Chaque neurone composant une carte auto-organisée est un vecteur. Par conséquent, pour un espace d'entrée de dimension  $n$  supérieure à deux, il est impossible de visualiser exactement toute l'information en une seule représentation. L'article [179] dresse un inventaire des méthodes développées et utilisées à cette fin.

Les paragraphes suivants décrivent en détail les différentes représentations adaptées à la surveillance de processus. Pour être complet, il faudrait rajouter les méthodes dites de projections. La technique la plus courante d'entre elles est connue sous le nom de "Sammon's mapping" en anglais [134]. Son principe est d'effectuer une projection non linéaire de l'ensemble d'entrée sur une surface bidimensionnelle en essayant de conserver les distances relatives entre les vecteurs d'entrée. Cette technique pourrait directement être appliquée avec les données de l'ensemble d'entrée. Dans la pratique, les opérations numériques nécessaires étant relativement lourdes, elle est mise en œuvre beaucoup plus avantageusement avec un ensemble de données réduit. Peu adaptés à une console de surveillance, les résultats donnent non seulement une approximation grossière de la forme de l'ensemble de données, mais aussi une évaluation graphique de l'organisation générale de la carte.

### 6.4.2 Décomposition en couches

L'approche la plus élémentaire consiste à considérer un tel réseau comme la superposition de  $n$  couches représentant respectivement chaque composante des neurones [172]. L'organisation des neurones est la même pour toutes les couches ; elle est fixée par la topologie du réseau de neurones lui-même. Par exemple, dans le cas d'un transformateur où le vecteur d'entrée serait défini par les quatre mesures du courant de charge  $I_{ch}(k)$ , de la température externe  $T_{air}(k)$ , de la température de l'huile  $T_{huile}(k)$  et du taux de gaz  $Gaz(k)$ , le poids de la deuxième couche représenterait la variable  $T_{air}$ .

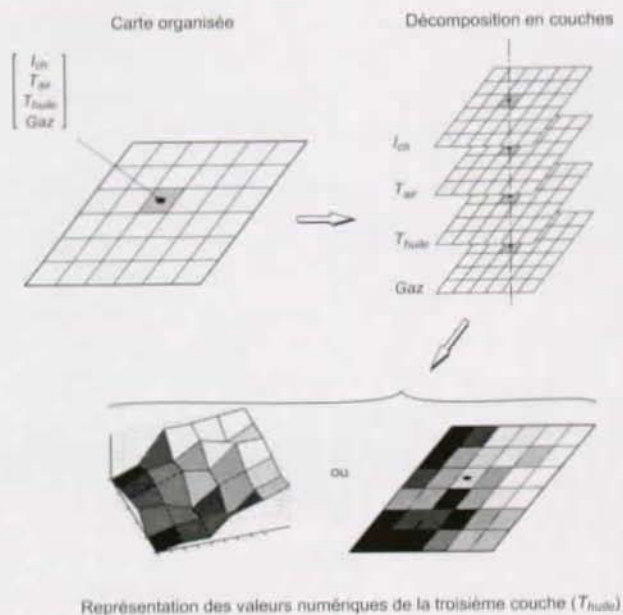


FIG. 6.8 - Principe de la décomposition en couches pour la représentation des cartes auto-organisées.

Chaque couche peut ainsi être représentée séparément sous forme d'un graphique tridimensionnel. Les trois axes sont composés des coordonnées  $x$  et  $y$  des neurones sur la carte et de la valeur de la variable considérée. De la même manière, les informations peuvent être projetées



sur une surface plane où la troisième dimension est remplacée par une coloration différenciée des points (fig. 6.8). L'étude de la combinaison de ces derniers graphiques fournit une excellente indication de l'organisation générale de la carte.

Par exemple, sur la couche représentée à la figure 6.8, les zones claires représentent les valeurs les plus élevées et les zones foncées les valeurs les plus basses. Par conséquent, les températures de l'huile les plus hautes sont modélisées par les neurones du coin supérieur droit de la carte et les températures les plus basses par ceux du coin inférieur gauche.

### 6.4.3 Matrice des distances

La technique la plus appropriée pour l'identification de groupes distincts au sein de l'ensemble d'entrée est l'exploitation des distances euclidiennes inter-neuronales. La méthode consiste à construire une matrice dont les valeurs représentent les distances euclidiennes entre chaque neurones adjacents [177]. Cette méthode, appelée "u-matrix" fournit une matrice qui peut être représentée par une surface colorée selon le même principe que celui décrit au paragraphe précédent. La matrice peut également être simplifiée de façon à ne contenir pour chaque neurone que la valeur moyenne de la distance qui le sépare de tous ses voisins du premier ordre [95].

Soit une échelle de couleur définie de telle sorte que les valeurs élevées correspondent aux zones claires et les valeurs faibles aux zones foncées. Si une démarcation existe pour un ensemble donné, celle-ci se manifestera par une zone de séparation claire. Les zones plus foncées indiquent quant à elles que les régions concernées sont constituées de neurones proches. Au contraire de l'affichage en couches, cette représentation a l'avantage de synthétiser toute l'information relative à l'identification de régions distinctes sur une carte. En revanche, elle ne permet pas de mettre en évidence les corrélations entre les différentes grandeurs. Les graphiques de la figure 6.9 représentent l'affichage des matrices des distances (version complète) pour les deux distributions disjointes vues au paragraphe 6.3.1. Toutes deux ont été apprises avec une carte kMER. La structure du réseau en fonction de la topologie de l'espace d'entrée apparaît clairement. Une ligne blanche coupant la carte (a) en deux permet de distinguer la subdivision de l'espace d'entrée. Pour le cas (b), la

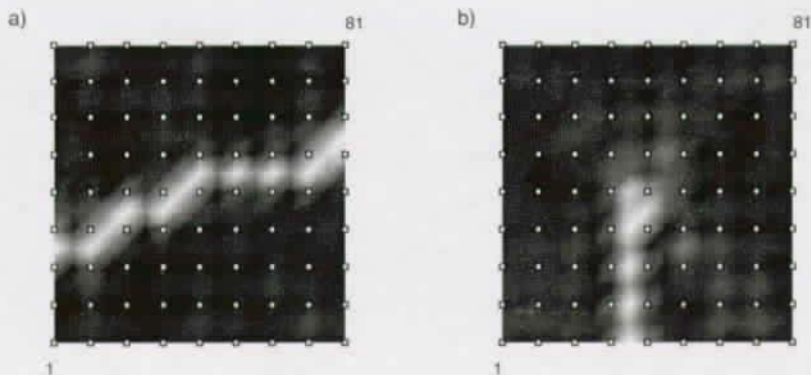


FIG. 6.9 – Représentation par le calcul de la matrice des distances : réseaux modélisant les deux carrés disjoints (a) et la surface en "U" (b); les neurones sont représentés par les carrés blancs et numérotés en ligne de 1 à 81.

distribution particulière de l'ensemble d'apprentissage est mise en évidence par une nette démarcation verticale dans la partie inférieure du réseau.

#### 6.4.4 Histogramme d'activation

Les propriétés des cartes auto-organisées kMER sont telles qu'à l'issue de l'apprentissage, la probabilité d'activation de chaque neurone est égale. En d'autres termes, si le réseau est interrogé pour tous les vecteurs de l'ensemble d'apprentissage, chaque neurone aura été activé le même nombre de fois.

Dans le cas d'un processus stationnaire, pour tout nouvel ensemble de données, la probabilité d'activation de chacun des neurones est effectivement uniforme. En revanche, pour un processus dont le régime de fonctionnement est variable, cette propriété n'est plus respectée. Les analyses décrites plus loin au chapitre 7 montrent que les transformateurs de puissance doivent être assimilés à ce dernier cas : une périodicité annuelle est clairement établie. Elle se manifeste par une organisation dépendant de la saison.

Par conséquent, les neurones représentant des points de fonctionnement typiquement mesurés en hiver n'auront pour ainsi dire aucune probabilité d'être activés en été et vice-versa. Il ne s'agit pas là de la seule structure qui peut être dégagée. L'évolution d'une zone de la carte vers une autre peut dépendre de nombreux autres paramètres. Aussi, afin d'évaluer l'état d'un processus dans un tel cas, il convient de repérer les zones d'activité les plus probables compte tenu de la situation présente.

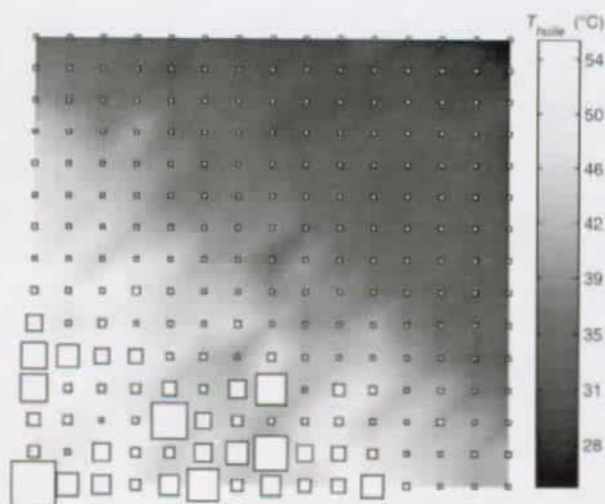


FIG. 6.10 – Exemple d'histogramme d'activation (superposé à la représentation colorée d'une couche).

Le principe de l'histogramme d'activation (fig. 6.10) est de représenter la probabilité d'activation de chaque neurone directement sur la carte construite selon l'une des représentations décrites aux paragraphes précédents. À l'emplacement de chaque neurone, un carré dont la surface est proportionnelle à sa probabilité d'activation est dessiné. Le calcul de l'histogramme dépend des critères relatifs aux conditions de fonctionnement qui doivent être englobées. Typiquement, en tenant compte du fait que le comportement d'un transformateur est périodique (cycle annuel), la probabilité d'activation locale est obtenue en ne considérant l'acti-

vation du réseau que pour les éléments d'apprentissage ou les valeurs des états qualifiés de sains lors des quelques jours précédant les mesures courantes.

## 6.5 Point de fonctionnement et trajectoire

### 6.5.1 Neurone gagnant

Lors de l'apprentissage selon l'algorithme kMER, plusieurs neurones peuvent être activés simultanément. Si un vecteur  $\mathbf{v}$  tombe dans la zone d'activation du neurone  $i$  définie par son rayon  $\sigma_i$ , celui-ci est retenu indépendamment du fait que d'autres neurones soient activés ou non. En revanche, si le vecteur ne tombe dans aucune zone d'activation, il n'y a pas de gagnant. Pour l'exploitation de la carte, le neurone gagnant  $i^*$  est simplement désigné, comme pour les cartes de Kohonen, en utilisant la distance euclidienne rappelée ici :

$$i^* = \arg \min_i \|\mathbf{w}_i - \mathbf{v}\| \quad (6.2)$$

où  $\mathbf{w}_i$  est le vecteur-poids du  $i^{\text{ème}}$  neurone.

### 6.5.2 Représentation graphique

Les méthodes de représentation qui viennent d'être développées ont toutes en commun le fait que l'ensemble des neurones est visualisé sous la forme d'une surface rectangulaire (le plus souvent carrée). Les neurones y sont disposés régulièrement et peuvent être repérés par leurs coordonnées spatiales. D'un point de vue ergonomique, s'agissant de phénomènes non-linéaires, cette représentation est excellente : elle rend des données par nature irrégulières exploitables visuellement de manière instinctive ; la complexité relative à la dimension des données est réduite sans perdre l'essentiel des informations.

Quel que soit le mode de représentation choisi (matrice des distances, décomposition en couches, etc.), la structure sous-tendant l'affichage est identique. Ce dernier peut être considéré comme un "papier peint" aidant à l'interprétation des résultats. Celui-ci sera accompagné de la dé-

limitation des zones de fonctionnement (§ 6.1.1) ou/et de l'histogramme d'activation (§ 6.4.4).

La délimitation des zones est appelée étiquetage et suit l'apprentissage. Cette phase consiste à établir la structure de la carte en fonction des paramètres d'exploitation jugés pertinents. Par exemple, il peut être souhaitable d'identifier la zone de la carte représentant l'exploitation durant une certaine période de l'année. Pour ce faire, les vecteurs d'apprentissage correspondant aux jours considérés sont présentés à la carte. Les neurones gagnants définissent alors la zone correspondant aux conditions choisies.

Afin de synthétiser l'information au maximum, pour certaines grandeurs critiques, les résultats de l'analyse en couches peuvent également être utilisés. Par exemple, les neurones peuvent être étiquetés suivant que la valeur d'un capteur dépasse ou non un certain seuil jugé critique. Lorsque toutes ces opérations sont réalisées, la carte est prête à être exploitée avec des données mesurées en continu. À chaque instant, le vecteur de mesure est présenté à la carte. Le neurone le plus proche est alors recherché. Celui-ci définit l'emplacement du point de fonctionnement sur la représentation graphique. Si pendant une période définie chaque point de fonctionnement est relié par un trait, on obtient alors une trajectoire. Celle-ci permet de mettre en évidence non seulement l'état du système mais aussi son évolution (fig. 6.1).

## 6.6 Processus décisionnel

### 6.6.1 Interprétation directe

La première méthode d'interprétation des résultats de surveillance continue est l'analyse visuelle de l'évolution du point de fonctionnement et de la trajectoire du processus. Les interprétations sont alors pondérées par :

- l'organisation de la carte suivant la valeur d'une ou plusieurs variables spécifiques ;
- les plages de fonctionnement distinctes identifiées par la matrice des distances et la procédure d'étiquetage ;

- l'histogramme d'activation pour les conditions de fonctionnement proches de celle considérée.

Cette liste reprend l'ensemble des techniques citées plus haut. Chaque apporte une information essentielle qu'il est difficile de synthétiser. D'un point de vue systématique, seule la localisation du point de fonctionnement dans l'une des zones classifiées permet d'avoir des résultats directs. L'interprétation de l'évolution de la trajectoire vers une zone ou une autre demande une interprétation assez fine. Tout au plus, on peut imaginer un signal d'alarme dès que le point de fonctionnement se rapproche d'une des limites avec un ordre de voisinage prédéterminé. Si la trajectoire s'éloigne subitement et de manière durable de la zone d'activité repérée par les carrés les plus gros, un changement est intervenu. En dehors de ces cas, l'interprétation directe permet essentiellement de fournir des indications qualitatives à l'opérateur surveillant un processus. Elles lui permettront de prendre des décisions rapidement lors d'alertes ou lorsqu'il s'agira de mettre le système sous un stress particulier (par exemple s'il s'agit de surcharger momentanément un transformateur).

### 6.6.2 Distance et indicateur de confiance

Pour l'application particulière aux transformateurs de puissance, l'approche classique développée jusqu'ici nécessite une adaptation. En effet, en raison des contraintes économiques évoquées à plusieurs reprises, seuls les états de fonctionnement corrects sont modélisés. L'identification de problèmes suivant la zone dans laquelle se trouve le point de fonctionnement n'est rigoureusement réalisable que lorsque des données correspondant aux défauts que l'on veut identifier sont disponibles.

La figure 6.11 illustre un tel cas pour un processus caractérisé par deux variables. Les états sains forment un "U"; ils ont été appris par le réseau de neurones. Tout état situé hors de cette zone doit être considéré comme défaillant. L'évolution du processus d'une situation normale vers un défaut peut être observée par le déplacement du neurone gagnant vers les bords du "U". Sur la représentation de la matrice des distances (fig. 6.9b), ces points sont situés sur le pourtour de la carte et aux abords des surfaces claires. Cette information ne peut être que qualitative: rigoureusement, ces neurones représentent des états admis. En revanche, comme cela a été vu au chapitre 5, chaque couple  $\{\mathbf{v}; \mathbf{w}_i\}$  associant

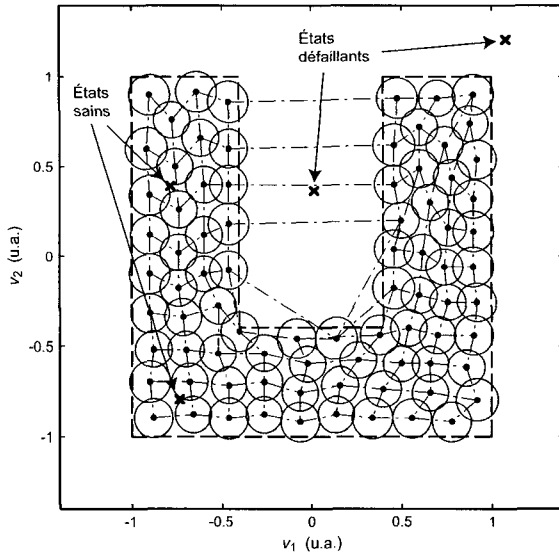


FIG. 6.11 – Définition des zones de fonctionnement modélisées par un réseau kMER (cas simplifié avec deux instruments).

un vecteur d'entrée à son neurone gagnant est caractérisé par une distance. En fixant un seuil sur cette distance, il est possible d'identifier les états qui ne correspondent manifestement à aucun mode de fonctionnement modélisé [57]. Dans ce cas, une alarme doit être déclenchée et une analyse plus fine doit être conduite. De part sa structure, l'algorithme kMER apporte une précision supplémentaire. En effet, comme les résultats du paragraphe 6.3.1 le montrent, l'indicateur de confiance  $IC$  (§ 5.7.7) fournit une mesure plus robuste que la simple indication de la distance.

Le graphique de la figure 6.12a a été obtenu pour le cas modélisant l'espace d'entrée dont la répartition des points n'est pas uniforme (fig. 6.5). Une coupe a été définie selon la diagonale traversant l'espace de fonctionnement du coin inférieur gauche au coin supérieur droit. Les distances et l'indicateur  $IC$  sont représentés en fonction du type de carte (KSOM ou kMER) et du couple  $\{\mathbf{v}; \mathbf{w}_{i^*}\}$ . Le nombre de neurones étant

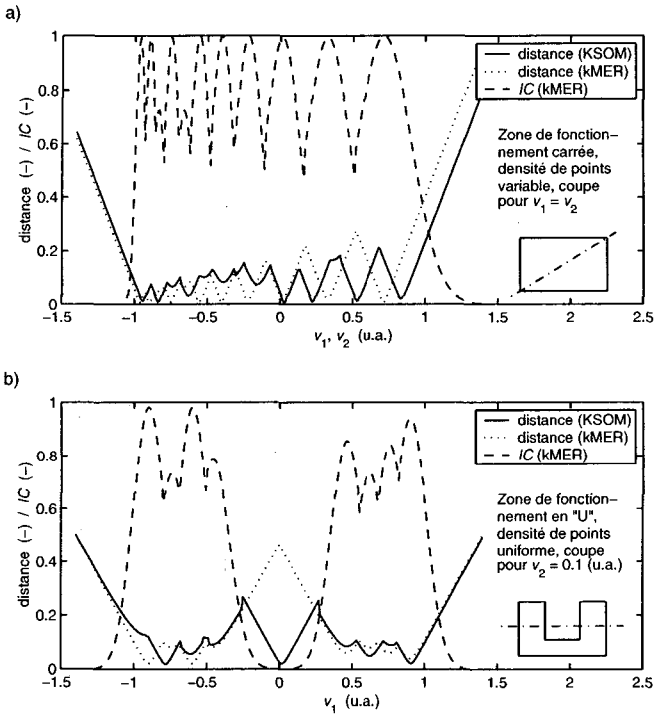


FIG. 6.12 – Illustration de l'utilisation de la distance et de l'indicateur de confiance IC pour l'identification des états défaillants (cas simplifié de deux instruments).

fixe, ces valeurs oscillent respectivement dans l'intervalle  $[0; 1]$  pour les distances et  $[1; 0]$  pour IC. La variation de ces mesures le long de la coupe se caractérise par une augmentation plus grande de la variance dans le cas des distances que dans celui de l'indicateur de confiance. Le seuil d'alarme doit impérativement tenir compte de ces variations. Celui-ci pourra donc être plus sensible s'il est fixé sur IC que sur la distance. Si cette dernière est utilisée, la distance maximale le long de la coupe pour un vecteur  $\mathbf{v}$  appartenant à l'ensemble d'apprentissage fixe le niveau du seuil. Dans le cas de l'indice de confiance, c'est son niveau minimal sur le même ensemble qui détermine le seuil. En utilisant respectivement la



carte KSOM et la carte kMER, la sensibilité du système est comparable pour les valeurs situées à droite. En revanche, à gauche pour les zones où la densité de points est plus grande, le seuil de détection basé sur  $IC$  s'avère être plus sensible.

Les courbes du graphique 6.12b sont construites selon le même principe pour une coupe horizontale. Elles montrent les différentes réponses dans le cas de la zone de fonctionnement définie par les mesures réparties uniformément en "U". Les neurones "morts" mis en évidence au paragraphe 6.3.1 introduisent un effet indésirable sur la variation de la distance. L'interprétation de celle-ci amène le système basé sur la carte KSOM à interpréter de manière erronée la majorité des points situés au centre du "U" (en dehors de l'ensemble d'apprentissage). Pour le réseau kMER, que la distance ou l'indice de confiance soit utilisé, l'apprentissage sans neurone "mort" permet d'éliminer ce problème.

### 6.6.3 Interprétation robuste de l'indicateur $IC$

Conformément au problème soulevé à la section 4.5, la détermination du niveau du seuil et l'ajustement des paramètres de calcul dépend d'un compromis établi entre la sensibilité de détection et la probabilité de fausse alarme. Ces deux caractéristiques sont toujours opposées. Il s'agit donc de trouver un optimum.

Le graphique de la figure 6.13 montre que l'indicateur de confiance est relativement sensible aux variations : qu'il s'agisse de son évolution typique dans des conditions normales ou lors de l'apparition d'un défaut, le signal est relativement bruité. Le lissage de la sortie ne peut être obtenu qu'au prix d'une réduction de sensibilité de détection. Les deux alternatives possibles consistent soit à modifier la fonction d'estimation de la densité de probabilité, soit à traiter la sortie à l'aide d'une méthode d'interprétation robuste (§ 4.5.4). La première technique, revient à modifier le niveau d'alarme de manière non-linéaire en introduisant un facteur au dénominateur de la fonction d'estimation. La sensibilité du système s'en trouve systématiquement réduite. La seconde méthode est de nature intégrale : l'augmentation d'un signal d'alarme dépend à la fois de la vraisemblance de l'erreur mais aussi de sa durée. Cette technique retarde certes l'apparition des alarmes mais la sensibilité de

détection n'est pas affectée. Elle est donc particulièrement bien adaptée au problème considéré.

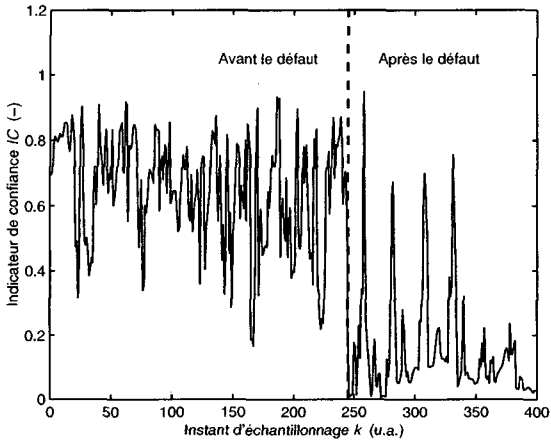


FIG. 6.13 – Évolution typique de l'indicateur de confiance  $IC$  dans le cadre de la surveillance de transformateur (chap. 7) : élévation soudaine de la température de l'huile pour  $k = 245$  (u.a.).

Considérant le grand nombre de cas tests en fonctionnement normal, l'indicateur de confiance  $IC$  peut être considéré comme une variable gaussienne. Sa valeur moyenne et sa variance dépendent de la qualité du réseau de neurones et peuvent être établies expérimentalement. Lors de l'apparition d'un défaut, la valeur de  $IC$  diminue; le niveau dépend de la sévérité du problème. Des études de sensibilité comme celles qui sont décrites au paragraphe 7.4.2 permettent de déterminer les paramètres de la méthode d'interprétation robuste de façon adéquate. En fixant un seuil d'alarme sur cette sortie intégrale, le système permet de détecter automatiquement toute évolution vers un état inconnu. Le niveau de ce seuil n'étant pas limité à une seule valeur, un certain nombre d'alertes peuvent également être définies.

## Chapitre 7

# Évaluation du système de surveillance

### 7.1 Installation pilote

#### 7.1.1 Caractéristiques principales

La validation de l'approche développée a été réalisée avec les données fournies par une unité pilote spécialement équipée de capteurs permettant de mesurer en continu un certain nombre de paramètres de fonctionnement.

Le transformateur de puissance utilisé (fig. 7.1) est situé à Fiesch en Valais. Son rôle est essentiellement le raccordement de la production locale des usines (turbines hydrauliques au fil de l'eau) sur le réseau électrique haute tension. La charge du transformateur dépend donc uniquement de la production des usines et est, par la nature même des unités de production, irrégulière.

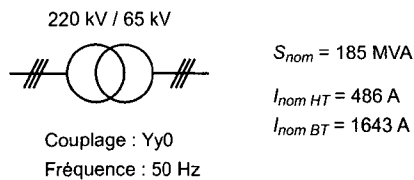


FIG. 7.1 – *Caractéristiques principales du transformateur pilote.*

La situation géographique de l'unité engendre un grand stress de fonctionnement ; l'activité orageuse est particulièrement importante en été et occasionne de fréquentes surtensions sur la ligne de transport à laquelle il est raccordé. Celles-ci ont été à l'origine d'une défaillance majeure ayant nécessité la révision complète du transformateur chez le fabricant.

### 7.1.2 Grandeurs mesurées

Les capteurs ont été sélectionnés suivant des critères relatifs à la détection d'un maximum de défauts de nature différentes, leur fiabilité et leur coût. Sans entrer dans les détails techniques, la solution implémentée regroupe les mesures suivantes :

- courant de charge  $I_{ch}$  ;
- courants de court-circuit  $I_{cc}$  ;
- tension de service  $U_{BT}$  (borne 65 kV) ;
- température ambiante  $T_{air}$  ;
- température de la cuve  $T_{cuve}$  ;
- température de l'huile  $T_{huile}$  ;
- gaz dissous (hydrogène essentiellement)  $Gaz$  ;
- taux d'humidité  $Hum$  ;
- vibrations de la cuve  $Vib$  ;
- surtensions ;
- alarmes de service.

L'essentiel des données est récolté par un système d'acquisition commercial. Les mesures de surtensions et de vibrations y font exception ; leur acquisition est contrôlée par un oscilloscope relié à un PC. Les deux systèmes sont reliés au réseau téléphonique par modem. Les données peuvent ainsi être consultées en temps réel ou téléchargées depuis un site distant (fig. 7.2).

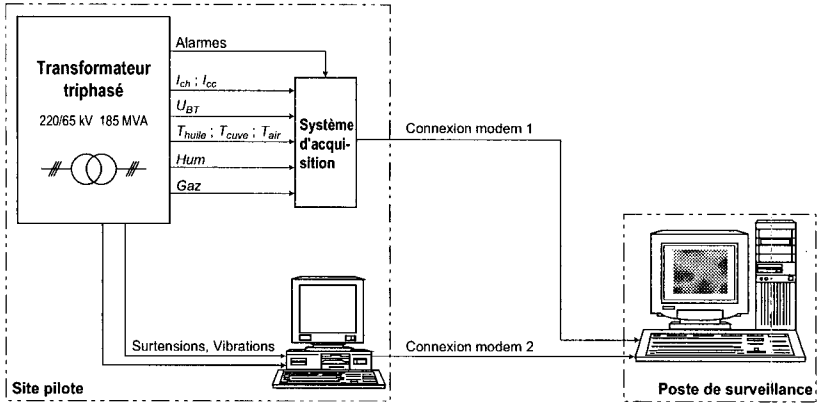


FIG. 7.2 – Dispositif de mesure installé sur le transformateur pilote.

## 7.2 Définition du vecteur d'entrée

### 7.2.1 Mesures brutes

Pour des raisons relatives à l'optimisation de l'utilisation de la mémoire tampon disponible sur les systèmes d'acquisition, toutes les données, à l'exception des mesures de surtensions et de vibrations, sont mesurées en continu et moyennées à raison d'un échantillon par heure. La base de données ainsi constituée représente plus de 35000 valeurs pour chacun des capteurs, recueillies de janvier 1997 à décembre 2000. De par leur nature irrégulières, les mesures de surtensions ne sont pas disponibles en continu. De ce fait, le traitement de ces données n'a pas été considéré. En ce qui concerne les mesures de vibrations, le signal fourni par le capteur est numérisé périodiquement. Une transformée de Fourier (FFT) est directement effectuée et seules les amplitudes des neuf premières harmoniques de rang pair sont conservées (cf. [2] pour plus de détails).

Les données de la figure 7.3 représentent les variations de cinq des paramètres mesurés au cours du temps. Les plages correspondant à des va-

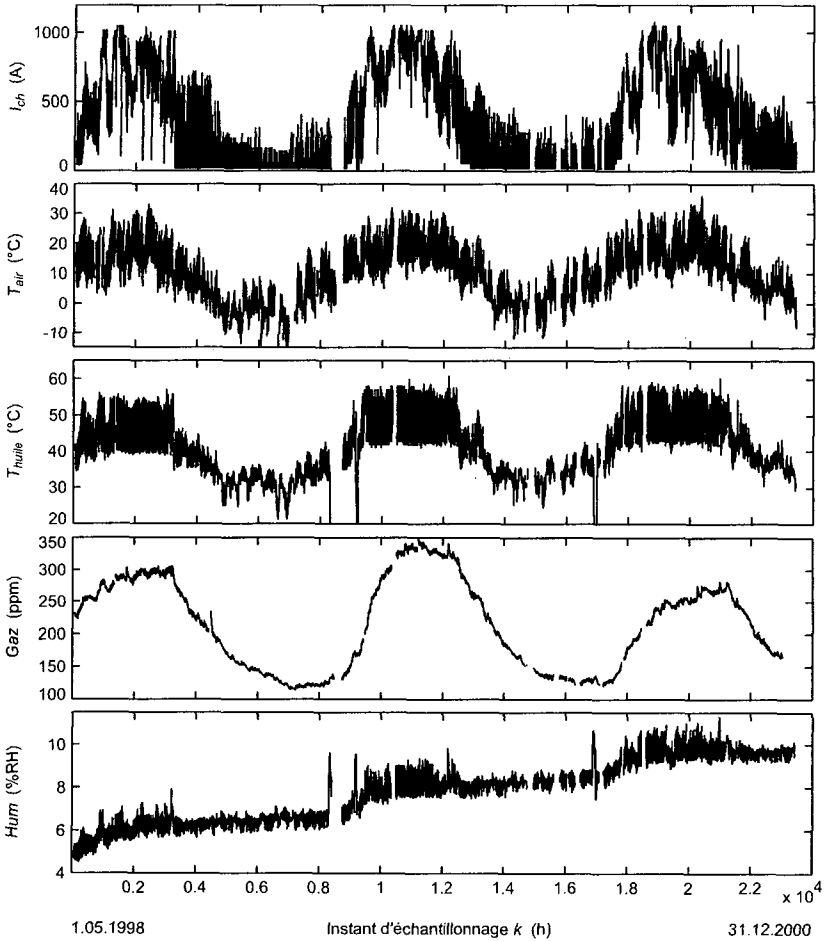


FIG. 7.3 – Mesures du courant de charge, de la température ambiante, de la température de l'huile, du taux de gaz dissous et de l'humidité entre le 1<sup>er</sup> mai 1998 et le 31 décembre 2000 sur le transformateur de Fiesch. Les zones vides correspondent à des mesures erronées ou manquantes. Les trois sauts brusques du niveau d'humidité coïncident avec les périodes de maintenance pour lesquelles la température de l'huile chute (transformateur arrêté).

leurs erronées ou manquantes ont été laissées en blanc. L'analyse visuelle de ces grandeurs permet de tirer un certain nombre d'enseignements :

- la charge varie selon un cycle annuel ;
- la température de l'huile est fortement corrélée à celle de l'air en hiver et est régulée en été entre 40 °C et 55 °C environ ;
- la variation du taux de gaz suit une périodicité annuelle à laquelle peut être ajoutée une composante journalière ;
- le taux d'humidité présente une forte tendance à l'augmentation et sa variation locale est nettement plus marquée en été qu'en hiver.

### 7.2.2 Analyses préliminaires des données

Les constatations qui viennent d'être énoncées sont confirmées par les analyses qui suivent. Tout d'abord, les graphiques de la figure 7.4 mettent en évidence la différence de types de fonctionnement entre l'été et l'hiver. L'été, le transformateur est chargé (fig. 7.4b) alors que l'hiver, il n'est sollicité que par intermittence et à des niveaux de courant relativement peu élevés (fig. 7.4a). En hiver, la variation de la température de l'huile est principalement influencée par l'évolution de la température ambiante. La faible charge ne suffit généralement pas à élever la température au-delà de 40 °C. En revanche, en été la charge contribue à une élévation importante de la température de l'huile. Comme cela a été mentionné plus haut, ni la variation de la charge, ni la variation de la température de l'air ne peuvent expliquer l'allure de la variation de la température de l'huile. L'explication à la succession de flancs exponentiels croissants et décroissants apparaissant sur la courbe de température de l'huile, est sa régulation au moyen des groupes de refroidissement à disposition. Sur la base de ces observations, la constante de temps typique du phénomène peut être estimée à quelques heures. Conformément à ce que montre la figure 7.4, compte tenu de la valeur de la constante de temps thermique ainsi que de la prédominance du mode de fonctionnement des groupes de refroidissement sur le système, une variation de charge limitée dans le temps n'a pas d'influence visible sur la température de l'huile. Ces graphiques mettent encore en évidence, en été comme en hiver, la corrélation négative entre la température de l'huile et le taux d'humidité évoquée au paragraphe 2.5.4.

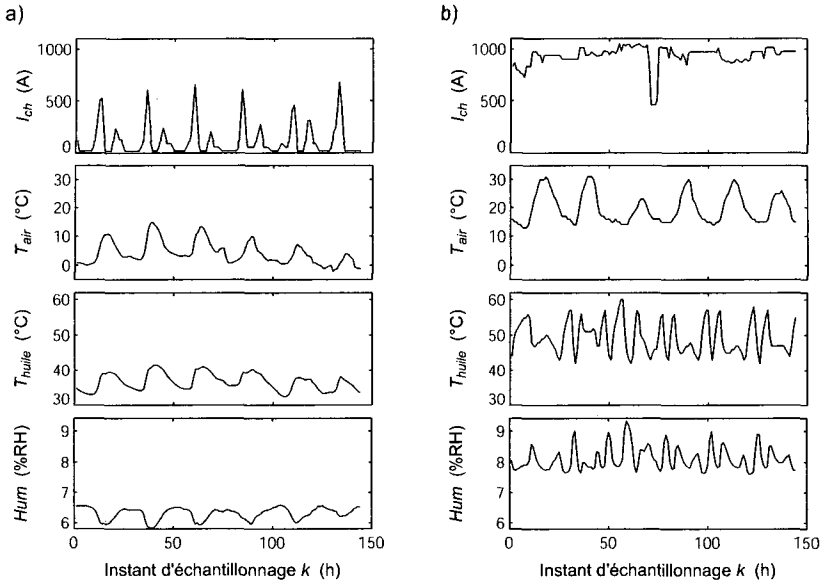


FIG. 7.4 – Mesures typiques effectuées sur le transformateur pilote : (a) fonctionnement hivernal, période du 21-26.10.1998, (b) fonctionnement estival, période du 26-31.07.1999.

La figure 7.5 montre les corrélations linéaires existant entre les différentes variables. Le graphique est à lire comme un tableau où les éléments de la diagonale désignent les variables. Le coefficient de corrélation linéaire de chaque paire de variable est représenté par un rectangle. Celui-ci est situé à l'intersection des axes désignés respectivement par les deux variables. Les dimensions de chaque rectangle sont proportionnelles à la valeur absolue du coefficient de corrélation linéaire associé. Les deux couleurs définies distinguent les corrélations positives (gris foncé) des négatives (gris clair). En outre, le graphique est divisé en deux parties : la zone inférieure gauche représente les corrélations calculées pour l'ensemble des données mesurées sur deux ans (la tendance observée sur l'humidité a été supprimée à l'aide d'une fonction linéaire) ; la partie supérieure droite a été obtenue quant à elle pour les données filtrées. Le filtrage effectué consiste simplement à soustraire aux valeurs originales



leur moyenne mobile calculée sur 3 heures. Seules les variations locales sont ainsi prises en compte.

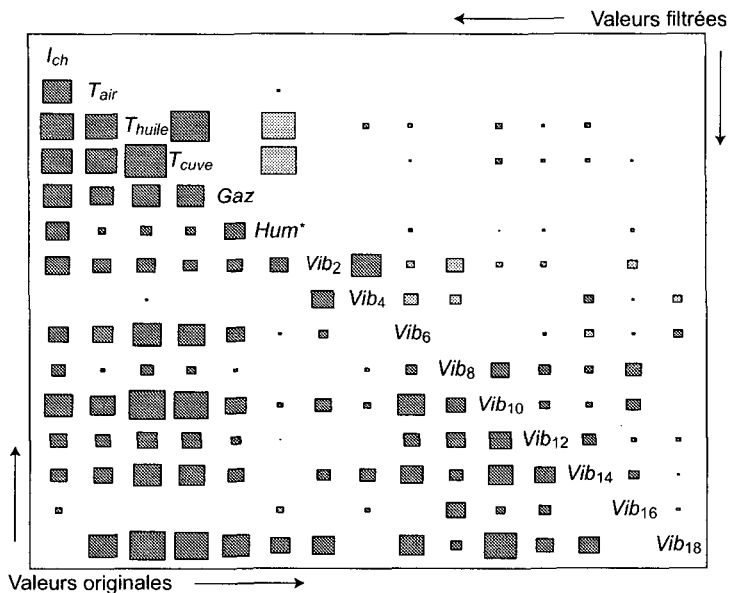


FIG. 7.5 – *Corrélations linéaires calculées pour l'ensemble des données récoltées sur 2 ans. Les côtés de chaque rectangle sont proportionnels aux coefficients des deux variables associées. Le gris clair représente les corrélations négatives et le gris foncé les corrélations positives.*

Au niveau global, l'analyse de cette figure montre une corrélation linéaire généralement positive entre toutes les variables. Ceci était attendu compte tenu de l'allure générale des mesures (fig. 7.3). Il faut tout de même noter, toujours globalement, que le taux d'humidité ainsi que plus de la moitié des harmoniques des vibrations ne sont corrélés de manière significative à aucune variable. Les valeurs filtrées mettent quant à elles en évidence la forte corrélation positive entre la température de l'huile  $T_{huile}$  et la température de la cuve  $T_{cuve}$  ainsi que la corrélation négative entre ces deux valeurs et le taux d'humidité relative  $Hum^* 1$ . Une modé-

1. Taux d'humidité auquel la tendance a été éliminée (fonction linéaire).

lisation relativement fine du processus est donc nécessaire pour prendre en compte cette dernière corrélation.

En se basant sur les connaissances théoriques et les observations réalisées, le lien entre les différentes variables mesurées est établi. Il est évident que les relations intervenant dans le processus dépendent de façon non linéaire du mode de fonctionnement du transformateur. Les deux graphiques de la figure 7.6 illustrent parfaitement cette constatation. La corrélation linéaire pour les deux couples de variables  $\{T_{huile}; T_{air}\}$  et  $\{T_{huile}; Hum\}$  est représentée localement en fonction du temps pour un calcul effectué sur une période (glissante) de deux semaines centrée à l'instant d'échantillonnage  $k$ . De manière générale, l'anticorrélation entre la température de l'huile et le taux d'humidité est confirmée. Cependant, elle n'est réellement significative que lors des périodes hivernales. En ce qui concerne la température ambiante, sa corrélation avec la température de l'huile n'est observable qu'en hiver (graphique du bas, corrélation de 0.9). En été, le même graphique montre clairement que ces deux valeurs ne sont plus corrélées linéairement de façon significative.

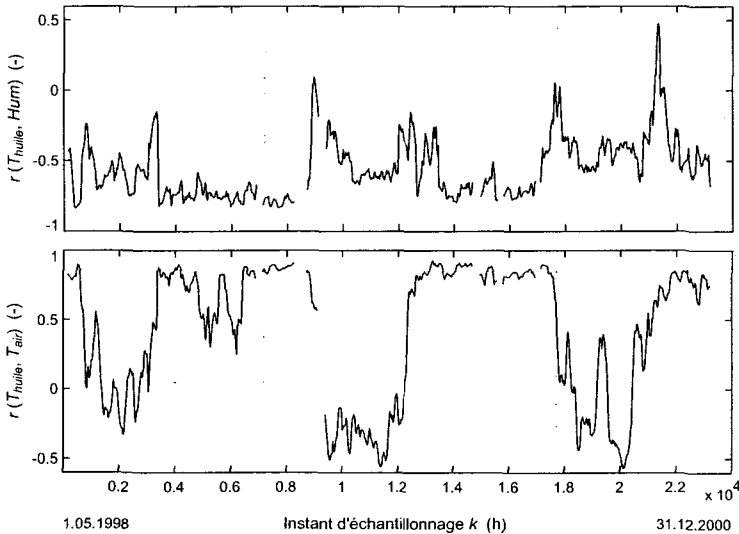


FIG. 7.6 – Évolution des corrélations linéaires au cours du temps (calculées pour une fenêtre glissante de 2 semaines).

Toutes les interactions ne peuvent pas être mises en évidence avec les techniques utilisées jusqu'ici. L'influence croisée des différents paramètres ainsi que celle de variables non mesurées (p.ex. les paramètres de fonctionnement du système de refroidissement) sont difficilement décelables. Les mesures de vibrations devraient fournir une information utile concernant le mode de fonctionnement des ventilateurs. Cependant, leur observation et l'analyse des corrélations ne donnent aucun résultat significatif. Comme le montre la figure 7.7, des phénomènes physiques supplémentaires viennent rendre la modélisation encore plus complexe. Ce graphique montre, pour un cycle de fonctionnement régulier, l'évolution du taux d'humidité en fonction de la température de l'huile pendant un jour. Les symboles "\*" représentent les premier et dernier points pris en compte. Le sens de la trajectoire est donné par la flèche. Les dimensions de la forme elliptique décrite par les points reflètent la dynamique de libération et d'adsorption de l'eau par la cellulose.

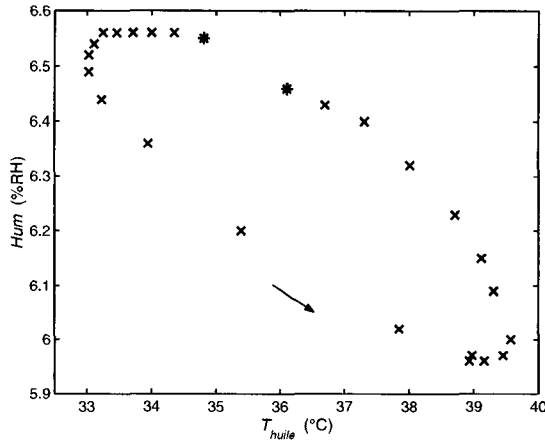


FIG. 7.7 – Dynamique entre le taux d'humidité *Hum* et la température de l'huile *T<sub>huile</sub>* pour un jour (21.10.1998).

Les cartes auto-organisées, on l'a vu, ont la propriété de réduire la complexité d'un ensemble de données en effectuant une projection sur un espace restreint. Elles ont aussi la particularité de prendre en compte

automatiquement les corrélations locales [120]. Leur utilisation selon le principe général décrit au chapitre 6 permet ainsi de construire un modèle global pour l'ensemble des données qui ne se limite pas exclusivement à la modélisation de la variation annuelle.

### 7.2.3 Sélection et prétraitement des variables

Les informations collectées par les analyses classiques constituent la première étape de sélection des variables et de leur prétraitement. Chaque état a été considéré jusqu'ici sans tenir compte de l'évolution temporelle et des constantes de temps liant les variables entre elles. L'approche qui consiste à évaluer les corrélogrammes<sup>2</sup> ne peut pas être mise en œuvre de manière systématique. Les expériences faites ont montré qu'il est possible de mettre en évidence une périodicité journalière. En revanche, le décalage maximisant la corrélation lorsqu'elle est positive et la minimisant lorsqu'elle est négative s'avère être très difficile à trouver. En ajoutant un degré de complexité supplémentaire, l'application de cette technique se heurte aux problèmes déjà évoqués (variation de la corrélation locale, influences mutuelles de plusieurs variables, relations cachées). Il n'est donc pas envisageable de l'utiliser de manière généralisée.

A priori, il est important de n'écarter aucune relation intervenant entre les différentes variables. La méthodologie utilisée pour construire le vecteur définissant l'état de fonctionnement consiste donc à prendre en compte toutes les valeurs. En procédant par apprentissages successifs, c'est l'organisation du réseau qui va guider le processus de sélection et de prétraitement de ces données. Le vecteur de fonctionnement initial est défini par

$$\zeta(k) = [v_1(k), v_2(k), \dots, v_n(k), v_1(k-1), v_2(k-1), \dots, v_n(k-1)] \quad (7.1)$$

où  $n$  est le nombre de variables collectées et  $k$  est l'instant d'échantillonnage. À ce stade, à l'exception de l'humidité, toutes les valeurs sont directement fournies par le système d'acquisition. L'humidité, comme

---

2. Un corrélogramme est la représentation graphique du coefficient de corrélation linéaire de deux variables en fonction de leur décalage temporel [119]. Cette technique permet de mettre en évidence les constantes de temps existant entre deux variables liées.

pour les analyses de corrélations, subit un prétraitement afin d'éliminer sa tendance (augmentation à long terme). L'expérience a montré que l'approximation de cette tendance par une droite est appropriée.

À l'issue de chaque apprentissage, l'organisation du réseau de neurones est évaluée de la façon suivante. Deux ensembles de tests sont définis. Le premier est constitué d'états faisant partie de l'ensemble d'apprentissage. Le deuxième comporte une série de mesures proches de cet ensemble mais n'ayant pas été prises en compte lors de l'apprentissage. Concrètement, si l'on prend en compte treize mois de mesures, les douze premiers mois constitueront l'ensemble d'apprentissage et le treizième mois sera l'ensemble de validation. Comme on le verra plus loin, cette manière de faire est conforme au contexte d'utilisation. Elle permet non seulement d'évaluer la précision de la réponse obtenue mais aussi la généralisation du modèle. La qualité de la réponse est évaluée en fonction de l'erreur absolue moyenne de chaque variable observée séparément. Pour le capteur  $j$ , il vient :

$$qual_j = \frac{\sum_{k=1}^m |w_{i^*j}(k) - v_j(k)|}{m} \quad (7.2)$$

où  $m$  est le nombre d'échantillons,  $v_j$  est la valeur effectivement mesurée,  $i^*$  est l'indice du neurone gagnant et  $w_{i^*j}$  le poids du neurone gagnant associé.

Pour s'assurer de prendre en compte de la même manière toutes les variables, leur valeurs sont centrées et réduites conformément à la procédure décrite au paragraphe 6.2.3. Compte tenu des propriétés de l'algorithme d'apprentissage, les variables dont l'indicateur de qualité  $qual_j$  est manifestement plus élevé que les autres ne peuvent pas être modélisées. Leur variation n'est pas liée à celle des autres variables. Elles représentent du bruit pour le réseau de neurones et affectent la qualité de la sortie tout en ne contribuant aucunement à la modélisation. En éliminant itérativement les valeurs qui n'ont manifestement pas été modélisées correctement, le vecteur d'état se réduit alors à la prise en compte des six variables  $I_{ch}$ ,  $T_{air}$ ,  $T_{huile}$ ,  $T_{cuve}$ ,  $Gaz$  et  $Hum$ .

L'analyse des poids du réseau montre que la forte corrélation entre les variables  $T_{huile}$  et  $T_{cuve}$  se traduit par une organisation identique de leurs couches respectives. Du point de vue de la robustesse de détection d'erreur, cette redondance rend le système moins sensible à la défaillance

de l'un de ces deux capteurs de température. Cependant, le total de l'information peut être obtenu en ne prenant en compte qu'une seule de ces deux températures. C'est pourquoi, afin d'avoir une répartition des variables la plus uniforme possible, la température de la cuve représentant une mesure indirecte de la température de l'huile a été retirée du vecteur  $\zeta(k)$ . Par ailleurs, l'étude de l'organisation des couches met en évidence le fait que la notion d'évolution temporelle n'intervient pas. En effet, les valeurs des poids modélisant les grandeurs  $v_j(k)$  et  $v_j(k-1)$  s'avèrent être identiques pour chaque neurone. Pour simplifier l'apprentissage, le vecteur d'état est encore réduit ; seul l'instant d'échantillonnage  $k$  est considéré.

Enfin, l'analyse de la figure 7.3 montre que le taux de gaz varie selon une périodicité annuelle avec une amplitude variable. Compte tenu du faible nombre d'années de mesures à disposition, il n'est statistiquement pas envisageable de modéliser avec précision ces variations. L'objectif étant de détecter des changements de faible amplitude, il a été décidé de rendre le système indépendant de ces variations en appliquant un filtre relativement élémentaire :

$$Gaz_{\Delta}(k) = Gaz(k) - \frac{1}{72} \sum_{i=k-71}^k Gaz(i). \quad (7.3)$$

Étant donné que l'humidité est également affectée par une variation à long terme (mais sous forme de tendance), la même technique de filtrage lui est appliquée :

$$Hum_{\Delta}(k) = Hum(k) - \frac{1}{72} \sum_{i=k-71}^k Hum(i). \quad (7.4)$$

Finalement, le vecteur  $\zeta(k)$  retenu pour représenter l'état du transformateur à l'instant  $k$  est défini par

$$\zeta(k) = [I_{ch}(k), T_{air}(k), T_{huile}(k), Gaz_{\Delta}(k), Hum_{\Delta}(k)]. \quad (7.5)$$

## 7.3 Implémentation

### 7.3.1 Console de surveillance

Comme cela a été défini au paragraphe 6.1.2, la méthodologie développée ici met en œuvre les cartes auto-organisées dans le cadre de la détection de changements rapides (de l'ordre de 72 heures et moins, compte tenu de la définition de  $\zeta$ ). Le traitement des tendances à plus long terme ne sera pas abordé, essentiellement en raison du manque de données expérimentales : les premières constatations faites sur les mesures du transformateur pilote de Fiesch montrent que pour avoir une base de données statistiquement exploitable, le nombre d'années de mesures nécessaires est bien supérieur aux quatre ans à disposition.

Comme le montrent les cas traités à la fin de ce chapitre, les informations fournies par la méthode sont les suivantes :

- gestion intelligente des alarmes (seuils adaptatifs dépendant de l'état de fonctionnement) ;
- interface graphique principale permettant de situer le point de fonctionnement du transformateur et son évolution par rapport à l'ensemble des états modélisés ;
- visualisation graphique de l'évolution de la fonction de décision permettant d'évaluer la qualité de la surveillance à l'aide de seuils intermédiaires dit "d'alerte".

### 7.3.2 Zones de fonctionnement

La figure 7.8 montre l'organisation typique de la carte auto-organisée kMER retenue pour la modélisation du comportement du transformateur de Fiesch. Le réseau est composé de 225 neurones qui ont été entraînés pour une base de données de deux ans. Le premier graphique (a) représente la matrice des distances (§ 6.4.3). La position des neurones est indiquée par les carrés blancs. Les zones noires représentent les valeurs faibles et les zones jaunes voire blanches, les grandeurs les plus élevées. Les cinq graphiques suivants illustrent l'organisation des couches constituées en fonction du vecteur  $\zeta$  (§ 6.4.2). Pour marquer la différence avec le premier graphique, l'échelle de couleur est définie du bleu pour les valeurs les plus basses au rouge pour les valeurs les plus

élevées. Ces couleurs sont définies indépendamment pour chaque couche en fonction de leurs valeurs minimale et maximale respectives.

L'analyse de ces graphiques permet tout d'abord de confirmer les premières constatations faites au paragraphe 7.2.2 : deux régimes de fonctionnement distincts sont mis en évidence et la variation de certaines corrélations locales est observée. Plus précisément, comme le montre le graphique (b), la zone supérieure droite est caractérisée par une charge faible et uniforme. Pour cette région de la carte, la température de l'air est faible. La comparaison des graphiques (c) et (d) pour cette même zone met encore en évidence le fait que la température de l'huile est étroitement liée à celle de l'air. La zone inférieure gauche représente quant à elle les états les plus chargés. Si les niveaux moyens des températures sont supérieurs à ceux observés pour la première zone, l'influence de la température ambiante  $T_{air}$  sur la température de l'huile  $T_{huile}$  n'est quasiment pas observée. On peut montrer facilement que les deux régions de la carte qui viennent d'être décrites correspondent respectivement à l'hiver et à l'été. La carte montre également que la délimitation entre ces deux zones n'est pas nette. Un certain nombre de neurones, répartis selon la diagonale qui va du coin supérieur gauche au coin inférieur droit de la carte, représentent des états de transition pour lesquels le transformateur est à mi-charge.

L'interprétation de l'organisation des couches représentant les taux de gaz dissous (e) et d'humidité dans l'huile (f) est plus difficile. Les zones pour lesquelles le taux de gaz est le plus élevé se situent sur le bord inférieur de la carte. Ces points représentent des états pour lesquels la température de l'huile et le courant de charge sont élevés. En ce qui concerne l'humidité, les valeurs les plus grandes sont localisées essentiellement dans la partie droite de la carte. Son organisation est relativement indépendante des deux modes de fonctionnement : l'humidité varie de la même manière durant l'été et l'hiver ; pour ces deux périodes, la carte montre que le taux d'humidité est anticorrélé à la température de l'huile. Les variations de cette dernière variable étant plus faibles en hiver qu'en été, le même phénomène est observé pour le taux d'humidité entre les quarts supérieur et inférieur droits du graphique.

La visualisation de la matrice des distances (fig. 7.8a) confirme qu'il n'y a pas de délimitation marquée entre les neurones représentant les



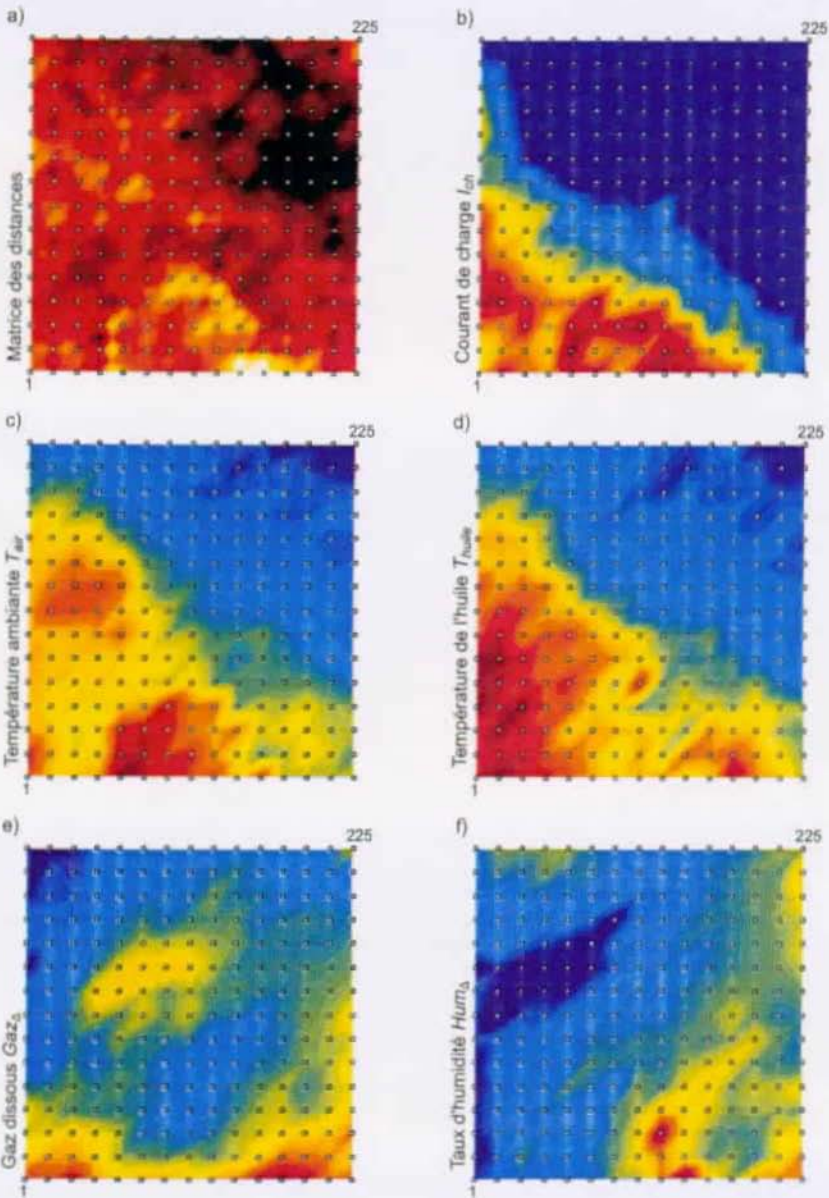


FIG. 7.8 – Organisation de la carte (transformateur de Fiesch 1998–1999) où les carrés blancs représentent la localisation des neurones. Échelles (selon ordre croissant des valeurs) : noir → blanc/jaune pour la matrice des distances ; bleu → rouge pour la représentation des couches.

deux modes de fonctionnement différents. La seule zone qui semble se détacher correspond aux températures ambiantes les plus élevées : la ligne claire observée pour la matrice des distances délimite assez précisément cette zone sur la représentation de la couche  $T_{air}$ .

Les constatations faites montrent que, après l'apprentissage, la carte représente non seulement les états de fonctionnement mais aussi qu'elle permet d'analyser la structure des données. Les états modélisés semblent constituer un espace relativement régulier où aucune zone ne se détache particulièrement. Par conséquent, le repérage précis des régions de la carte en fonction de critères d'exploitation est difficile dans ce contexte. En revanche, les corrélations locales sont bien mises en évidence par la représentation graphique. Cette méthode montre ici son efficacité à représenter de manière synthétique une information très riche. Cette propriété est particulièrement intéressante dans l'optique d'une représentation graphique efficace de l'état d'un transformateur.

### 7.3.3 Taille de la base de données

La méthodologie développée à la section 6.3 a mis en évidence le fait que, pour être efficace, la carte kMER doit avoir été entraînée avec des données comparables à celles qu'elle doit surveiller. Concrètement, en adoptant une approche qui consiste à ne construire qu'un seul réseau pour l'ensemble des états de fonctionnement, compte tenu du cycle annuel de variations de fonctionnement du transformateur, cela signifie que la base de données minimale nécessaire est d'une année.

Cette contrainte peut être assouplie si le domaine d'utilisation de la carte est réduit. Il est par exemple envisageable de définir une carte pour chaque mois. La base d'apprentissage de chacune de ces cartes pourrait être constituée des données récoltées lors du mois précédent et, si elles sont disponibles, des valeurs collectées les années précédentes à la même période. Au vu des variations statistiques d'un mois à l'autre mais également d'une année à l'autre, cette manière de procéder est de toute évidence moins robuste que l'approche globale. Sa mise en œuvre a néanmoins été explorée avec les données du transformateur de Fiesch pour une carte auto-organisée classique [139]. Les bons résultats obtenus semblent valider l'approche pour autant que les réapprentissage soient suffisamment fréquents (§ 6.3.2 et 7.3.4). Cependant, cette technique ne

paraît véritablement justifiée qu'au cours de la première année dans le cadre de la mise en place du système de surveillance.

Pour cette raison, afin d'avoir une modélisation la plus générale possible, la modélisation globale a été préférée. Pour l'apprentissage, un ensemble de données de deux ans est utilisé. Sa justification tient essentiellement à des considérations pratiques telles que la disponibilité de mesures fiables. En effet, il a été constaté, en procédant à la suppression des tendances sur le taux de gaz dissous et de l'huile, que l'augmentation de la base de données influence positivement les qualités de généralisation du réseau. Malheureusement, bien que l'unité pilote soit équipée de capteurs depuis 1997, suite à un mode de fonctionnement hivernal particulier en 1997-1998, seules les valeurs mesurées à partir du mois de mai 1998 ont pu être prises en compte. Un apprentissage est donc typiquement réalisé pour un ensemble de plus de 17500 vecteurs, soit deux ans de données.

### 7.3.4 Apprentissage et adaptation de la carte

Après avoir fixé le nombre de neurones du réseau (cf. § 7.4.3 pour une analyse de sensibilité), les paramètres d'apprentissage sont fixés selon la procédure décrite aux paragraphes 5.7.5 et 6.3.2. Pour la structure retenue ces paramètres sont les suivants : 225 neurones ;  $\rho = 1.5$  ;  $\sigma_{\Lambda 0} = 7.5$  ;  $\eta = 0.0005$  ;  $\gamma_{OV} = 1.03$  ;  $t_{max} = 320$ . La technique de détermination unique des paramètres lors du premier apprentissage s'avère fonctionner parfaitement : des déterminations effectuées pour des bases de données ultérieures fournissent systématiquement des paramètres du même ordre de grandeur.

Le taux d'actualisation proposé pour la carte est d'un réapprentissage par mois. Celui-ci est le fruit d'un compromis entre une adaptation rapide permettant au réseau de neurones de tenir compte au plus vite de certaines variations statistiques<sup>3</sup> et une actualisation plus lente visant à déceler l'évolution anormale de paramètres sur une période de plusieurs jours.

L'utilisation de la méthode de pré-organisation des poids décrite au paragraphe 6.3.2 permet d'assurer efficacement la continuité de l'orga-

---

3. La variation de la température extérieure d'une année à l'autre est un bon exemple de telles variations.

nisation générale du réseau. L'évolution n'est que locale ; les rotations et autres symétries sont évitées.

## 7.4 Sensibilité du système

Avant de tester le système avec des défauts dynamiques, sa sensibilité et son pouvoir de détection ont été évalués. Le but est de quantifier la réponse du système lorsque des valeurs biaisées (ou faussées) lui sont présentées en entrée. Ces valeurs faussées sont obtenues en additionnant sélectivement une erreur à une variable déterminée.

TAB. 7.1 – *Données statistiques des variables considérées (valeur moyenne  $\mu$  et variation annuelle  $\mathcal{V}$ , calculées pour 1999).*

	$I_{ch}$ (A)	$T_{air}$ (°C)	$T_{huile}$ (°C)	$Gaz$ (ppm)	$Hum$ (%RH)	$Gaz_{\Delta}$ (ppm)	$Hum_{\Delta}$ (%RH)
$\mu_i$	393	10.4	42.5	213	7.3	-0.1	0.0
$\mathcal{V}_i$	1070	52.2	40.0	220	3.5	37.0	2.83

Les tests ont été réalisés avec des valeurs artificiellement faussées pour un ou deux paramètres à la fois. Tous les résultats présentés dans cette section ont été obtenus en appliquant la méthodologie suivante :

- le réseau est adapté chaque mois avec les données des 24 mois précédents ;
- pour un mois donné, les valeurs testées sont systématiquement prises hors de l'ensemble d'apprentissage (simulation de l'application réelle) ;
- pour chaque variable, les biais sont exprimés en valeurs relatives par rapport à la variation annuelle (tab. 7.1).

### 7.4.1 Réponse de la carte pour une valeur faussée

La figure 7.9 montre les évolutions moyennes de la distance euclidienne  $d(\zeta(k), w_{i^*j})$  et de l'indicateur de confiance  $IC$  associés au neurone gagnant  $i^*$  durant une année. Les performances sont calculées pour l'ensemble de données original ainsi que pour cinq groupes de vecteurs

faussés. Ceux-ci sont composés des valeurs originales auxquelles a systématiquement été ajouté un biais  $Er$  respectivement de 5, 10, 15, 20 et 25 % sur les mesures de la température de l'huile  $T_{huile}$ .

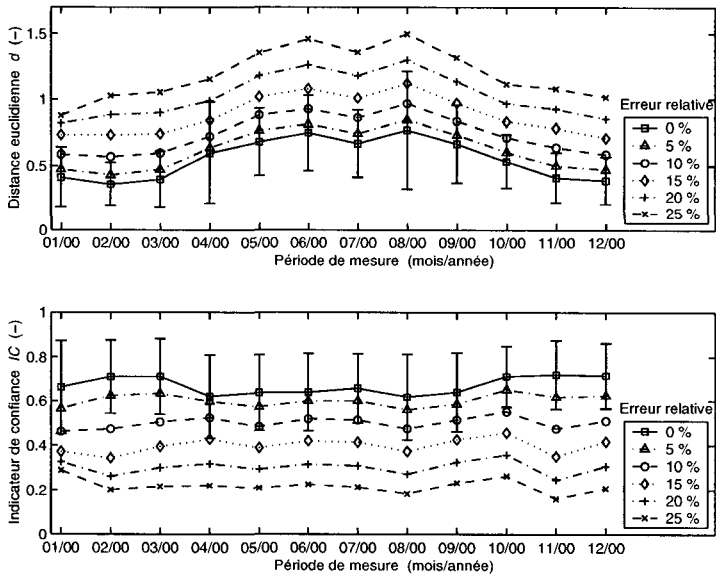


FIG. 7.9 – Variation des valeurs de l'indicateur de confiance  $IC$  et de la distance euclidienne  $d$  au cours d'une année de mesures en fonction de l'erreur relative ajoutée  $Er$  (biais) sur la mesure de la température de l'huile  $T_{huile}$ .

Chaque point représente la valeur moyenne de  $IC$ , respectivement  $d$  sur le mois considéré. Les écarts-types sont représentés pour les deux courbes obtenues avec les éléments de la base de données. Tout au long de l'année, la courbe de la distance euclidienne s'avère être moins stable que celle de l'indicateur de confiance  $IC$  : le réseau semble avoir modélisé avec plus de précision les états rencontrés en hiver que ceux rencontrés en été. Par conséquent, même sans défaut, il est normal que la distance euclidienne moyenne  $d(\zeta(k), w_{i^*j})$  augmente.

L'indicateur de confiance  $IC$  vient normaliser cette distance dans la mesure où il tient compte du rayon d'activation local. En terme de dé-

tection d'erreur, le cas considéré à la figure 7.9 confirme que la sensibilité du système est meilleure pour l'utilisation du signal  $IC$  que pour celle de la distance euclidienne. En définissant le seuil de détection de sorte à ce que la mesure ne soit pas comprise dans l'intervalle maximal défini par les couples  $\{\mu_i; \sigma_i\}$ , il peut être fixé aux environs de 1.2 pour la distance euclidienne et 0.4 pour  $IC$ .

Dans ces conditions, en été la sensibilité du système est supérieure de 5 % pour  $IC$  par rapport à  $d(\zeta(k), w_{i \cdot j})$ . Cependant, c'est en hiver que l'avantage de la méthode est le plus flagrant. Avec l'indicateur de confiance, la sensibilité est la même qu'en été. En revanche, l'utilisation de la distance euclidienne rend le système insensible à des décalages de la température de l'ordre de 25 % de la variation annuelle.

#### 7.4.2 Seuil de détection

La figure 7.10 montre le niveau de l'indicateur de confiance annuel moyen en fonction de l'erreur relative ajoutée  $Er$ . Les calculs ont été effectués d'après le même procédé que celui décrit ci-dessus ; les niveaux sont estimés pour chaque mois avec leur modèle respectif et sont ensuite moyennés. Dans ce cas, chacune des cinq variables a été prise en compte et est représentée par une courbe (saut simulé pour une variable à la fois).

TAB. 7.2 – *Sensibilité moyenne de détection en fonction de la valeur du seuil de l'indicateur de confiance IC.*

$IC$ (-)	$I_{ch}$ (%)	$T_{air}$ (%)	$T_{huile}$ (%)	$Gaz_{\Delta}$ (%)	$Hum_{\Delta}$ (%)
0.5	23.0	10.3	9.8	12.0	7.6
0.4	35.6	15.2	16.1	17.2	11.5
0.3	47.8	20.9	24.0	22.2	15.6
0.2	> 50	28.4	31.0	28.0	21.7

Les résultats montrent qu'à l'exception du courant de charge, la variation de l'indicateur de confiance en fonction de l'erreur relative ajoutée a la même allure pour toutes les variables prises en compte. En construisant l'histogramme des valeurs de chacune de ces grandeurs, il est facilement démontrable que  $T_{air}$ ,  $T_{huile}$ ,  $Gaz_{\Delta}$  et  $Hum_{\Delta}$  ont une

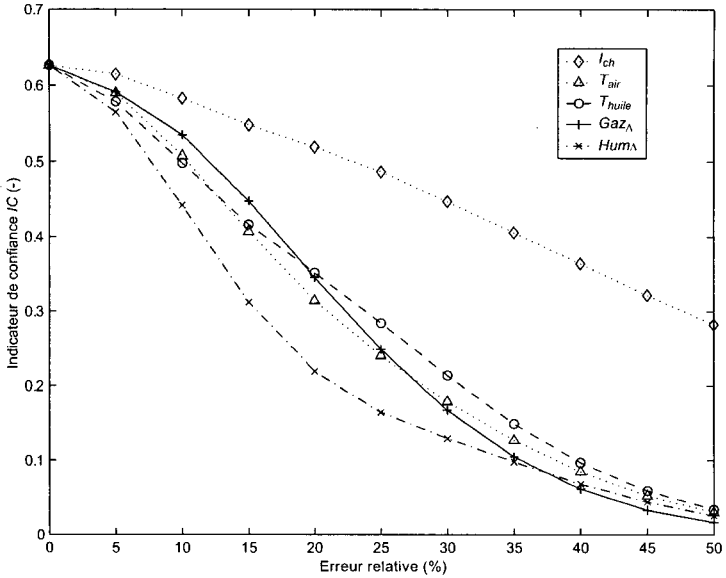


FIG. 7.10 – Valeur numérique de l'indicateur de confiance en fonction de l'erreur relative simulée (biais) pour les différentes variables. Écart-type de l'indicateur de confiance : 0.19 (valeur initiale sans erreur).

répartition statistique qui peut être assimilée à un phénomène gaussien. Dans la mesure où le vecteur d'entrée a été choisi de telle sorte que toutes ses valeurs aient une influence sur l'organisation, que les erreurs sont exprimées en fonction de la variation annuelle de ces variables et que l'apprentissage est réalisé à partir des valeurs normalisées, il n'est pas surprenant de trouver ce résultat.

Le tableau 7.2 regroupe sous forme numérique les principaux résultats qui peuvent être tirés de la figure 7.10 : les moyennes des erreurs ajoutées pour chaque variable ont été interpolées pour quatre niveaux d'indicateur de confiance définis. Ces valeurs peuvent être interprétées comme des mesures de sensibilité moyenne. L'écart-type de l'indicateur de confiance pour les données originales (non biaisées) est  $\sigma_{IC0} = 0.19$ . En fixant le seuil de détection à  $\mu_{IC0} \pm \sigma_{IC0}$  et en prenant une marge

de sécurité pour tenir compte de la variation annuelle (fig. 7.9), il paraît raisonnable de fixer le seuil de détection du système à  $IC = 0.4$ .

Les graphiques de la figure 7.11 montrent les résultats obtenus si un défaut ne se manifeste plus par l'augmentation d'une seule mais de deux variables. La procédure de test reste la même. La seule différence réside dans le fait que les cas sont calculés à partir d'un état de base tenant déjà compte d'un biais sur une variable. Trois décalages initiaux de 5, respectivement 10 et 15 % de la variation annuelle sont considérés pour cette première variable. Pour chacune de ces valeurs et pour les quatre autres mesures prises en compte, l'erreur (ou biais) qui doit être ajoutée pour atteindre le seuil de  $IC = 0.4$  est recherchée et représentée graphiquement. Parmi les cinq paramètres pris en compte par le modèle, les deux variables d'état  $I_{ch}$  et  $T_{air}$  n'ont pas été retenues en tant que premières variables.

Niveau limite de l'indicateur de confiance  
(seuil d'erreur)  $IC = 0.4$

Erreur relative simulée  
sur la première variable :

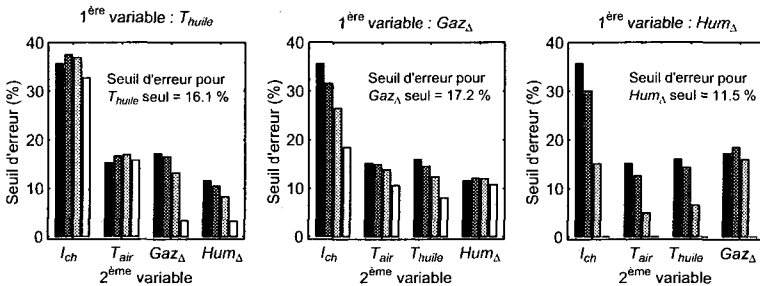


FIG. 7.11 – Seuil d'alarme pour deux variables faussées simultanément. Le niveau limite de l'indicateur de confiance est fixé à  $IC = 0.4$ .

Sur chacune des représentations de la figure 7.11, la barre noire représente le seuil de détection en fonction d'une erreur apparaissant exclusivement sur la deuxième variable (0 % sur la première variable). À quelques exceptions près, l'apparition d'erreurs simultanées sur deux variables améliore la sensibilité de détection. Il faut encore remarquer que



sur le troisième graphique, la barre blanche correspondant à une erreur  $Er$  de 15 % sur  $Hum_{\Delta}$  n'apparaît jamais. Cela est simplement dû au fait que le seuil d'erreur de la variable  $Hum_{\Delta}$  est de 11.5 %; en lui assignant un biais de 15 %, le système détectera une erreur, quel que soit le biais simulé sur la deuxième variable.

### 7.4.3 Influence du nombre de neurones

La détermination du nombre de neurones nécessaires pour modéliser le système a été faite en utilisant les mêmes tests que ceux décrits plus haut. La figure 7.12 illustre les résultats obtenus pour l'évolution de l'indicateur de confiance en fonction du nombre de neurones. Les valeurs tracées sont les moyennes annuelles calculées selon le principe utilisé jusqu'ici. Chaque courbe représente un niveau d'erreur relative distinct exprimé en pour cent de la variation annuelle de la variable  $T_{huile}$ .

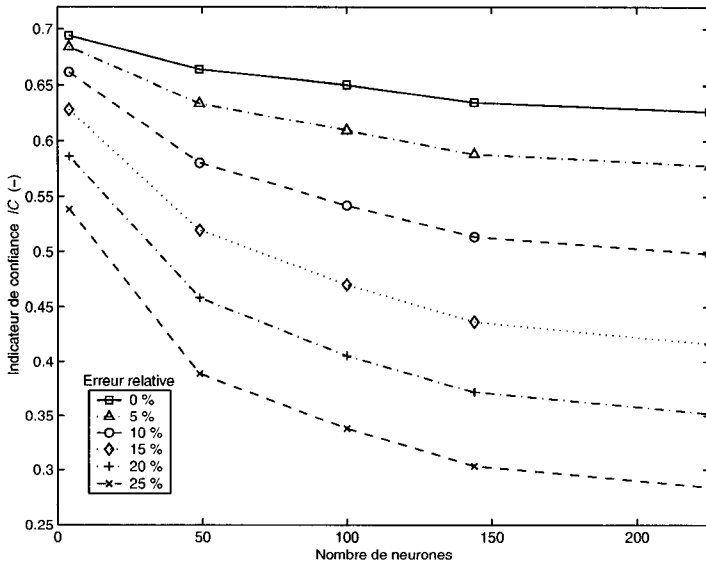


FIG. 7.12 – Influence du nombre de neurones de la carte auto-organisée sur l'indicateur de confiance. Valeurs typiques avec des erreurs simulées (biais) pour la température de l'huile  $T_{huile}$ .

L'analyse du graphique montre qu'avec l'augmentation du nombre de neurones, le niveau initial de l'indicateur de confiance a tendance à diminuer légèrement. Ceci s'explique par la subdivision toujours plus précise de l'espace d'entrée par des hypersphères de rayons toujours plus réduits. La précision locale augmente, mais au prix d'une légère perte de généralisation. Quoi qu'il en soit, le pouvoir de détection du réseau est caractérisé par la différence entre la valeur de l'indicateur *IC* pour le cas sans erreur et celui avec une erreur donnée. La figure 7.12 montre clairement que ce pouvoir augmente en fonction du nombre de neurones et semble se stabiliser autour de 225 neurones. La taille du réseau n'a pas été explorée au-delà de cette valeur : la quantité de calculs devient très importante et, compte tenu de la taille de la carte, le nombre d'éléments d'apprentissage exigé atteint les limites de la base de données à disposition.

## 7.5 Efficacité du système

### 7.5.1 Procédure de test

Les analyses faites jusqu'ici ont mis en évidence la sensibilité moyenne de la carte auto-organisée par rapport à l'identification de défauts se manifestant par l'augmentation du niveau d'un ou deux des paramètres pris en compte. Dans le contexte de l'implémentation du système en continu, le point essentiel est la gestion dynamique des alarmes. Dans ce sens, l'évaluation de la méthode doit être faite aussi bien en termes de sensibilité, de probabilité d'erreur de classification et de temps de détection.

Avant d'illustrer l'utilisation du système de surveillance intelligent par quelques cas, son pouvoir de détection automatique a été systématiquement testé. Ne disposant d'aucune donnée relatant un incident majeur pour l'unité pilote, les tests ont été effectués sur la base de simulations de défauts élémentaires.

Un seul cas est considéré : l'augmentation brusque du niveau d'une seule variable. À partir de l'ensemble de données réelles, les événements sont générés (ou simulés) et ajoutant durablement une valeur constante à l'une des variables enregistrées (fig. 7.13).

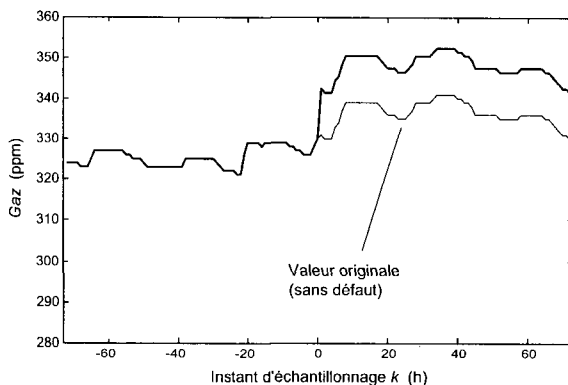


FIG. 7.13 – Principe de génération de défaut élémentaire : saut brusque à l'instant d'échantillonnage  $k = 0$  sur le taux de gaz ; évolutions de la valeur réelle (trait fin) et du défaut simulé (trait gras).

Les résultats du paragraphe précédent ont montré que par rapport à la modification simultanée de deux variables, cette approche pouvait être considérée comme un cas défavorable. Le deuxième type de défaut élémentaire pouvant être envisagé est la dérive rapide de certains paramètres. Celui-ci n'a pas été testé ici. Cependant, en mettant en relation les résultats obtenus pour le cas d'une augmentation brusque avec ceux attendus pour une augmentation graduelle, il est possible de faire une évaluation pessimiste de la sensibilité du système pour ce deuxième scénario.

Afin de prendre en compte le temps de détection, l'approche adoptée ici consiste à tirer aléatoirement un certain nombre de périodes de mesures dans la base de données, simuler un saut brusque sur l'une des variables à partir d'un instant précis et relever la réponse du système de détection. Ensuite, comme précédemment, le réseau est adapté chaque mois avec les données des 24 mois précédents (données originales). Les valeurs testées sont systématiquement obtenues à partir des données du premier mois suivant la fin de l'ensemble d'apprentissage. Pour chaque variable, les valeurs des sauts sont exprimées en pour cent de la variation annuelle de la mesure considérée (tab. 7.1).

Comme l'explique la méthodologie présentée au paragraphe 6.6.3, la valeur de  $IC$  n'est pas utilisée directement. Son écart-type moyen  $\sigma = 0.19$  obtenu pour un indicateur de confiance  $IC = 0.63$  dans le cas sans défaillance est trop grand pour que l'interprétation directe de  $IC$  donne des résultats suffisamment robustes. C'est pourquoi il est traité par l'algorithme cumulatif décrit au paragraphe 4.5.4. Les paramètres utilisés sont  $\mu_0 = 0.63$ ,  $\mu_1 = 0.4$  et  $\sigma = 0.2$ .

### 7.5.2 Erreur de classification

Le graphique de la figure 7.14 représente la probabilité d'erreur de classification moyenne en fonction du seuil d'alarme  $h$  fixé sur la sortie  $g_k$  (fonction de décision ou somme cumulée au sens de l'équation 4.38). Il est établi pour la détection d'un saut brusque simulé sur une variable.

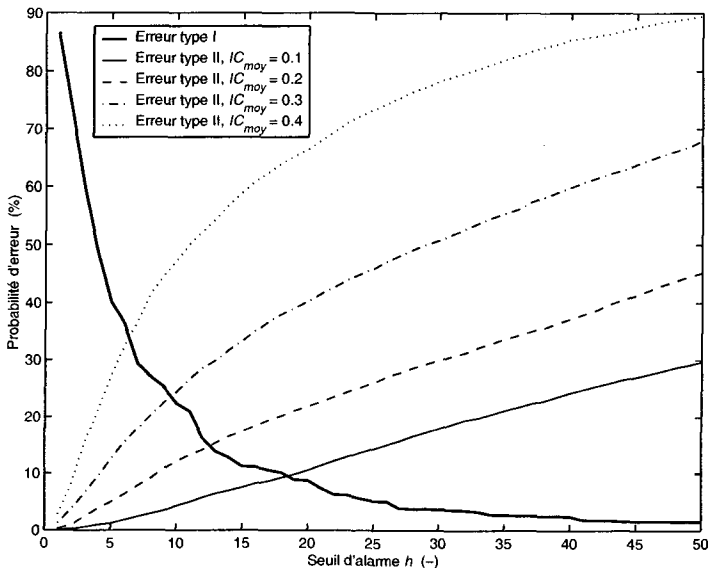


FIG. 7.14 – Performances moyennes du système en fonction du seuil d'alarme  $h$  : probabilité de fausses alarmes (erreur de type I) ; probabilité de ne pas détecter un problème (erreur de type II) pour des erreurs simulées pour différents niveaux de l'indicateur de confiance  $IC$ .

Chaque courbe de la figure 7.14 est le résultat de  $5 \times 4000$  simulations réalisées suivant la procédure de test décrite au paragraphe 7.5.1. La base de données utilisée est constituée des mesures collectées entre 1998 et 2000. Seules les 72 premières heures après l'apparition du défaut sont prises en compte et la variation des paramètres est considérée de manière globale; les regroupements sont faits en fonction de la valeur moyenne de l'indicateur de confiance. Par exemple, d'après le tableau 7.2, l'abaque  $IC_{moy} = 0.4$  représente des sauts de 35.6, 15.2, 16.1, 17.2 et 11.5 % respectivement sur les valeurs de  $I_{ch}$ ,  $T_{air}$ ,  $T_{huile}$ ,  $Gaz_{\Delta}$  et  $Hum_{\Delta}$  (valeurs relatives par rapport aux variations annuelles respectives).

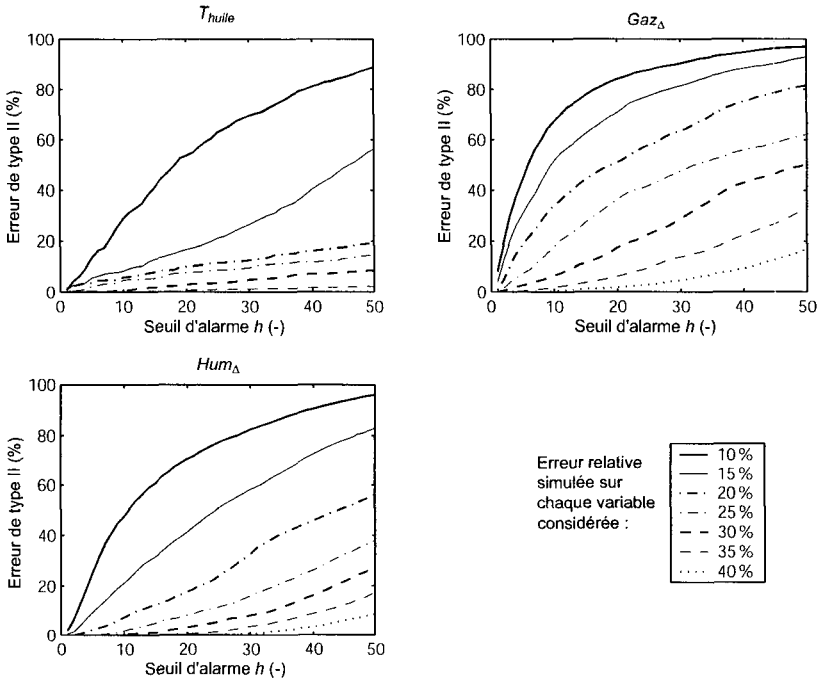


FIG. 7.15 – Probabilité de ne pas détecter un problème (erreur de type II) en fonction du seuil d'alarme  $h$ .

La figure 7.14 illustre bien le compromis qui doit être trouvé pour optimiser le rapport entre la probabilité de déclencher une fausse alarme (erreur de première espèce ou de type I) et celle de ne pas détecter un défaut (erreur de deuxième espèce ou de type II). Il met en évidence le fait que le seuil d'alarme ne devrait pas être inférieur à  $h = 15$  afin de limiter la probabilité d'erreur de type I à 10 %. L'analyse de la probabilité d'erreur de deuxième espèce peut être réalisée à l'aide des trois graphiques de la figure 7.15. Les mêmes données sont représentées sélectivement en fonction des trois paramètres de fonctionnement quantifiant directement l'état du transformateur. La lecture de ces graphiques montre que le système est beaucoup plus sensible à la détection de sauts sur la température de l'huile que sur le taux de gaz ou, dans une moindre mesure sur celui de l'humidité.

La figure 7.16 présente pour le cas de la température de l'huile une dernière alternative de représentation des données. Ce graphique met en évidence, quel que soit le seuil, l'aptitude du réseau à identifier des sauts de l'ordre de 30 % de la variation annuelle avec une probabilité d'erreur de type II inférieure à 10 %.

### 7.5.3 Rapidité de détection

La rapidité de détection n'a pas encore été considérée. Les résultats établis jusqu'ici ne sont pas discriminés suivant l'instant de détection du défaut. Pour autant que celui-ci soit identifié dans les 72 heures qui suivent son apparition, il est considéré comme détecté. La figure 7.17 permet de préciser la situation dans le cas de la température de l'huile. Le temps de détection est représenté pour la valeur médiane  $\Delta t_{50\%}$  et le quartile<sup>4</sup>  $\Delta t_{75\%}$  en fonction de l'amplitude du saut simulé. Il est clair que le seuil d'alarme influence le temps de détection. Pour des erreurs conséquentes (par exemple  $Er = 50\%$ ), la figure 7.17 montre que le temps de détection peut être supérieur à 10 heures. Ceci s'explique par la vitesse maximale de détection de la méthode qui dépend du seuil  $h$ . En effet, lorsqu'un problème est sévère,  $IC$  est nul et le logarithme du rapport de vraisemblance (4.28) est constant. Ainsi, le temps d'intégration nécessaire pour atteindre le seuil définit l'instant de détection.

---

4. Un quartile  $\hat{q}(p)$  est défini comme étant la valeur telle qu'une proportion de 25 ou 75 % des observations soit plus petite que  $\hat{q}(p)$  [101].

TAB. 7.3 – Performances du système de détection robuste pour l'identification de sauts brusques survenant indépendamment sur chacune des grandeurs surveillées ( $\mu_0 = 0.63$ ,  $\mu_1 = 0.4$ ,  $\sigma = 0.2$ ).

<i>Seuil h = 10</i>		<i>Erreur type I = 22.4 %, Erreur type II = 10.0 %</i>		
<b>Variable</b>	<b>Erreur relative (%)</b>	<b>Erreur absolue</b>	$\Delta t_{50\%}$ (h)	$\Delta t_{75\%}$ (h)
<i>I<sub>charge</sub></i>	43.8	469 A	9	22
<i>T<sub>air</sub></i>	16.0	8.4 °C	13	22
<i>T<sub>huile</sub></i>	14.5	5.8 °C	15	25
<i>Gaz<sub>Δ</sub></i>	28.5	10.5 ppm	5	8
<i>Hum<sub>Δ</sub></i>	19.0	0.54 %RH	6	11

<i>Seuil h = 15</i>		<i>Erreur type I = 11.2 %, Erreur type II = 10.0 %</i>		
<b>Variable</b>	<b>Erreur relative (%)</b>	<b>Erreur absolue</b>	$\Delta t_{50\%}$ (h)	$\Delta t_{75\%}$ (h)
<i>I<sub>charge</sub></i>	50.0	535 A	11	21
<i>T<sub>air</sub></i>	17.7	9.2 °C	17	27
<i>T<sub>huile</sub></i>	18.1	7.2 °C	16	23
<i>Gaz<sub>Δ</sub></i>	31.0	11.5 ppm	7	10
<i>Hum<sub>Δ</sub></i>	21.8	0.62 %RH	9	13

<i>Seuil h = 20</i>		<i>Erreur type I = 8.7 %, Erreur type II = 10.0 %</i>		
<b>Variable</b>	<b>Erreur relative (%)</b>	<b>Erreur absolue</b>	$\Delta t_{50\%}$ (h)	$\Delta t_{75\%}$ (h)
<i>I<sub>charge</sub></i>	> 50.0	> 535 A	-	-
<i>T<sub>air</sub></i>	18.8	9.8 °C	20	28
<i>T<sub>huile</sub></i>	20.0	8.0 °C	18	23
<i>Gaz<sub>Δ</sub></i>	33.3	12.5 ppm	8	11
<i>Hum<sub>Δ</sub></i>	24.0	0.68 %RH	11	16

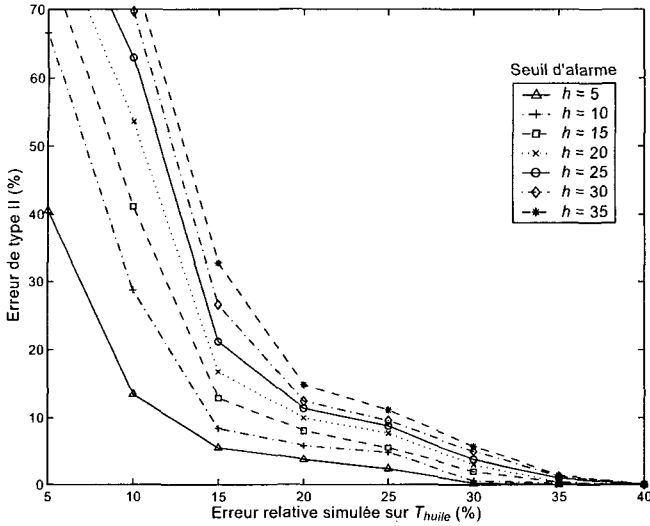


FIG. 7.16 – Probabilité de ne pas détecter un problème (erreur de type II) en fonction de l'erreur relative sur une variable : cas pour l'ajout d'un biais sur les valeurs de température de l'huile  $T_{huile}$ .

Le tableau 7.3 synthétise les résultats numériques pour trois seuils  $h$  différents et une erreur de type II fixée à 10 %. Tous les indicateurs qui viennent d'être discutés y figurent. Leur analyse tend à montrer que le seuil  $h = 15$  est le plus adapté au monitoring du transformateur pilote ; c'est pour cette valeur que le meilleur compromis entre la sensibilité, le temps de détection et l'erreur de première espèce est obtenu. Compte tenu des prétraitements effectués, le seuil de détection absolu pour les gaz dissous est très faible par rapport à sa valeur moyenne (tab. 7.1). Pour cette variable, ainsi que pour le taux d'humidité, le prétraitement appliqué (suppression de la moyenne mobile sur 72 heures) élimine progressivement l'effet d'un défaut caractérisé par l'augmentation soudaine d'une de ces valeurs. Compte tenu du seuil de détection déterminé au paragraphe 7.4.2, le système sera le plus efficace pour les instants proches de l'apparition du défaut. Ceci explique pourquoi les défaillances sur ces deux valeurs sont identifiées plus rapidement que les autres. Les histo-



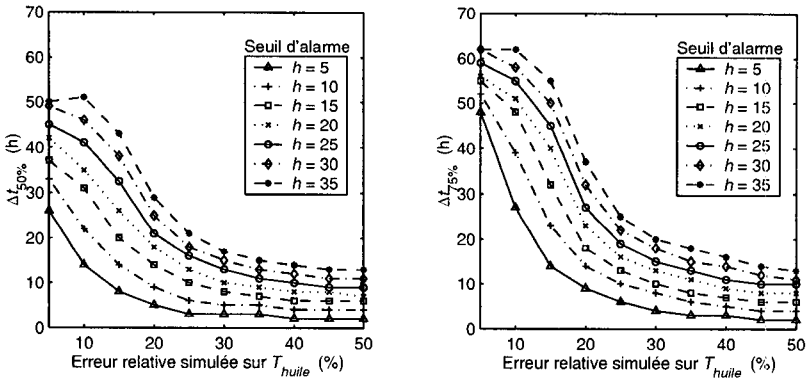


FIG. 7.17 – Rapidité de détection d'un problème : valeur médiane  $\Delta t_{50\%}$  (à gauche) et quartile  $\Delta t_{75\%}$  (à droite) du temps de détection en fonction de l'erreur relative sur la température de l'huile  $T_{huile}$  (saut brusque).

grammes de la figure 7.18 viennent illustrer ces propos. La répartition du temps de détection pour une erreur simulée sur la température de l'huile est plus large que celle observée sur le taux de gaz. Cette dernière est très étroite et répartie principalement autour de  $\Delta t = 6$  heures. Compte tenu du seuil fixé, cette valeur correspond par ailleurs à la vitesse de détection maximale du système.

Dans un contexte plus large où l'évolution d'un défaut n'est pas instantanée, seuls les événements ayant une évolution rapide peuvent être pris en compte pour ces deux paramètres (taux de gaz et d'humidité). À nouveau, l'effet de la moyenne élimine progressivement l'influence du défaut. Si la dérive est linéaire, après 72 heures, seule l'augmentation survenue durant cette période sera prise en compte. Dans ce cas, la détection n'est pas limitée aux 72 premières heures. Cependant, bien qu'en valeur absolue l'amplitude du défaut augmente toujours, le système ne pourra pas la prendre en compte. Quelle que soit la durée de la période observée, la limite de détection reste constante. Pour les variables  $Gaz_{\Delta}$  et  $Hum_{\Delta}$ , le seuil de détection pour les dérives linéaires est obtenu en divisant la limite établie pour une augmentation brusque par la période sur laquelle la moyenne mobile est calculée.

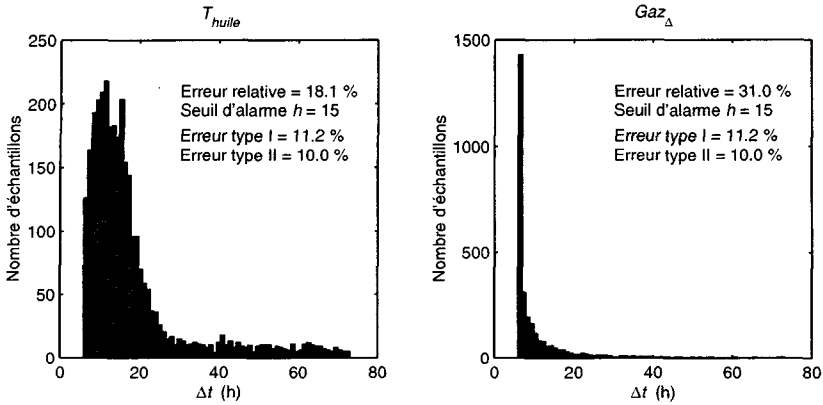


FIG. 7.18 – Histogrammes représentant la répartition des temps de détection  $\Delta t$  pour des sauts brusques générés sur les variables  $T_{huile}$  et  $Gaz_{\Delta}$ .

## 7.6 Études de cas

La mise en œuvre de la méthode de surveillance, tenant compte à la fois du signal d'alarme adaptatif étudié jusqu'ici et de l'interface graphique, est illustrée ci-après pour trois cas différents. Ceux-ci sont obtenus d'après la procédure de simulation décrite au paragraphe 7.5.1. Les défauts considérés sont les suivants : deux types d'apparitions soudaines de décalages (cf. fig. 7.13 pour une courbe typique) et une dérive rapide sur une variable. L'allure typique d'un défaut simulé dans ce dernier cas est représentée à la figure 7.19.

### 7.6.1 Apparition soudaine d'un décalage

Les figures 7.20 et 7.21 illustrent la réponse du système de monitoring développé pour un défaut se manifestant par l'augmentation soudaine de la température de l'huile  $T_{huile}$ . Ce saut est simulé sur la base des valeurs réellement mesurées sur le transformateur pilote de Fiesch. À partir du 22 octobre 1999, un incrément de 25 % de la variation annuelle moyenne, soit 10 °C, est ajouté sur la mesure de la température de l'huile.

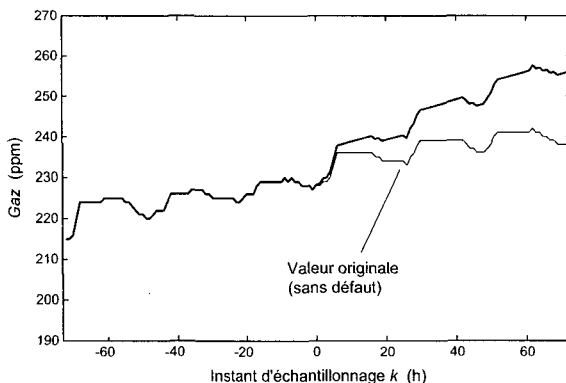


FIG. 7.19 – Principe de génération de défaut élémentaire : dérive du taux de gaz apparaissant à l'instant d'échantillonnage  $k = 0$  ; évolutions de la valeur réelle (trait fin) et du défaut simulé (trait gras).

L'instant d'échantillonnage correspondant à l'apparition du défaut définit l'origine des abscisses des graphiques de la figure 7.20. Dès cet instant, l'allure de l'indicateur de confiance  $IC$  est modifiée. La valeur moyenne passe d'environ 0.6 à 0.3. Parallèlement, la fonction de décision, nulle pour  $k = 0$ , croît. Le seuil d'alarme est finalement dépassé 8 heures après l'apparition du défaut. Pour un seuil d'alarme fixé sur la température de l'huile, le mode de fonctionnement estival impose un seuil au-delà de  $65\text{ }^{\circ}\text{C}$ . Or, dans le cas présent, même après l'apparition du défaut, on peut montrer que la température de l'huile n'atteint jamais  $60\text{ }^{\circ}\text{C}$ . Par conséquent, l'approche classique n'aurait pas permis de déceler ce problème.

Parallèlement au signal d'alarme, l'interface graphique permet d'interpréter qualitativement les résultats (fig. 7.21). Le fond coloré représente l'organisation de la couche correspondant à la température de l'huile  $T_{\text{huile}}$ . L'emplacement des neurones est repéré par les carrés blancs. Conformément à la technique décrite au paragraphe 6.4.4, les dimensions de ces carrés dépendent du taux d'activation de chacun des neurones pour l'ensemble des dix jours précédant les mesures courantes (à l'exception des 72 dernières heures). L'histogramme est réactualisé

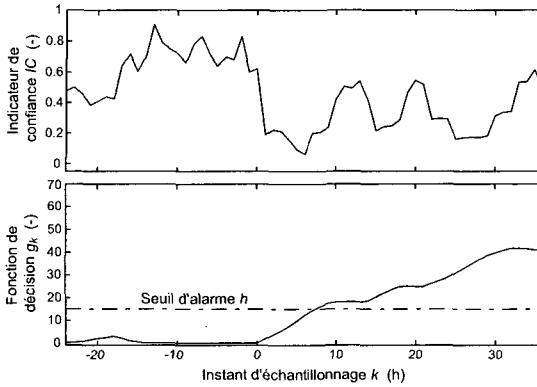


FIG. 7.20 – Valeurs de l'indicateur de confiance et du signal d'alarme pour le cas d'une augmentation soudaine de 25 % de la température de l'huile.

chaque heure afin de prendre en compte l'évolution du transformateur. Afin de ne pas biaiser l'histogramme avec des valeurs erronées, les mesures des 72 dernières heures ne sont jamais prises en compte.

Dans le cas présent, l'histogramme d'activation indique que le transformateur est caractérisé par une température située au milieu de la plage modélisée. L'analyse des autres couches montre une organisation générale relativement similaire à celle décrite au paragraphe 7.3.2. Il est ainsi possible d'affirmer qu'avant le défaut, le transformateur est moyennement chargé. Il se trouve en fait dans la période de transition entre les modes de fonctionnement estival et hivernal. La trajectoire du point de fonctionnement est représentée par une ligne dégradée. Le blanc représente les états initiaux observés 12 heures avant le défaut et le noir l'état du transformateur 36 heures après l'apparition du problème. L'histogramme représenté à la figure 7.21 correspond à celui établi pour ce dernier instant d'échantillonnage.

La trajectoire montre clairement une évolution de l'état du transformateur de la zone repérée par les carrés les plus gros vers une région de la carte où la température de l'huile est élevée. Par conséquent, sur cette base et compte tenu de la valeur élevée de la fonction de décision, il est

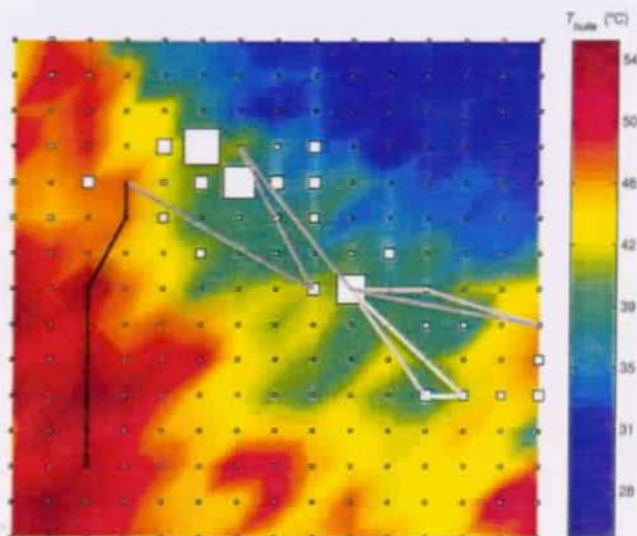


FIG. 7.21 – Représentation graphique du cas où la température de l'huile augmente soudainement. La trajectoire correspond à l'évolution du transformateur de 12 h avant (blanc) à 36 h après l'apparition du problème (noir).

non seulement possible déceler le fait qu'il y a un défaut mais aussi la nature de celui-ci.

L'identification du même type de problème (augmentation soudaine d'une valeur) simulé pour le taux de gaz dissous est représenté par les figures 7.22 et 7.23. À partir des valeurs originales, le taux de gaz subit une augmentation rapide de 11.4 ppm (31% de la variation annuelle) le 22 juillet 1999. L'histogramme d'activation montre que le transformateur est chargé et qu'il se situe dans une zone où le taux de gaz varie fortement. L'effet du retranchement de la moyenne mobile discuté au paragraphe 7.5.3 apparaît clairement sur l'évolution de la valeur de  $IC$  (fig. 7.22). Après une chute rapide de l'indicateur de confiance  $IC$  et une stabilisation de sa valeur autour de zéro pendant 20 heures, celle-ci commence à croître. Quoi qu'il en soit, lorsque ce phénomène apparaît,

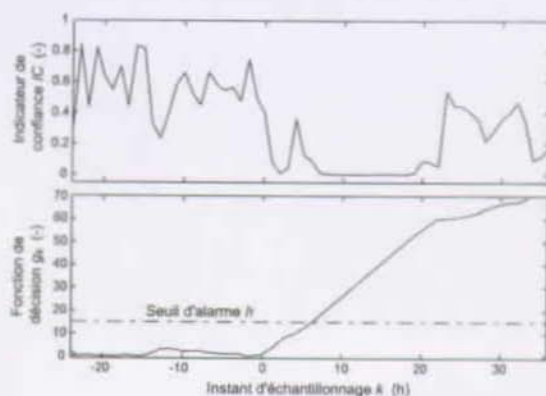


FIG. 7.22 – Valeurs de l'indicateur de confiance et du signal d'alarme pour le cas d'une augmentation soudaine de 31 % du taux de gaz dissous.

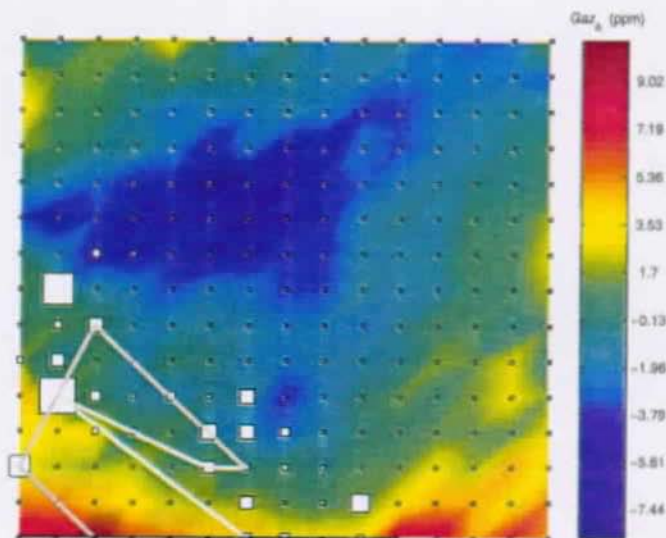


FIG. 7.23 – Représentation graphique du cas où le taux de gaz augmente soudainement. La trajectoire correspond à l'évolution du transformateur de 12 h avant (blanc) à 36 h après l'apparition du problème (noir).

l'alarme est déjà déclenchée (le seuil  $h$  est dépassé six heures après l'apparition du défaut).

Pour l'interprétation des résultats, la carte est représentée en fonction de l'organisation de la couche correspondant au taux de gaz (fig. 7.23). Comme pour le cas précédent, l'évolution de la trajectoire montre le déplacement du point de fonctionnement d'une zone où le taux de gaz est relativement faible vers une région caractérisée par un fort taux de gaz.

### 7.6.2 Dérive rapide d'une variable

Le dernier cas considéré concerne une défaillance qui se manifesterait par l'augmentation progressive d'une variable. Pour illustrer ce type d'événement, une dérive de 0.62 %RH par jour (22 % de la variation annuelle de  $Hum_{\Delta}$ ) a été ajoutée aux mesures d'humidité relative collectées à partir du 23 janvier 2000.

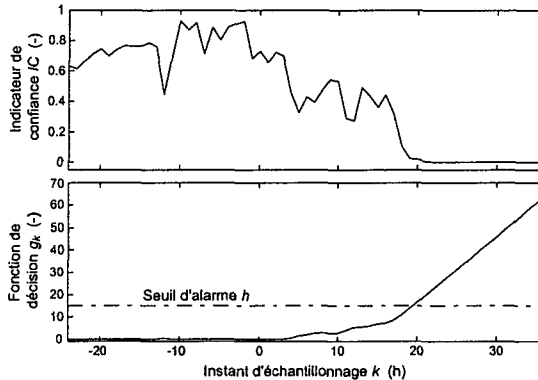


FIG. 7.24 – Valeurs de l'indicateur de confiance et du signal d'alarme pour le cas d'une augmentation progressive du taux d'humidité, dérive de 0.62 %RH.

La figure 7.24 montre les évolutions respectives de l'indicateur de confiance  $IC$  et de la fonction de décision  $g_k$ . L'effet progressif du défaut est bien visible. Contrairement aux cas précédents, la valeur moyenne de l'indicateur de confiance ne chute pas d'un coup. Le même phéno-

mène est observable pour la fonction de décision : l'augmentation se fait d'abord avec une pente faible qui s'infléchit nettement 15 heures après le début du défaut. Le seuil d'alarme est franchi 20 heures après l'apparition du problème. Dans ce cas, la tendance journalière simulée est suffisamment importante pour que le prétraitement des mesures d'humidité n'affecte pas la qualité de la détection.

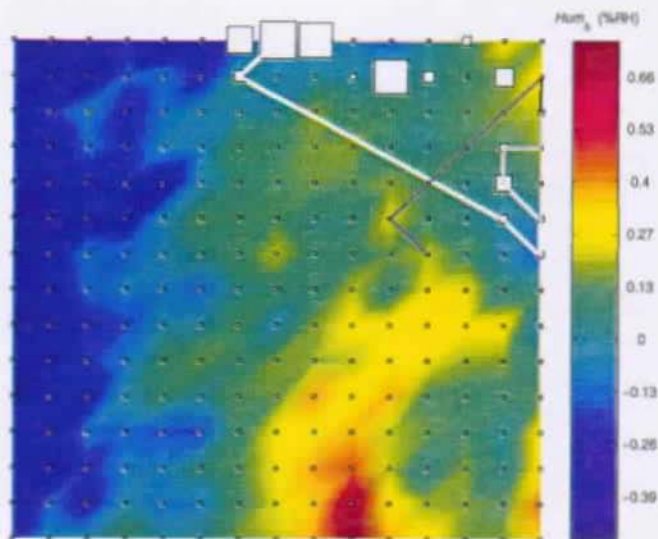


FIG. 7.25 – Représentation graphique du cas où le taux d'humidité relative augmente progressivement (dérive). La trajectoire correspond à l'évolution du transformateur de 12 h avant (blanc) à 36 h après l'apparition du problème (noir).

À nouveau, la représentation graphique (fig. 7.25) permet de mettre qualitativement en évidence le fait que le taux d'humidité est responsable de l'alarme. Alors qu'avant l'apparition du défaut, l'historique montre que le transformateur fonctionne selon le mode hivernal typique (coin supérieur droit de la carte), la trajectoire du point de fonctionnement se déplace vers la zone où l'humidité est élevée dès que le défaut devient important.



## 7.7 Discussion

Les cas qui viennent d'être présentés valident l'approche développée au chapitre 6 de manière pratique. Les différentes défaillances traitées sont des cas très simples. Néanmoins, elles sont basées sur des mesures réelles, sélectionnées de manière à représenter les différents modes de fonctionnement du transformateur modélisé (été, hiver et période de transition). En pratique, l'interdépendance entre les variables se traduit certainement par des situations plus complexes que celles considérées ici. Par contre, les variations statistiques dues entre autres à la précision des instruments de mesures et aux variations météorologiques ( $T_{air}$ ) sont implicitement prises en compte par la modélisation des cartes.

Le temps limite de détection de 6 heures déterminé pour le seuil  $h = 15$  peut s'avérer dans certains cas beaucoup trop long. On a vu que ce temps dépend uniquement du niveau du seuil. Afin d'éviter les fausses alarmes, celui-ci ne devrait pas être baissé. En revanche, si ce temps est trop long, c'est généralement que le problème est sévère. Dans ce cas, l'indicateur de confiance doit normalement tomber à des valeurs proches de zéro. Il est alors possible d'améliorer la sélectivité du traitement par les sommes cumulées en définissant une deuxième fonction de décision. Dans ce cas, la valeur de la moyenne après changement pourrait être définie par  $\mu_1 = 0$ . La probabilité d'obtenir un indicateur de confiance nul alors que l'état est sain est quasiment nulle ; le seuil  $h$  fixé sur cette variable pourrait ainsi être faible et influencer positivement le temps de détection. À la limite, un seuil très proche de zéro pourrait être directement défini sur  $IC$ .

Tous les résultats présentés ici ont été obtenus à partir d'une base de données constituée depuis décembre 1997. Les calculs ont été réalisés "off-line". En revanche, pour autant que l'acquisition des mesures, la gestion de la base de données et le prétraitement des valeurs manquantes ou aberrantes soient en place, la technique développée peut être mise en œuvre telle quelle en continu.



# Chapitre 8

## Conclusion

Le travail développé dans cette thèse apporte une contribution originale au problème de surveillance de processus en général. La méthodologie et les résultats présentés s'appuient sur un cas concret : l'interprétation des données provenant d'un système de surveillance de transformateur. La connaissance des paramètres de fonctionnement de ces derniers aide à la compréhension et à l'interprétation des résultats. Cependant, l'approche systématique développée pour la mise en œuvre des algorithmes est indépendante de ces considérations. La méthode développée peut donc être étendue à tout problème de surveillance dont le cahier des charges est :

- nécessité de détecter automatiquement les défaillances d'un processus industriel ou chimique sur la base d'informations multivariées délivrées en continu par un système d'acquisition ;
- machine ou processus complexe dont la modélisation mathématique n'est pas envisageable ;
- environnement tel que l'évolution des paramètres en cas de défaut n'est pas connu.

### 8.1 Résultats obtenus

Après avoir discuté de l'ensemble des techniques de mesures relatives à l'évaluation de l'état des transformateurs et dressé l'état de l'art dans le domaine de la surveillance de processus, l'application des cartes auto-organisées au domaine a été discutée. En analysant les propriétés de ces

cartes, un algorithme d'apprentissage modifié par rapport à l'algorithme original de Kohonen a pu être sélectionné. Éliminant le problème de neurones "morts" en introduisant la notion de zone d'activation, les cartes organisées selon ces nouvelles règles (kMER) conservent les propriétés de représentation topologique de l'algorithme de Kohonen.

Le cadre général d'application des cartes auto-organisées à la surveillance des transformateurs a ensuite été défini. Compte tenu des contraintes citées plus haut, l'objectif était de pouvoir identifier automatiquement toute excursion du transformateur hors de l'ensemble d'états connus et de pouvoir l'expliquer qualitativement. La mise en œuvre des cartes dans ce contexte s'est matérialisée par deux sorties distinctes qui sont : une interface graphique servant de support à l'interprétation qualitative des résultats et un indicateur de confiance permettant de déclencher automatiquement une alarme. La méthodologie générale d'interprétation des résultats représente la première contribution originale de cette thèse. L'interprétation qualitative des résultats à l'aide de l'histogramme d'activation en constitue l'un des points essentiels.

La seconde contribution originale est l'étude et le développement de l'indicateur de confiance qui vient d'être cité. L'exploitation des paramètres liés à la structure des cartes kMER a permis de tirer l'expression d'un indicateur normalisé évaluant, selon des critères statistiques, la qualité de la réponse du réseau à une sollicitation. Cette valeur est robuste dans le sens où elle tient compte de l'organisation du réseau dans la région du neurone gagnant. L'interprétation de cet indicateur de confiance à l'aide d'un seuil fournit finalement un signal d'erreur qui prend en compte le mode de fonctionnement du transformateur et les interactions entre les différents paramètres mesurés.

Le développement de la technique de réapprentissage contrôlé des cartes constitue une autre contribution. Il a pu être démontré qu'en partant du même état initial partiellement organisé, l'état final pouvait représenter parfaitement un ensemble de données, tout en conservant la même topologie générale entre deux apprentissages successifs.

Finalement, la mise en œuvre de ces techniques pour les données d'un transformateur réel a permis de démontrer l'aptitude de la méthode à détecter des états faussés (situés hors de la zone de fonctionnement normal). Sur la base de tests faits pour une seule variable faussée à la

fois, l'interprétation graphique de la carte montre qu'il est possible de déterminer quelle est cette variable en un coup d'œil. La carte en tant que principal interface graphique permet de représenter l'évolution de la situation courante de manière synthétique. Plus lisible que le tracé des évolutions temporelles de chaque variable, elle permet de caractériser immédiatement un point de fonctionnement par rapport à l'ensemble des états possibles.

La méthode proposée n'a été testée en toute rigueur que pour l'évolution anormale d'une seule variable. Ceci peut résulter, dans la plupart des cas, de la défaillance d'un capteur. Cependant, il n'est pas possible de l'affirmer a priori. Des connaissances plus larges des phénomènes modélisés sont nécessaires pour pouvoir se prononcer. En revanche, on peut affirmer que si le système permet de détecter de faibles décalages introduits sur une seule variable, il aura d'autant moins de peine, dans le cas d'une défaillance réelle, à déceler un état se caractérisant par la modification de plusieurs paramètres.

## 8.2 Perspectives d'avenir

Malgré le manque de connaissances du système original, la méthode développée donne de bons résultats. Elle améliore nettement l'approche qui consiste à fixer des seuils fixes sur chacune des variables surveillées. Le système décrit est limité à la prise en compte de l'évolution du transformateur sur un horizon relativement court (72 heures). Ceci n'est en aucun cas une limitation introduite par les cartes. L'incertitude statistique sur les variations annuelles motive la décision qui a été prise de ne pas considérer les variations sur une période plus longue.

L'amélioration de la qualité de l'outil passe inévitablement par une meilleure connaissance de l'évolution des paramètres de fonctionnement. La détermination des valeurs de fonctionnement anormal est également nécessaire pour pouvoir entraîner et tester le réseau plus efficacement. Dans ces conditions, l'utilisation des cartes auto-organisées pour la surveillance des transformateurs de puissance a le potentiel de fournir des résultats encore plus riches que ceux décrits dans cette thèse. En tenant compte des états défaillants dans l'apprentissage de la carte, celle-ci contiendra des zones clairement délimitées qui pourront être reconnues

et conférer au réseau de neurones des qualités de diagnostic. Connaissant mieux les défauts auxquels on doit faire face, il sera également possible d'étudier la discrimination des problèmes en fonction de leur nature (défaut réel ou capteur défaillant).

Ces objectifs ne pourront être atteints qu'en poursuivant les recherches suivant trois axes principaux :

- l'étude approfondie des phénomènes physiques afin de pouvoir établir des modèles permettant d'augmenter la base de données d'apprentissage ;
- l'établissement de modèles statistiques permettant de fournir des entrées supplémentaires (défauts essentiellement) ;
- la multiplication des sites de mesures afin d'augmenter les connaissances relatives aux paramètres opérationnels des transformateurs.

La mise en œuvre de ces techniques devrait permettre d'étendre le domaine d'utilisation du système actuel à l'identification de phénomènes se développant plus lentement et d'introduire le pouvoir prédictif qui lui fait encore partiellement défaut.

# Bibliographie

- [1] E. Alhoniemi, J. Hollmén, O. Simula, J. Vesanto. Process monitoring and modeling using the self-organizing map. *Integrated Computer-Aided Engineering*, 6(1):3–14, 1999.
- [2] Th. Aschwanden, M. Hässig, J. Fuhr, P. Lorin, V. Der Houhanessian, W. Zaengl, A. Schenk, P. Zweiacker, A. Piras. Development and application of new condition assessment methods for power transformers. *37th CIGRÉ session reports*, September 1998. Paper 12–207.
- [3] J. Aubin. Effect of particles on transformer dielectric strength. *Electra*, 190:134–140, 2000.
- [4] N. Ball, L. Kierman, K. Warwick, E Cahill, D. Esp, J. Macqueen. Neural networks for power systems alarm handling. *Neurocomputing*, 4(1-2):5–8, 1992.
- [5] P. Baroni, G. Guida, S. Mussi. State assessment and preventive diagnosis of power transformers: A knowledge-based approach. *Engineering Intelligent Systems*, 5(2):91–105, 1997.
- [6] M. Basseville. Detecting changes in signals and systems – A survey. *Automatica*, 24(3):309–326, 1988.
- [7] M. Basseville, I.V. Nikiforov. *Detection of Abrupt Changes*. PTR Prentice-Hall, 1993.
- [8] H.-U. Bauer, K.R. Pawelzik. Quantifying the neighborhood preservation of self-organizing feature maps. *IEEE Transactions on Neural Networks*, 3(4):570–579, 1992.
- [9] T. Baumann, A. Germond. Application of the Kohonen network to short-term load forecasting. *Proceedings of the Second Internatio-*

*nal Forum on Applications of Neural Networks to Power Systems ANNPS'93*, pages 407–412, 1993.

- [10] T. Baumann, A. Germond, D. Tschudi. Impulse test fault diagnosis on power transformer using Kohonen's self-organizing neural network. *Symposium on Expert System Application to Power Systems*, pages 324–329, 1991.
- [11] G. Belanger, M. Duval. Monitor for hydrogen dissolved in transformer oil. *IEEE Transactions on Electrical Insulation*, EI-12(5):334–340, 1977.
- [12] C. Bengtsson. Status and trends in transformer monitoring. *IEEE Transactions on Power Delivery*, 11(3):1379–1384, July 1996.
- [13] T. Bengtsson, H. Kols, L. Martinsson, B.O. Stenestam, M. Foata, F. Léonard, C. Rajotte, J. Aubin. Acoustic diagnosis of tap changers. *36th CIGRÉ session reports*, 1996. Paper 12–101.
- [14] J. Binggeli, J. Froidevaux. Surveillance en service de l'état de l'isolement des transformateurs. *Bulletin Sécheron*, 5, 1967.
- [15] C.M. Bishop, M. Svensen, C.K.I Williams. Developments of the generative topographic mapping. *Neurocomputing*, 21:203–224, 1998.
- [16] A. Bognar, L. Kalocsai, G. Csepes, E. Nemeth, J. Schmidt. Diagnostic tests of high voltage oil-paper insulating systems (in particular transformer insulation) using DC dielectrometrics. *33th CIGRÉ session reports*, 1990. Paper 15/33–08.
- [17] C. Boisseau, T. Fortin, V. Hilario. Instrument transformers partial discharges measurements as a monitoring method. *8th International Symposium on High Voltage Engineering*, pages 21–24, 1993. 61.01.
- [18] C. Booth, J.R. McDonald. The use of artificial neural networks for condition monitoring of electrical power transformers. *Neurocomputing*, 23:97–109, 1998.
- [19] P. Boss, P. Lorin, A. Viscardi, J.W. Harley, J. Isecke. Economical aspects and practical experiences of power transformer on-line monitoring. *38th CIGRÉ session reports*, September 2000. Paper 12–202.
- [20] P. Boss, J.F. Ravot, P. Mathis, H. Braendle. Mesure du profil de température dans les transformateurs à l'aide d'un capteur dis-



- tribué à fibre optique. *CIGRÉ Symposium Berlin*, 1993. Article 110-07.
- [21] D. Chu, A. Lux. On-line monitoring of power transformers and components: A review of key parameters. *Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Conference*, pages 669-675, 1999. IEEE.
- [22] M. Cottrell, J.C. Fort. Étude d'un algorithme d'auto-organisation. *Annales de l'Institut Henri Poincaré*, 23(1):1-20, 1987.
- [23] M. Cottrell, J.C. Fort, G. Pagès. Theoretical aspects of the SOM algorithm. *Proceedings of WSOM'97*, pages 246-267, 1997.
- [24] M. Cottrell, B. Girard, P. Rousset. Long term forecasting by combining Kohonen algorithm and standard prevision. *International Conference on Artificial Neural Networks ICANN'97*, pages 993-998, 1997.
- [25] T.H. Crowley. *Automated Diagnosis of Large Power Transformers Using Adaptive Model-Based Monitoring*. Massachusetts Institute of Technology, LEES Technical Report, 1990.
- [26] D. Dahlgren, G. Rosenwald, C.C. Liu, S. Muchlinski, A. Eide, D. Sobajic. Model-based synthesis and suppression of transformer alarms in a control center environment. *IEEE Transactions on Power Delivery*, 13(3):843-848, 1998.
- [27] F. de Viron, M. De Vlaminck, T. Minns, Th. Weber. A feasibility study for nuclear safety improvement: A neural network to perform early warning. *Expert System Applications for the Electric Power Industry: International Conference and Exhibition*, 1993. EPRI.
- [28] P. Demartines. *Analyse de données par réseaux de neurones auto-organisés*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1994.
- [29] V. Demjanenko, R.A. Valtin, M. Soumekh, H. Naidu, A. Antur, D.P. Hess and A. Soom, M.K. Tangri, S.Y. Park, D.M. Benenson, S.E. Wright. A noninvasive diagnostic instrument for power circuit breakers. *IEEE Transactions on Power Delivery*, 7(2):656-663, 1992.
- [30] D. DeSieno. Adding a conscience to competitive learning. *IEEE International Conference on Neural Networks*, 1:117-124, 1988.

- [31] E.P. Dick, C.C. Erven. Transformer diagnostic testing by frequency response analysis. *IEEE Transactions on Power Apparatus and Systems*, PAS-97(6):2144–2153, Nov/Dec 1978.
- [32] H.E. Dijk, N.V. Kema. Exformer, an expert system for transformer faults diagnosis. *Proceedings of the 9th Power Systems Computation Conference (PSCC) 1987*, pages 715–721, 1987.
- [33] Y. Du, M. Zahn, B.C. Lesieutre, A.V. Mamishev, S.R. Lindgren. Moisture equilibrium in transformer paper-oil systems. *IEEE Electrical Insulation Magazine*, 15(1):11–20, 1999.
- [34] B. Dubuisson. *Diagnostic et reconnaissance des formes*. Hermès, 1990.
- [35] M. Duval. Dissolved gas analysis – It can save your transformer. *IEEE Electrical Insulation Magazine*, 5(6):22–27, 1989.
- [36] P.M. Eleftherion. Partial discharge XXI: Acoustic emission-based PD source location in transformers. *IEEE Electrical Insulation Magazine*, 11(6):22–26, 1995.
- [37] E. Erwin, K. Obermayer, K. Schulten. Self-organizing maps: Ordering, convergence properties and energy functions. *Biological Cybernetics*, 67:47–55, 1992.
- [38] E. Erwin, K. Obermayer, K. Schulten. Self-organizing maps: Stationary states, metastability and convergence rates. *Biological Cybernetics*, 67:35–45, 1992.
- [39] D.G. Esp, A.J. McGrail. Advances in data mining for dissolved gas analysis. *IEEE International Symposium on Electrical Insulation*, pages 456–459, 2000.
- [40] K. Feser, H.A. Maier, B. Feuchter, H. Freund, U. Rosenow, A. Baur, H. Mieske. On-line diagnostic system for monitoring the thermal behaviour of transformers. *CIGRÉ Symposium Berlin*, 1993. Paper 110–08.
- [41] B. Feuchter, K. Feser. On-line diagnostic of the thermal behaviour of power transformers. *8th International Symposium on High Voltage Engineering*, 1993. Paper 65.01.
- [42] P.M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – A survey and some new results. *Automatica*, 26(3):459–474, 1990.

- [43] J. Fuhr, M. Haessig, P. Boss, D. Tschudi, R.A. King. Detection and location of internal defects in the insulation of power transformers. *IEEE Transactions on Electrical Insulation*, 28(6):1057–1067, 1993.
- [44] H. Furukawa, T. Ueda, M. Kitamura. A systematic method for rational definition of plant diagnostic symptoms by self-organizing neural networks. *3rd International Conference on Fuzzy Logic, Neural Networks and Soft Computing*, pages 555–556, 1994.
- [45] C. Gandillon. *Analyse et diagnostic des gaz dissous dans l'huile isolante de transformateur*. BBC-Sécheron SA, 1987.
- [46] A. Gersho, R. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Press, London, 1992.
- [47] J.J. Gertler. A survey of model-based failure detection and isolation in complex plants. *IEEE Control Systems Magazine*, 8:3–11, 1988.
- [48] P. Gervais, J. Aubin. On-line monitoring of key fault gases in power transformers and shunt reactors. *62nd Annual International Conference of Doble Clients*, 1995.
- [49] U. Gäfvert, L. Adeen, M. Tapper, P. Ghasemi, B. Jönsson. Dielectric spectroscopy in time and frequency domain applied to diagnostic of power transformers. *6th International Conference on Properties and Applications of Dielectric Materials*, 2000.
- [50] F. Girosi, T. Poggio. Networks and the best approximation property. *Biological Cybernetics*, 63:169–176, 1990.
- [51] M. Givelberg, S. Gurevich, N. Lazar, R. Zelichonok. An algorithm and software for distribution transformer operating age monitoring. *Nineteenth Convention of Electrical and Electronics Engineers in Israel*, 1996. IEEE.
- [52] P.V. Goosen. Transformers accessories. *36th CIGRÉ session reports*, 1996. Paper 12–104.
- [53] W.H. Hagman, M.J. Boyd. Model-based analysis of a failed transformer. *EPRI Substation Equipment Diagnostics Conference VII*, 1999.
- [54] W.H. Hagman, T.H. Crowley, R.D. Tabors, F.C. Schweppe. An adaptive transformer monitoring system. *International Sympos-*

*sium for Demonstrations of Expert System Applications to the Power Industry*, 1989.

- [55] A.C. Hall. Discussion on "Diagnostic methods for transformers in service". Part 3.5: "Surveillance using microprocessor systems". *Keynote Presentation, CIGRÉ SC12*, Septembre 1985.
- [56] E. Handschin, D. Kuhlmann, C. Rehtanz. Visualization and analysis of voltage stability using self-organizing neural networks. *Artificial Neural Networks ICANN'97*, pages 1113–1118, 1997.
- [57] T. Harris. A Kohonen S.O.M. based, machine health monitoring system which enables diagnosis of fault not seen in the training set. *International Joint Conference on Neural Networks IJCNN'93*, 1:947–950, 1993.
- [58] A.C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, 1990.
- [59] S. Haykin. *Neural Networks, A Comprehensive Foundation*. Prentice-Hall Inc., 2nd edition, 1999.
- [60] Q. He, J. Si, D.J. Tylavsky. Prediction of top-oil temperature for transformers using neural networks. *IEEE Transactions on Power Delivery*, 15(4):1205–1211, 2000.
- [61] M. Heathcote. *The J & P Transformer Book, A Practical Technology of the Power Transformer*. Newnes, 12th edition, 1998.
- [62] J. Hertz, A. Krogh, R.G. Palmer. *Introduction to the Theory of Neural Computation*. Addison-Wesley, 1991.
- [63] A.J. Höglund, K. Hätönen. Computer network user behaviour visualisation using self organizing maps. *International Conference on Artificial Neural Networks ICANN'98*, 2:899–904, 1998.
- [64] D. Hill, T.T. Le, M. Darveniza, T. Saha. A study of the degradation of cellulosic insulation materials in a power transformer. Part iii: Degradation products of cellulose insulation paper. *Polymer Degradation and Stability*, 51:211–218, 1996.
- [65] B. Hochard. *Le transformateur de puissance*. Technique et documentation (Lavoisier), Paris, 1988.
- [66] K. Hornik, M. Stinchcombe, H. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.

- [67] V. Der Houhanessian. *Measurement and Analysis of Dielectric Response in Oil-Paper Insulation Systems*. PhD thesis, Eidgenössische Technische Hochschule Zürich, 1998.
- [68] Y.-Y. Hsu, C.-C. Yang. Design of artificial neural networks for short-term load forecasting. Part I: Self-organizing feature maps for day type identification. *IEE Proceedings C (Generation, Transmission and Distribution)*, 138(5):407–413, 1991.
- [69] Y.-C. Huang, H.-T. Yang, C.-L. Huang. Developing a new transformer fault diagnosis system through evolutionary fuzzy logic. *IEEE Transactions on Power Delivery*, 12(2):761–767, 1997.
- [70] M.M. Van Hulle. Topology-preserving map formation achieved with a purely local unsupervised competitive learning rule. *Neural Networks*, 10(3):431–446, 1997.
- [71] M.M. Van Hulle. Kernel-based equiprobabilistic topographic map formation. *Neural Computation*, 10:1847–1871, 1998.
- [72] M.M. Van Hulle. Faithful representations with topographic maps. *Neural Networks*, 12(6):803–823, 1999.
- [73] M.M. Van Hulle. *Faithful Representations and Topographic Maps*. John Wiley & Sons Inc., 2000.
- [74] M.M. Van Hulle, D. Martinez. On a novel unsupervised competitive learning algorithm for scalar quantization. *IEEE Transactions on Neural Networks*, 5(3):498–501, 1994.
- [75] J.E. Jackson. *A User's Guide to Principal Components*. John Wiley & Sons Inc., 1991.
- [76] J. Jalbert, R. Gilbert, P. Tétreault. Determination of the analytical performance of a headspace capillary gas chromatographic technique and karl fischer coulometric titration by system calibration using oil samples containing known amounts of moisture. *Analytical Chemistry*, 71(15):3283–3291, 1999.
- [77] E.R. Kandel, J.H. Schwartz, T.M. Jessell. *Principles of Neural Science*. Prentice-Hall International Inc., 3rd edition, 1991.
- [78] P. Kang, D. Birtwhistle, D. McCulloch. Non-invasive on-line condition monitoring of on load tap changers. *IEEE Power Engineering Society Winter Meeting*, pages 2223–2228, 2000.
- [79] J. Kangas, S. Kaski. *3043 works that have been based on the Self-Organizing Map (SOM) Method developed by Kohonen*. Helsinki

University of Technology, Department of Computer Science and Engineering, Laboratory of Computer and Information Science, 1998. Report A49.

- [80] M. Kasslin, J. Kangas, O. Simula. Process state monitoring using self-organizing maps. *Artificial Neural Networks 2*, 2:1532–1534, 1992.
- [81] J.J. Kelly. Transformer fault diagnosis by dissolved-gas analysis. *IEEE Transactions on Industry Applications*, IA-16(6):777–782, November/December 1980.
- [82] J.L. Kirtley, W.H. Hagman, B.C. Lesieutre, M.J. Boyd, E.P. Warren, H.P. Chou, R.D. Tabors. Monitoring the health of power transformers. *IEEE Computer Applications in Power*, 9(1):18–23, 1998.
- [83] J.L. Kirtley, J.C. Lavalley, D.J. McCarthy. Acoustic monitoring of transformer structures. *CIGRÉ Symposium Vienne*, 1987. Paper 820–03.
- [84] T. Kohonen. *Self-Organizing Maps*. Springer-Verlag, 1995.
- [85] T. Kohonen, E. Oja, O. Simula, A. Visa, J. Kangas. Engineering applications of the self-organizing map. *Proceedings of the IEEE*, 84(10):1358–1384, 1996.
- [86] T. Leibfried. Online monitors keep transformers in service. *IEEE Computer Applications in Power*, 11(3):36–42, 1998.
- [87] T. Leibfried, K. Feser. Monitoring of power transformers using the transfer function method. *IEEE Transactions on Power Delivery*, 14(4):1333–1341, 1999.
- [88] T. Leibfried, W. Knorr, K. Viereck, D. Dohnal, A. Kosmata, U. Sundrmann, B. Breitenbauch. On-line monitoring of power transformers – Trends, new developments and first experiences. *37th CIGRÉ session reports*, September 1998. Paper 12–211.
- [89] B.C. Lesieutre, W.H. Hagman, J.L. Kirtley. An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system. *IEEE Transactions on Power Delivery*, 12(1):249–256, 1997.
- [90] C.E. Lin, J.M. Ling, C.L. Huang. An expert system for transformer fault diagnosis using dissolved gas analysis. *IEEE Transactions on Power Delivery*, 8(1):231–238, 1993.

- [91] S.R. Lindgren, S.C. Pyke, M.A. Reynolds. Analyseur direct de la charge d'un transformateur et des gaz de défaut en vue de déterminer une surcharge fiable. *Rapports de la 36ème session CIGRÉ*, 1996. Article 12–110.
- [92] S.P. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, IT-28:129–137, 1982.
- [93] N. Macabrey, T. Baumann, A.J. Germond. Prédiction de charge dans un réseau électrique à l'aide du réseau de neurones de Kohonen. *Bulletin ASE/AES*, 83(5):13–19, 1992.
- [94] R. Malewski, B. Poulin. Impulse testing of power transformers using the transfer function method. *IEEE Transactions on Power Delivery*, 3(2):476–489, 1988.
- [95] J. Mao, A.K. Jain. Artificial neural networks for feature extraction and multivariate data projection. *IEEE Transactions on Neural Networks*, 6(2):296–317, 1995.
- [96] D. Maquin, J. Ragot. *Diagnostic des systèmes linéaires*. Hermès Science, 2000.
- [97] J. Max. Quantizing for minimum distortion. *IRE Transactions on Information Theory*, IT-28:7–12, 1960.
- [98] W.J. McNutt, T.O. Rouse, G.H. Kaufmann. Mathematical modelling of bubble evolution in transformers. *IEEE Transactions on Power Apparatus and Systems*, 104(2):21–48, February 1985.
- [99] G.K. Mechefske. Correlating power transformer tank vibration characteristics to winding looseness. *INSIGHT*, 37(8):599–604, 1995.
- [100] A. Möllmann, B. Pahlavanpour. New guidelines for interpretation of dissolved gas analysis in oil-filled transformers. *Electra*, 186:30–51, 1999.
- [101] S. Morgenthaler. *Introduction à la statistique*. Presses Polytechniques et Universitaires Romandes, 1997.
- [102] H. Mori, Y. Tamaru, S. Tsuzuki. An artificial neural-net based technique for power system dynamic stability with the Kohonen model. *IEEE Transactions on Power Systems*, 7(2):856–864, 1992.
- [103] A. Muñoz. *Aplicación de técnicas de redes neuronales artificiales al diagnóstico de procesos industriales*. Tesis Doctoral, Universidad Pontificia Comillas de Madrid, 1996.

- [104] A. Muñoz, M.A. Sanz-Bobi. An incipient fault detection system based on the probabilistic radial basis function network: Application to the diagnosis of the condenser of a coal power plant. *Neurocomputing*, 23:177–194, 1998.
- [105] A. Muñoz, J. Villar, A. Sanz-Bobi, A. Gimeno, L. Zarauza. Neural network approach to the diagnosis of the boiler combustion in a coal power plant. *International Conference on Artificial Neural Networks ICANN'95*, 1995.
- [106] D. Niebur. Artificial neural networks for power systems: A literature survey. *International Journal of Engineering Intelligent Systems*, 1(3):133–158, 1993.
- [107] D. Niebur. *Kohonen Self-Organizing Neural Network for Power System Security Assessment*. PhD thesis, École Polytechnique Fédérale de Lausanne, 1994.
- [108] D. Niebur, A.J. Germond. Power system static security assessment using the Kohonen neural network classifier. *IEEE Transactions on Power Systems*, 7(2):865–872, 1992.
- [109] Working Group 05 of CIGRÉ Study Committee 12. An international survey on failures in large power transformers in service. *Electra*, 88:21–48, 1983.
- [110] Working Group 09 of CIGRÉ Study Committee 12. Lifetime evaluation of transformers. *Electra*, 150:38–52, October 1993.
- [111] H. Ogi, H. Tanaka, Y. Akimoto, Y. Izui. Abnormality detection for GIS using self-organizing neural network. *Proceedings of the 11th Power Systems Computation Conference (PSCC) 1993*, pages 1171–1177, August 1993.
- [112] T.V. Oommen. Moisture equilibrium in paper-oil systems. *Proceedings of the Electrical/Electronics Insulation Conference*, pages 162–166, 1983.
- [113] T.V. Oommen. On-line moisture sensing in transformers. *Proceedings of the 20th Electrical Electronics Insulation Conference*, pages 236–240, 1991.
- [114] T.V. Oommen. On-line moisture monitoring in transformers and oil processing systems. *CIGRÉ Symposium Berlin*, 1993. Paper 110–03.



- [115] R. Otte, K. Goser. New approaches of process visualization and analysis in power plants. *Proceedings of WSOM'97*, pages 44–50, 1997.
- [116] A. De Pablo, A. Möllmann. New guidelines for furans analysis as well as dissolved gas analysis on oil-filled transformers. *36th CIGRÉ session reports*, 1996. Paper 15/21/33–19.
- [117] E.S. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.
- [118] R.J. Patton, J. Chen, T.M. Siew. Fault diagnosis in nonlinear dynamic systems via neural networks. *Proceedings of CONTROL'94*, pages 1346–1351, 1994.
- [119] A. Piras. *A Multiresponse Structural Connectionist Model for Short Term Electrical Load Forecasting*. PhD thesis, École Polytechnique Fédérale de Lausanne, 1996.
- [120] A. Piras, A. Germond. Local linear correlation analysis with the SOM. *Neurocomputing*, 21:79–80, 1998.
- [121] J. Poittevin, P. André. Nouvelles mesures de décharges partielles numériques appliquées au transformateur. *Rapports de la 36ème session CIGRÉ*, 1996. Article 15/21/33–17.
- [122] M.M. Polycarpou, A.T. Vemuri. Learning methodology for failure detection and accommodation. *IEEE Control Systems Magazine*, 15(3):16–24, 1995.
- [123] T.D. Poyser. An on-line microprocessor based transformer analysis system to improve the availability of power transformers. *IEEE Transactions on Power Apparatus and Systems*, PAS-102(4):957–962, April 1983.
- [124] T.D. Poyser, D.A. Yannucci, J.B. Templeton, B.N. Lendrking. On-line monitoring of power transformers. *IEEE Transactions on Power Apparatus and Systems*, PAS-104(1):207–211, January 1985.
- [125] J.H. Provanzana, P.R. Gattens, W.H. Hagman, H.R. Moore, J.W. Harley, J.E. Triner. Transformer condition monitoring – Realizing an integrated adaptive analysis system. *34th CIGRÉ session reports*, 1992. Paper 12–105.
- [126] B. Richardson. Transformer life management, bushings and tap-changers. *Colloquium on Transformer Life Management*, pages 8/1–8/4, October 1998.

- [127] H. Ritter. Asymptotic level density for a class of vector quantization process. *IEEE Transactions on Neural Networks*, 2(1):173–175, 1994.
- [128] H. Ritter, T. Martinetz, K. Schulten. *Neural Computation and Self-Organizing Maps*. Addison-Wesley, 1992.
- [129] H. Ritter, K. Schulten. Convergence properties of Kohonen's topology conserving maps : Fluctuations, stability and dimension selection. *Biological Cybernetics*, 60:59–71, 1988.
- [130] O. Roizman, V. Davydov. Neuro-fuzzy computing for large power transformers monitoring and diagnostics. *18th International Conference of the North American Fuzzy Information Processing Society NAFIPS*, pages 248–252, 1999. IEEE.
- [131] R. Rojas. *Neural Networks, A Systematic Introduction*. Springer-Verlag, 1996.
- [132] T.O. Rouse. Mineral insulating oil in transformers. *IEEE Electrical Insulation Magazine*, 14(3):6–16, 1998.
- [133] M. Runde, G.E. Ottesen, B. Skyberg, M. Ohlen. Vibration analysis for diagnostic testing of circuit-breakers. *IEEE Transactions on Power Delivery*, 11(4):1816–1823, 1996.
- [134] J.W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, C-18(5):401–409, 1969.
- [135] J. Sans, K. M. Bilgin, J.J. Kelly. Large-scale survey of furanic compounds in operating transformers and implications for estimating service life. *Conference Record of the 1998 IEEE International Symposium on Electrical Insulation*, pages 543–553, 1998.
- [136] M.A. Sanz-Bobi, A. García-Cerrada, R. Palacios, J. Villar. Experiences learned from the on-line internal monitoring of the behaviour of a transformer. *International Electric Machines and Drives Conference*, 1997.
- [137] M.A. Sanz-Bobi, A. Garcia-Cerrada, J. Rolan, A. Luengo, A.M. Alonso J.C. Burgos, F. Fernandez. TRAFES : Expert system for diagnosis of power transformers. *CIGRÉ Symposium Berlin*, 1993. Paper 110–20.
- [138] A. Schenk, A. Germond, P. Boss, P. Lorin. Surveillance continue des transformateurs : application des méthodes d'intelligence artificielle. *Bulletin ASE/AES*, 4:29–33, Mars 2000.

- [139] A. Schenk, A. Piras, A. Germond, P. Boss, P. Lorin, J. Dutoit. On-line monitoring of power transformers using self-organizing maps. *Proceedings of the 13th Power Systems Computation Conference (PSCC) 1999*, 2:945–951, July 1999.
- [140] A. Schnettler, M. Kurrat. Partial discharge diagnosis using an artificial neural network. *8th International Symposium on High Voltage Engineering*, pages 57–60, 1993.
- [141] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986.
- [142] O. Simula, E. Alhoniemi, J. Hollmén, J. Vesanto. Monitoring and modelling of complex process using hierarchical self-organizing maps. *IEEE International Symposium on Circuits and Systems*, Supplement:73–76, 1996.
- [143] B. Sparling, J. Aubin, A. Glodjo. Field experience with multi-gas on-line monitoring of power transformers. *Transmission and Distribution Conference*, 2:895–900, 1999. IEEE.
- [144] M. Stace, S.M. Islam. Condition monitoring of power transformers in Australian state of New South Wales using transfer function measurements. *Proceedings of the 5th International Conference on Properties and Applications of Dielectric Materials*, pages 248–251, 1997.
- [145] IEC Standard. *Specification for cellulosic papers for electrical purposes. Part 1: Definitions and general requirements*. CEI/IEC 60554-1, 1977.
- [146] IEC Standard. *Measurement of relative permittivity, dielectric dissipation factor and D.C. resistivity of insulating liquids*. CEI/IEC 60247, 1978.
- [147] IEC Standard. *Specification for pressboard and presspaper for electrical purposes. Part 1: Definitions and general requirements*. CEI/IEC 60641-1, 1979.
- [148] IEC Standard. *Partial discharge measurements*. CEI/IEC 60270, 1981.
- [149] IEC Standard. *Guide to the lightning impulse and switching impulse testing of power transformers and reactors*. CEI/IEC 60722, 1982.

- [150] IEC Standard. *Specification for unused mineral insulating oils for transformers and switchgear*. CEI/IEC 60296, 1982.
- [151] IEC Standard. *Supervision and maintenance guide for insulating mineral insulating oils in electrical equipment*. CEI/IEC 60422, 1989.
- [152] IEC Standard. *Loading guide for oil-immersed power transformers*. CEI/IEC 60354, 1991.
- [153] IEC Standard. *Mineral insulating oils – Methods for the determination of 2-furfural and related compounds*. CEI/IEC 61198, 1993.
- [154] IEC Standard. *Power transformers – Part 2: Temperature rise*. CEI/IEC 60076-2, 1993.
- [155] IEC Standard. *Insulating liquids – Determination of the breakdown voltage at power frequency – Test method*. CEI/IEC 60156, 1995.
- [156] IEC Standard. *Insulating liquids – Oil-impregnated paper and pressboard – Determination of water by automatic coulometric Karl Fischer titration*. CEI/IEC 60814, 1997.
- [157] IEC Standard. *Mineral oil-impregnated electrical equipment in service – Guide to the interpretation of dissolved and free gases analysis*. CEI/IEC 60599, 1999.
- [158] IEC Standard. *Power transformers – Part 3: Insulation levels, dielectric tests and external clearances in air*. CEI/IEC 60076-3, 2000.
- [159] IEEE Standards. *IEEE Guide for the Reclamation of Insulating Oil and Criteria for Its Use*. IEEE Std 637–1985, 1985.
- [160] IEEE Standards. *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*. IEEE Std C57.104–1991, 1991.
- [161] IEEE Standards. *IEEE Standard Terminology for Power and Distribution Transformers*. ANSI/IEEE Std C57.12.80–1978, 1992.
- [162] IEEE Standards. *IEEE Standard Guide for Diagnostic Field Testing of Electric Power Apparatus*. IEEE Std 62–1995, 1995.
- [163] IEEE Standards. *IEEE Guide for Loading Mineral-Oil-Immersed Transformers*. IEEE Std C57.91–1995, 1996.

- [164] S. Tamminen, S. Pirttikangas, J. Röning. Self-organizing maps in adaptive health monitoring. *International Joint Conference on Neural Networks IJCNN'2000*, 4:259–264, 2000.
- [165] M. Tanaka, M. Sakawa, I. Shiromaru, T. Matsumoto. Application of Kohonen's self-organizing network to the diagnosis system for rotating machinery. *IEEE International Conference on Systems, Man and Cybernetics*, 5:4039–4044, 1995.
- [166] S. Tenbohlen, F. Figel. On-line condition monitoring of power transformers. *IEEE Power Engineering Society Winter Meeting*, 2000.
- [167] S. Tenbohlen, D. Uhde, J. Poittevin, H. Borsi, P. Werle, U. Sundermann, H. Matthes. Enhanced diagnosis of power transformers using on- and off-line methods: Results, examples and future trends. *38th CIGRÉ session reports*, 2000. Paper 12–204.
- [168] K.F. Thang, R.K. Aggarwal, D.G. Esp, A.J. McGrail. Statistical and neural network analysis of dissolved gases in power transformers. *Eighth International Conference on Dielectric Materials, Measurements and Applications*, pages 324–329, 2000. IEE Conf. Publ. No. 473.
- [169] K. Tomsovic, M. Tapper, T. Ingvarsson. A fuzzy information approach to integrating different transformer diagnostic methods. *IEEE Transactions on Power Delivery*, 8(3):1638–1646, 1993.
- [170] Y. Tournier. *Transformateurs*. Alsthom-Savoisienne, 1969.
- [171] V. Tryba, K. Goser. Self-organizing feature maps for process control in chemistry. *Artificial Neural Networks*, pages 847–852, 1991.
- [172] V. Tryba, S. Metzen, K. Goser. Designing basic integrated circuits by self-organizing feature maps. *Proceedings of Neuro-Nîmes '89*, pages 225–235, 1989.
- [173] D. Tschudi. *Localisation de défauts de transformateurs de puissance lors de l'essai de choc*. Thèse de doctorat, École Polytechnique Fédérale de Lausanne, 1993.
- [174] H. Tsukioka, K. Sugawara, E. Mori. Apparatus for continuously monitoring hydrogen gas dissolved in transformer oil. *IEEE Transactions on Electrical Insulation*, EI-16(6):502–509, 1981.

- [175] D.J. Tylavsky, Q. He, J. Si, G.A. McCulla, J.R. Hunt. Transformer top-oil temperature modeling and simulation. *IEEE Transactions on Industry Applications*, 36(5):1219–1225, 2000.
- [176] A. Ultsch. Self organized feature maps for monitoring and knowledge acquisition of a chemical process. *International Conference on Artificial Neural Networks ICANN'93*, pages 864–867, 1993.
- [177] A. Ultsch, H.P. Siemon. Kohonen's self organizing feature maps for exploratory data analysis. *Proceedings of International Neural Network Conference INNC'90*, pages 305–308, 1990.
- [178] M. Vapola, O. Simula, T. Kohonen, P. Meriläinen. Representation and identification of fault conditions of an anaesthesia system by means of the self-organizing map. *International Conference on Artificial Neural Networks ICANN'94*, 1:350–353, 1994.
- [179] J. Vesanto. SOM-based data visualization methods. *Intelligent Data Analysis*, 3:111–126, 1999.
- [180] T. Villmann, R. Der, M. Herrmann, T.M. Martinetz. Topology preservation in self-organizing feature maps: Exact definition and measurement. *IEEE Transactions on Neural Networks*, 8(2):256–265, 1997.
- [181] X.Z. Wang. *Data Mining and Knowledge Discovery for Process Monitoring and Control*. Springer-Verlag, 1999.
- [182] Z. Wang, Y. Liu, P.J. Griffin. A combined ANN and expert system tool for transformer fault diagnosis. *IEEE Transactions on Power Delivery*, 13(4):1224–1229, 1998.
- [183] D.A. Waterman. *A Guide to Expert Systems*. Addison-Wesley, 1986.
- [184] M. Webb. Water in oils. *An Engineering Review of Liquid Insulation*, pages 2/1–2/5, 1997.
- [185] P. Werle, H. Borsi, E. Gockenbach. Hierarchical cluster analysis of broadband measured partial discharges as part of a modular structured monitoring system for transformers. *11st High Voltage Engineering Symposium*, 5:29–32, 1999.
- [186] D.J. Willshaw, C. von der Malsburg. How patterned neural connections can be set up by self-organisation. *Proceedings of the Royal Society of London*, B 194:431–445, 1976.

- [187] A.S. Willsky. A survey of design methods for failure detection in dynamics systems. *Automatica*, 12:601–611, 1976.
- [188] H.-T. Yang, C.-C. Liao. Adaptive fuzzy diagnosis system for dissolved gas analysis of power transformers. *IEEE Transactions on Power Delivery*, 14(4):1342–1350, 1999.
- [189] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [190] Y. Zhang, X. Ding, Y. Liu, P.J. Griffin. An artificial neural network approach to transformer fault diagnosis. *IEEE Power Engineering Society Winter Meeting*, January 1996.





# Index

- Activation, zone d', 91, 107
- Adaptation périodique, 112, 141
- Alarme robuste, 67
- Appartenance (fonction floue), 92
- Apprentissage, 74, 80, 92
  - paramètres d', 83, 93, 109
- Arrhenius, loi d', 16
- Augmentation brusque, 148, 156
  
- Base de données, 102, 104, 105, 127, 140, 163
- Batch (ou par paquets), apprentissage, 93
- Bornes de traversées, 42
- Buchholz, relais de, 34
  
- Cartes auto-organisées, 79, 99
  - organisation, 112
  - principe de surveillance, 102
- Changement rapide, 60, 103
- Changeurs de prises, 43
- Coalescence, 57
- Composantes principales, 54
- Conscience, 87
- Corrélation linéaire, 54, 130–132
- Cortex cérébral, 76
- Couches
  - organisation, 137
  - représentation des, 114
  
- Covariance, 53
  
- Décharges partielles, 36
  - localisation acoustique, 38
  - localisation électrique, 36, 39
- Décomposition structurelle, 58
- Défaillances
  - causes, 19
  - détection, 47, 58
  - répartition, 18
- Défaut élémentaire, 149, 157
- Densité de probabilité, 86, 96, 97
- Dérive d'une variable, 149, 155, 156, 161, 162
- Détection de changement, 65, 123
- Distance euclidienne, 76, 79, 93, 102, 121
- Dynamique, topologie, 87
  
- Écart type, 53
- Échantillon, 105
- Entropie, 89
- Entrée
  - espace d', 76, 79
  - vecteur d', 105
- Équipements auxiliaires, 41, 42
- Équiprobabilité, 87, 89, 107
- Erreur

- type I, 60, 66, 152
  - type II, 60, 66, 152
- Erreur de classification, 148, 150
- Erreur de quantification, 76, 79, 93, 102
- Erreur quadratique, minimisation, 64
- Espérance mathématique, 52
- Étiquetage, 119
- Filtrage, 136
- Gagnant, neurone, 79, 80, 118
- Gaz
  - bulles, 34
  - défaillances, 22, 32
  - gaz clés, 30
  - gaz dissous, 29, 159
  - méthode des rapports, 31
  - représentation graphique, 32
  - surveillance continue, 33
- GTM, 88
- Histogramme d'activation, 117, 158
- Huile, 8
  - additifs, 10
  - analyses physico-chimiques, 26
  - circulation, 13
  - coloration, 25
  - composition, 9
  - mesures diélectriques, 26
  - température de l', 131, 152, 157
- Humidité, 131, 133, 161
  - défaillances, 20
  - méthode de Karl Fischer, 34
  - modélisation, 63
  - taux relatif, 34
- Image thermique, 27
- Indicateur de confiance, 98, 107, 121, 142–147, 157–161
- Information, quantité d', 89
- Interface graphique, 157
- Interprétation de la carte, 119, 138
- Isolation, 8
- kMER
  - adaptation contrôlée du facteur de voisinage, 97
  - apprentissage, 92
  - apprentissage optimisé, 93
  - cartes, 91
  - paramètres, 94
- Kohonen
  - apprentissage, 80, 82
  - cartes de, 79
  - voisinage, 79
- Logique floue, 71
- Lorentz, forces de, 41
- Magnétostriction, 41
- Matrice des distances, 115
- Mesures, 25, 44, 126, 163
  - off-line, 25
  - on-line, 25
- Modèle
  - linéaire, 63
  - neuronal, 64
  - physique, 63
- Moindres carrés, 64

- Neurone gagnant, 79, 80, 118  
Nombre de neurones, 147  
Normalisation, 55, 81, 85, 106  
Organisation de la carte, 112, 139  
Oxygène, 22  
Paramètres d'apprentissage, 83,  
93, 109, 141  
PCA, 54  
Performances, 142, 150, 153  
Point chaud, 13, 27  
Point de fonctionnement, 119, 162  
Prétraitement, 134  
Produits de décomposition, 35  
Projection, 82  
Quantification vectorielle, 57, 75  
Rapidité de détection, 152  
Rapport de vraisemblance, 67  
Réapprentissage, 113, 140, 141  
Refroidissement, 11, 42  
Régime de fonctionnement, 138,  
158  
Réponse diélectrique, 40  
Représentation graphique, 159,  
160, 162  
Représentation graphique, 118,  
140, 162  
Réseau de neurones artificiels, 73  
Réseaux de neurones  
architectures, 73  
cartes de Kohonen, 79  
méthodes non supervisées,  
75  
méthodes supervisées, 74  
modèle biologique, 78  
Résidus, 62, 65  
Saut brusque, 128, 149, 150, 153,  
155, 156  
Sélection des données, 134  
Sensibilité, 144  
Seuil adaptatif, 66  
Seuil de détection, 121, 144, 154  
Seuil fixe, 61  
Statistique des variables consi-  
dérées, 142  
Surveillance continue, 47, 100  
Système expert, 70  
Temps de détection, 148, 152,  
163  
Température  
de l'huile, 131, 152, 157  
défaillances, 20  
modélisation, 63  
techniques de mesure, 27  
Tendance, 59, 62, 102  
Topologique, organisation, 84  
Trajectoire, 119, 158, 162  
Transformateurs  
catégories, 5  
constitution, 6  
Transformateur pilote, 125  
U-matrix, 115  
Unité pilote, 125  
Unités mortes, 87, 109, 123  
Valeur faussée, 142  
Valeurs instantanées, 102  
Variabilité de recouvrement, 96,  
109  
Variance, 52

- Vecteur d'entrée, 105
- Vecteur de fonctionnement, 134,  
136
- Vibrations, 41, 127, 133
- Vieillessement, 15, 17
  - détection, 58
- VQP, 88
  
- Zone d'activation, 91, 107

# Curriculum vitae

Alain Schenk

né le 27 février 1972

originaire d'Onex (GE) et Eggwil (BE)

marié

## Formation

*1996* Ingénieur dipl. EPF en électricité (École Polytechnique Fédérale de Lausanne)

*1992* Ingénieur dipl. ETS en électricité avec orientation en génie nucléaire (École d'Ingénieurs de Genève)

## Expérience professionnelle

*dès 1996* Assistant de recherche et d'enseignement au Laboratoire de Réseaux d'Énergie Électrique (LRE) de l'EPFL  
*domaines d'activité* : intelligence artificielle appliquée au domaine de l'énergie électrique ; simulation numérique des phénomènes dynamiques dans les réseaux électriques

*1996–2000* Administrateur du réseau informatique au LRE

*1992* Stagiaire au laboratoire de radioprotection du CERN  
*travail réalisé* : caractérisation d'une chaîne de mesure de contamination radioactive

## Publications

- A. Schenk, A. Germond, P. Boss, P. Lorin. Surveillance continue des transformateurs : application des méthodes d'intelligence artificielle. *Bulletin ASE/AES*, 4:29–33, mars 2000.

- A. Schenk, A. Piras, A. Germond, P. Boss, P. Lorin, J. Dutoit. On-line monitoring of power transformers using self-organizing maps. *Proceedings of the 13th Power Systems Computation Conference (PSCC) 1999*, 2:945–951, July 1999.
- R. Cherkaoui, P.-A. Chamorel, A. Schenk, F. Renaud. Étude du comportement dynamique de la centrale de Bieudron (Cleuson-Dixence). *Bulletin ASE/AES*, 15:14–18, juillet 1999.
- Th. Aschwanden, M. Hässig, J. Fuhr, P. Lorin, V. Der Houhanessian, W. Zaengl, A. Schenk, P. Zweiacker, A. Piras. Development and application of new condition assessment methods for power transformers. *37th CIGRÉ session reports*, September 1998. Paper 12–207.

### Activités associatives

- 2000* Professeur de ski degré III, brevet de l'Interassociation Suisse pour les Sports de Neige (IASS)
- dès 1999* Instructeur au centre de formation de l'Interassociation Genevoise pour l'Enseignement du Ski (IGES)
- dès 1997* Membre du comité de l'école de ski et de snowboard Glisse Puissance 3 (GP3) à Genève