

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



المدرسة الوطنية المتعددة التخصصات
Ecole Nationale Polytechnique

ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT DU GENIE ELECTRIQUE

MEMOIRE DE FIN D'ETUDES
D'INGENIEUR D'ETAT EN ELECTROTECHNIQUE

Sujet :

**Application des systèmes hybrides
neuro-flous au diagnostic
des huiles de transformateurs**

Présenté par :

Najet AOUCHAR & Chafika BEKHALED

Membres du jury :

F. BOUDJEMA Professeur, Président	(E.N.P)
A. BOUBAKEUR Professeur, Promoteur	(E.N.P)
L. MOKHNACHE Docteur d'Etat, Co-promotrice	(Université de Batna)
A.MEKHALDI Professeur, Examineur	(E.N.P)
M.BOUDOUR Docteur d'Etat, Examineur	(USTHB)
B.SAFTA , Membre invité	(Transfo-Centre, Blida)
D.RAHAL , Membre invité	(Transfo-Centre, Blida)

- Juin 2005 -

DEDICACES

-A mes très chers parents, que dieu me les garde éternellement heureux.

-A mes sœurs et mes frères.

-A ma grand-mère, et mon oncle.

-A tous mes amis.

-A tous ceux qui me sont chers.

-A ma ville, et à la mémoire de tous ses gens qui ont péri un 21 mai 2003.

A tous,

je dédie ce modeste travail.

Najet AOUCHAR.

A papa, maman, mani, Rafik et Rahid.

Je ne serai jamais arrivée là où je suis sans vous.

Merci pour tout.

Chafika BEKHALED

REMERCIEMENTS

Nous voulons exprimer notre gratitude au Professeur Ahmed BOUBAKEUR qui nous a honoré en dirigeant ce travail. Nous voulons le remercier d'avoir mis à notre disposition tous les moyens nécessaires pour mener à bien ce modeste travail. Ses insignes observations et suggestions ont énormément contribué dans la réussite de ce projet.

Nous adressons notre sincère reconnaissance, ainsi que notre profond respect, à Mademoiselle Leïla MOKHNACHE, docteur à l'université de Batna, qui malgré l'éloignement géographique a suivi, pas à pas, l'évolution de notre travail. Elle n'a pas cessé de nous soutenir par toutes ses orientations, ses contributions et ses précieuses remarques qui nous ont toujours servies.

C'est avec un énorme plaisir que nous formulons notre gratitude, et nos remerciements à la direction et aux travailleurs de la société de réparation des transformateurs électriques du centre (Transfo-Centre -unité de BLIDA-). Ils nous ont accueilli chaleureusement et ont été particulièrement serviables. Nos remerciements, les plus vifs vont à Madame Dalal RAHAL, Responsable du laboratoire d'analyses des huiles diélectriques, pour l'aide précieuse qu'elle nous a apportée et la patience dont elle a fait preuve avec nous.

Nous tenons à remercier Monsieur Farès BOUDJEMA, Professeur à l'Ecole Nationale Polytechnique, de nous avoir fait l'honneur de présider le jury et Monsieur Mohamed BOUDOUR, Maître de Conférence à l'Université des Sciences et Technologies Houari Boumedienne (USTHB), pour avoir bien voulu évaluer notre travail.

Nous remercions également Monsieur Abdelouaheb MEKHALDI, professeur à l'Ecole Nationale Polytechnique, d'avoir accepté de juger notre travail et tenons à lui exprimer notre gratitude pour tous ses encouragements.

Nos remerciements, sont adressés aux enseignants de l'Ecole Nationale Polytechnique qui ont participé, d'une manière où d'une autre, à notre formation ainsi qu'à l'élaboration et l'aboutissement de ce projet nous citons, plus particulièrement Monsieur Madjid TEGUAR, Maître de Conférence, Monsieur Mohamed OUADJAOUT, Chargé de Cours et Monsieur H.CHKIREB, Maître de conférence.

Nous témoignons notre profonde gratitude aux êtres qui nous sont les plus chers : nos parents : Madame et Monsieur AOUCHAR et Madame et Monsieur BEKHALED, pour leur soutien dans toutes les étapes de notre vie, mille mercis.

Et finalement, nous remercions tous nos amis, particulièrement les étudiants des départements de Génie Electrique et d'Electronique ainsi que toutes les personnes qui ont contribué de près ou de loin à ce modeste travail.

TABLE DES MATIERES

INTRODUCTION GENERALE..... 1

Chapitre I : INTRODUCTION AUX RESEAUX DE NEURONES

1.1. Historique..... 5

1.2. Le neurone artificiel..... 5

 1.2.1. Le neurone biologique..... 5

 1.2.2. Le neurone formel..... 6

 1.2.2.1. La nature des entrées..... 6

 1.2.2.2. La fonction d'entrée totale..... 6

 1.2.2.3. La fonction d'activation..... 7

 1.2.2.4. La fonction de sortie..... 7

 1.2.2.5. La nature de la sortie du neurone..... 7

 1.2.3. Exemple du fonctionnement d'un neurone formel..... 7

1.3. Les réseaux de neurones artificiels..... 8

 1.3.1. La structure des connexions..... 8

 1.3.1.1. Les réseaux entièrement connectés..... 8

 1.3.1.2. Les réseaux en couches..... 9

 1.3.2. La dynamique des connexions..... 9

 1.3.2.1. L'apprentissage supervisé..... 9

 1.3.2.1.1. Data adaptive learning (Feed-Forward Network pattern)..... 9

 1.3.2.1.2. Block adaptive learning (batch learning) 10

 1.3.2.2. L'apprentissage non supervisé..... 10

 1.3.2.2.1. Les réseaux compétitifs..... 10

 1.3.2.2.2. Les réseaux auto-organiseurs de Kohonen..... 10

1.4. Les techniques d'apprentissage adaptées à la classification..... 11

 1.4.1. La technique de la descente de gradient..... 11

 1.4.2. L'algorithme de rétro-propagation (Back-Propagation)..... 11

 1.4.2.1. La rétro-propagation avec Momentum..... 12

 1.4.2.2. La rétro-propagation à taux d'apprentissage adaptatif..... 12

 1.4.3. Les limites de la rétro-propagation..... 13

 1.4.4. La méthode de Levenberg-Marquardt..... 13

1.5. L'application des réseaux de neurones artificiels en Technique de la Haute Tension..... 14

 1.5.1. Les décharges partielles et les tensions disruptives..... 14

 1.5.2. Les systèmes d'isolations électriques..... 15

 1.5.2.1. L'isolation des transformateurs..... 15

 1.5.2.2. L'isolation des câbles électriques de haute tension..... 16

 1.5.3. Les autres applications..... 17

Chapitre II : INTRODUCTION A LA LOGIQUE FLOUE

2.1. Historique.....	19
2.2. Les bases générales de la logique floue.....	20
2.2.1. La logique booléenne.....	20
2.2.2. La théorie des ensembles flous.....	20
2.2.2.1. La fonction d'appartenance.....	21
2.2.2.2. Les différentes formes de fonctions d'appartenance.....	22
2.2.2.2.1. La fonction triangulaire.....	22
2.2.2.2.2. La fonction trapézoïdale.....	22
2.2.2.2.3. La fonction gaussienne.....	22
2.2.2.2.4. La fonction en forme de cloche.....	22
2.2.2.2.5. La fonction sigmoïde.....	22
2.2.2.3. Les ensembles flous.....	23
2.2.2.4. Les opérateurs de la logique floue.....	23
2.2.2.4.1. Les opérations algébriques.....	23
2.2.2.4.1.1. La multiplication cartésienne.....	23
2.2.2.4.1.2. Le produit algébrique.....	23
2.2.2.4.1.3. L'exposant.....	23
2.2.2.4.1.4. La concentration et la dilatation.....	24
2.2.2.4.1.5. La somme algébrique.....	24
2.2.2.4.2. Les opérateurs linguistiques.....	24
2.2.2.4.3. Les opérateurs logiques.....	24
2.2.2.4.3.1. La complémentation.....	25
2.2.2.4.3.2. L'intersection.....	25
2.2.2.4.3.3. L'union.....	26
2.3. La logique floue et les probabilités.....	27
2.4. Les systèmes basés sur la logique floue.....	27
2.4.1. L'univers de discours.....	28
2.4.2. Les termes et les variables linguistiques.....	28
2.4.3. Les règles linguistiques.....	28
2.4.4. L'implication floue.....	29
2.4.5. La structure générale d'un système flou.....	29
2.4.5.1. L'interface de fuzzification.....	29
2.4.5.2. La base de connaissance.....	29
2.4.5.3. Le bloc de décision ou le moteur d'inférence.....	30
2.4.5.4. L'interface de défuzzification.....	30
2.4.6. La procédure de raisonnement flou.....	30
2.4.6.1. La fuzzification.....	30
2.4.6.2. Les inférences floues.....	31
2.4.6.3. L'agrégation des règles linguistiques.....	32
2.4.6.4. Les types d'inférences floues.....	32
2.4.6.4.1. La méthode d'inférence max-min.....	32
2.4.6.4.2. La méthode d'inférence max-prod.....	32
2.4.6.4.3. La méthode d'inférence somme-prod.....	32
2.4.6.5. La défuzzification.....	32
2.4.6.5.1. La méthode du centre de gravité.....	33
2.4.6.5.2. La méthode de la moyenne des maxima.....	33

Chapitre V : DIAGNOSTIC DES HUILES PAR LES SYSTEMES ADOPTES

5.1. La combinaison des réseaux neuronaux et de la logique floue pour le diagnostic.....	52
5.2. Les entrées et les sorties des systèmes.....	52
5.3. Les architectures des systèmes neuro-flous adoptés.....	53
5.3.1. Le système flou.....	53
5.3.2. Le réseau de neurones artificiels.....	54
5.3.3. Le réseau de neurones flous.....	55
5.3.4. Le réseau de Pedrycz.....	55
5.3.5. Le réseau de neurones à initialisation floue.....	56
5.3.6. Les systèmes fusionnés.....	57
5.3.6.1. La structure fusionnée compacte.....	57
5.3.6.2. La structure fusionnée dissociée.....	58
5.3.6.2.1. Les réseaux de fuzzification.....	58
5.3.6.2.2. Le réseau d'inférences floues.....	59
5.4. Les résultats du diagnostic des huiles par les systèmes adoptés.....	60
5.4.1. Le diagnostic par la logique floue.....	60
5.4.2. Le diagnostic par le réseau de neurones artificiels.....	61
5.4.3. Le diagnostic par le réseau de neurones flous.....	62
5.4.4. Le diagnostic par le réseau de neurones à initialisation floue.....	64
5.4.5. Le diagnostic par le réseau de Pedrycz.....	66
5.4.6. Le diagnostic par le réseau fusionné compact.....	68
5.4.7. Le diagnostic par la structure fusionnée dissociée.....	69
5.5. Discussions et conclusions.....	73
CONCLUSION GENERALE.....	75
BIBLIOGRAPHIE.....	77
ANNEXES	
Annexe 1. La base de données utilisée pour l'entraînement des réseaux neuronaux.....	82
Annexe 2. Les règles utilisées pour le système flou.....	86

Introduction générale

INTRODUCTION GENERALE

L'électricité fait partie intégrante de notre vie. Il suffit que les lumières s'éteignent pour que la vie s'arrête. Cette dépendance impose aux entreprises de production et de transport de l'électricité un service permanent, surtout en ce qui concerne les infrastructures stratégiques. Une interruption de l'alimentation peut leur causer des dégâts irrémédiables. Pour un établissement industriel, par exemple, une coupure imprévue peut entraîner de lourdes pertes de la productivité.

Les interruptions électriques sont généralement causées par la vétusté des équipements électriques, en particulier les transformateurs.

Les transformateurs électriques jouent un rôle majeur dans la distribution et la transmission de l'électricité. Ils subissent régulièrement des défaillances techniques et peuvent exploser. Quand de telles avaries se produisent, la vie des personnes est mise en danger et les conséquences financières pour les sociétés d'électricité sont considérables.

Ces accidents sont favorisés par l'accroissement régulier de la consommation en électricité. En effet, la production augmente de 2% par ans dans le monde [1] et de 5.8% en Algérie [2]. Cette évolution représente pour les anciens transformateurs une surcharge, d'autant plus que l'âge de la plupart d'entre-eux dépasse les 30 ans. Ceci provoque leur dégradation prématurée.

Les pertes humaines et financières engendrées par les explosions des transformateurs mettent l'accent sur la nécessité de trouver des méthodes préventives permettant de réduire ces pertes sans causer l'arrêt inutile des installations. Ces méthodes doivent être assez efficaces pour permettre d'agir avant que l'accident ne survienne.

L'huile isolante est l'un des éléments les plus importants dans un transformateur. Elle requiert une surveillance et un entretien régulier pour pouvoir garantir une bonne isolation et un bon refroidissement. La surveillance de l'état de l'huile d'un transformateur en service est un moyen de prévention efficace. Elle peut également renseigner sur l'état intérieur de la partie active.

Pour pouvoir assurer ses rôles, l'huile diélectrique doit être conforme aux recommandations des normes. L'application de ces normes exige beaucoup d'expérience car il y a presque toujours des conflits entre les données des tests, ainsi que des cas d'indécision.

Pour permettre à des ingénieurs inexpérimentés de s'approcher de la qualité de diagnostic des experts, de nombreuses études ont été entreprises pour développer des systèmes intelligents d'aide à la décision. Le recours à ce type de systèmes artificiels permet également de réduire le temps de la prise de décision.

Les réseaux de neurones artificiels et la logique floue, sont deux techniques largement utilisées dans le domaine du diagnostic et de la surveillance des systèmes industriels. Leur usage a donné des résultats positifs, ce qui nous a incité à les introduire dans les systèmes de diagnostic des huiles de transformateurs. Leur combinaison permet de réunir les avantages de chaque technique et de pallier les inconvénients qu'elles présentent.

Le travail, présenté dans le cadre de ce projet de fin d'études, consiste à faire une synthèse et une comparaison de différentes combinaisons de réseaux neuronaux et de la logique floue pour la classification et le diagnostic. Il contribue également à enrichir ces méthodes en proposant deux nouvelles structures hybrides. Il est réparti comme suit :

Le premier chapitre traite des réseaux de neurones artificiels. Il comporte une brève introduction aux neurones biologiques, à leur modélisation et à leur connexion ainsi que l'algorithme détaillé de l'apprentissage par descente du gradient et ses variantes. Pour l'ensemble des réseaux programmés, nous avons opté pour l'adaptation des poids suivant la méthode de Levenberg-Marquardt.

Le second chapitre est consacré à la logique floue. Il contient un exposé des notions de base de cette théorie, de la structure générale d'un système flou et de la procédure du raisonnement flou.

Le troisième chapitre est une synthèse non exhaustive des travaux de recherches concernant l'application des structures neuro-floues à la classification et au diagnostic. Nous y recensons les systèmes hybrides proposés dans la littérature.

Le quatrième chapitre présente des notions fondamentales de maintenance et de diagnostic des huiles de transformateurs.

Enfin, le dernier chapitre comporte une étude détaillée de toutes les structures programmées et une analyse de leurs performances.

Nous concluons ce travail par un récapitulatif des systèmes étudiés, une synthèse des résultats obtenus ainsi que les perspectives envisagées.

Chapitre I

Introduction aux réseaux de neurones

1.1. HISTORIQUE :

Les réseaux de neurones artificiels (RNA) ont été évoqués pour la première fois en 1943, dans un article resté historique publié par Mc Culloch et Pitts [3]. Ils avaient réussi à créer le neurone formel qui est la modélisation du neurone biologique.

Peu de temps après, en 1949, D.O Hebb, dans un ouvrage de référence, a défini la notion de plasticité synaptique ou autrement dit d'apprentissage. Depuis, des travaux sur les neurones formels ont été élaborés et plusieurs articles furent publiés vers la fin des années cinquante (50). En 1958, Rosenblat a proposé pour la première fois le perceptron et en 1960, le modèle Adaline (Adaptative Linear Element) a été introduit par M. Hoff et Windrow.

Cependant, le manque d'approfondissement en matière d'études théoriques a engendré un ralentissement prolongé des travaux du connexionisme (connexion des neurones formels pour former un réseau neuronal). Entre 1970 et 1980 peu d'articles furent publiés en raison de la faible quantité de chercheurs intéressés.

Au début des années quatre-vingt (80), les recherches dans ce domaine ont repris surtout après la publication du travail de J.J. Hopfield en 1982.

Depuis et à nos jours, les RNA sont de plus en plus introduits dans différentes disciplines, à savoir la reconnaissance des formes statiques, la détection d'anomalies, le traitement dépendant du temps (la prédiction), l'identification et la commande des processus. Ils sont également utilisés dans certaines applications militaires, bancaires et financières, dans le domaine du traitement de la parole et des images, dans la robotique et même dans certaines industries de consommation.

En Technique de la Haute Tension, les RNA sont apparus en 1994, après la publication de A.A. Mazroua et al [4] dans 'The Institute of Electrical and Electronics Engineers (IEEE)' d'un article traitant de la reconnaissance des tailles d'impulsions de décharges partielles. Depuis, les travaux se sont succédés et ont touché tous les domaines de la Haute Tension.

1.2. LE NEURONE ARTIFICIEL :

La structure des neurones artificiels est inspirée du modèle du neurone biologique.

1.2.1. Le neurone biologique :

Les neurones ou les cellules nerveuses sont les éléments de base du système nerveux. Ils possèdent de nombreux points communs avec les autres cellules du corps humain et présentent des caractéristiques spécifiques qui leur permettent de recevoir des signaux en provenance des autres neurones, de les intégrer, d'engendrer un influx nerveux, de le conduire et de le transmettre. Le cerveau possède environ cent milliards de neurones reliés les uns aux autres dans des endroits appelés synapses. Leur interconnexion est dénommée *réseau de neurones*.

Un neurone est constitué de trois parties : le corps cellulaire, l'axone et les dendrites (Fig.1). Le rôle des dendrites est de capter l'information provenant des autres neurones. Cette information se présente sous la forme de médiateurs chimiques dits *neurotransmetteurs* libérés dans les synapses. La capture de ces substances par les dendrites donne naissance à un signal électrique appelé potentiel d'action qui est acheminé vers le corps cellulaire.

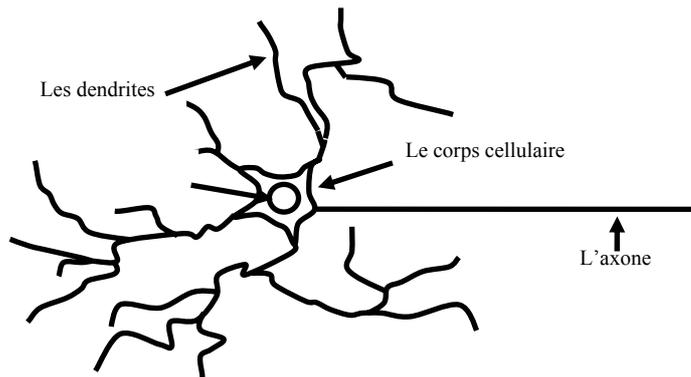


Fig.1 : Le neurone biologique [2].

1.2.2. Le neurone formel :

C'est un opérateur mathématique très simple qui possède une sortie et plusieurs entrées pouvant être les sorties d'autres neurones, ou des entrées de signaux extérieurs. Il est défini par les cinq éléments suivants (Fig.2) [5]:

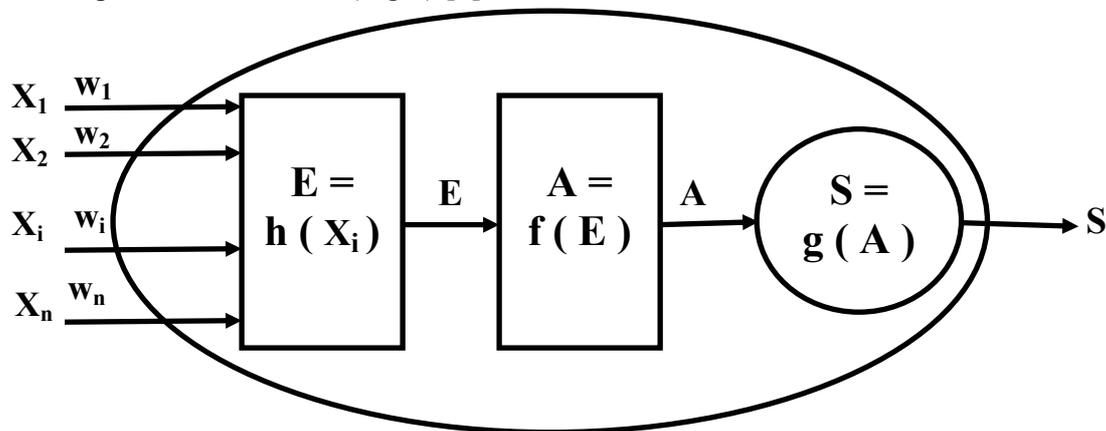


Fig.2 : Architecture du neurone artificiel .

1.2.2.1. La nature des entrées :

Elles peuvent être binaires ou réelles. Le vecteur d'entrée de chaque neurone possède une entrée spéciale appelée le *biais* (en anglais *bias*). Sa valeur est généralement égale à l'unité. Elle permet de déplacer la fonction d'activation sur l'axe (s) (Fig.3). Ce déplacement ne dépend que du poids attribué à cette entrée.

1.2.2.2. La fonction d'entrée totale :

Elle définit le pré-traitement des entrées. Elle peut être :

- booléenne ;
- linéaire : $E = h(e_1 \dots e_n) = \sum_{i=1 \dots n} W_i e_i$;
- affine : $E = h(e_1 \dots e_n) = \sum_{i=1 \dots n} W_i e_i - a$;
- polynomiale de degré supérieur à deux.

1.2.2.3. La fonction d'activation :

C'est une fonction non-linéaire qui définit l'état interne (A) du neurone en fonction de son entrée totale (E). Elle peut être une fonction binaire à seuil, linéaire à seuil ou multi-seuil, sigmoïde, stochastique ou une autre fonction généralement choisie croissante et impaire (Fig.3) [5].

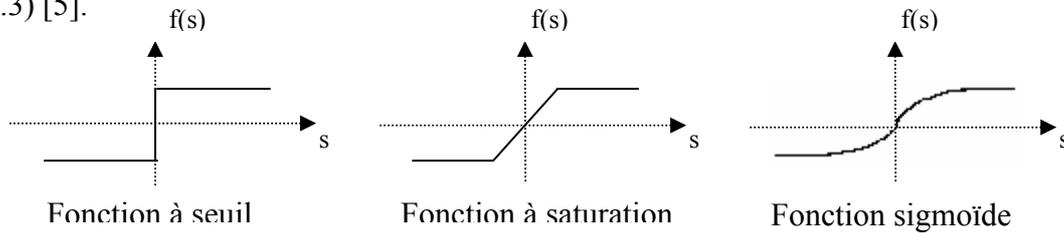


Fig.3 : Quelques fonctions d'activation.

1.2.2.4. La fonction de sortie :

Elle calcule la sortie (S) du neurone en fonction de son état d'activation (A). Elle est généralement considérée comme la fonction identité : $S = g(A) = A = f(E)$.

1.2.2.5. La nature de la sortie du neurone :

Elle peut être binaire ou réelle.

1.2.3. Exemple du fonctionnement d'un neurone formel :

Afin de comprendre le principe de fonctionnement, nous allons présenter l'exemple classique de la fonction logique OU. Sa structure neuronale est représentée à la figure 4.

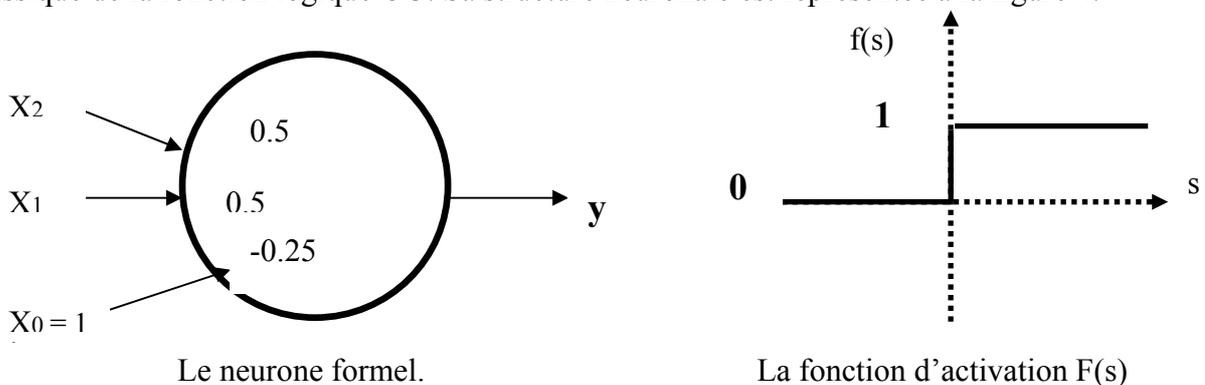


Fig.4 : La fonction logique et le neurone correspondant

Le poids attribué aux entrées X_1 et X_2 est 0.5 et le biais possède un poids de - 0.25. Cette forme nous donne :

- Si $X_1 = 0$ et $X_2 = 0$ donc $s = -0.25$ et $f(s) = 0$.
- Si $X_1 = 0$ et $X_2 = 1$ donc $s = 0.25$ et $f(s) = 1$.
- Si $X_1 = 1$ et $X_2 = 0$ donc $s = 0.25$ et $f(s) = 1$.
- Si $X_1 = 1$ et $X_2 = 1$ donc $s = 0.75$ et $f(s) = 1$.

Ce fonctionnement traduit parfaitement la fonction logique OU.

1.3. LES RESEAUX DE NEURONES ARTIFICIELS :

Le réseau de neurones artificiels est un ensemble de neurones formels connectés suivant une architecture bien définie afin de réaliser des tâches différentes. Ces réseaux sont introduits en s'inspirant du réseau biologique que constitue le cerveau humain.

Les huit composants principaux d'un réseau de neurones sont [6] :

- un ensemble de neurones ;
- un état d'activation pour chaque neurone (actif, inactif) ;
- une fonction de sortie pour chaque neurone ($f(S)$) ;
- un modèle de connectivité entre les neurones (chaque neurone est connecté à tous les autres, par exemple). C'est ce qu'on appelle la structure du connexionisme ;
- une règle de propagation pour propager les valeurs d'entrée à travers le réseau vers les sorties ;
- une règle d'activation pour combiner les entrées d'un neurone (très souvent une somme pondérée) ;
- une règle d'apprentissage ;
- un environnement d'opération (le système d'exploitation, par exemple).

Ces huit composants ont un rôle décisif dans le comportement des réseaux neuronaux. La variation d'un des paramètres change radicalement sa réponse.

1.3.1. La structure des connexions :

Les réseaux de neurones naturels sont organisés en couches et présentent un grand nombre de connexions. Les neurones d'une même couche sont reliés les uns aux autres ainsi qu'aux neurones des autres couches. Leur architecture est d'une très grande complexité et leur modélisation est difficile.

Les structures utilisées dans les réseaux formels sont très variées. Elles peuvent aller d'une connectivité totale où tous les neurones sont reliés les uns aux autres à une connectivité partielle où les neurones ne sont reliés qu'à leurs plus proches voisins :

1.3.1.1. Les réseaux entièrement connectés :

Chaque neurone a tous les bouclages possibles. Il est relié à tous les autres neurones et possède un retour sur lui-même (Fig.5).

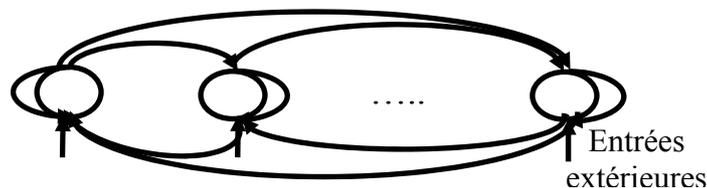


Fig.5 : Architecture générale d'un réseau entièrement connecté [4].

1.3.1.2. Les réseaux en couches :

Ces réseaux sont organisés en couches. Les neurones d'une même couche ne sont pas reliés entre eux. Ils reçoivent leurs entrées de la couche précédente et transmettent le résultat des traitements à la couche suivante (Fig.6). Les couches situées entre la couche d'entrée et la couche de sortie sont appelées couches cachées. Leur nombre est variable.

La modélisation multicouche des RNA permet de traiter des problèmes très complexes. Elle est apparue suite à l'impuissance marquée par le perceptron dans le domaine de la génération des sorties non linéaires (exemple de la fonction XOR) [7].

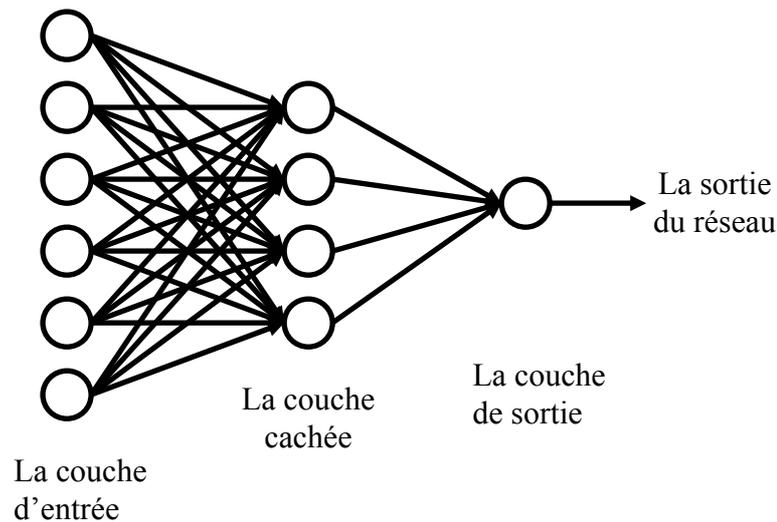


Fig.6 : Architecture générale d'un réseau multicouches.

1.3.2. La dynamique des connexions :

Les signaux transmis entre deux neurones sont pondérés en affectant à chaque connexion un coefficient dit *poids de la connexion* ou *poids synaptique* que le réseau modifie au fur et à mesure que de nouveaux exemples lui sont présentés afin de s'adapter aux situations désirées : c'est la phase d'apprentissage. L'apprentissage neuronal est la propriété la plus intéressante des réseaux neuronaux. Deux types existent : l'apprentissage supervisé et l'apprentissage non supervisé.

1.3.2.1. L'apprentissage supervisé :

L'apprentissage supervisé se fait en introduisant des couples d'entrées et leurs sorties désirées. Les performances du réseau sont déterminées par la suite, par l'intermédiaire d'un critère à optimiser.

L'intérêt de cet apprentissage est de trouver des poids synaptiques optimaux afin de ramener au réseau un comportement désiré à l'aide d'échantillons aléatoires. La réadaptation des poids peut se faire selon deux méthodes :

1.3.2.1.1. Data adaptive learning (Feed-Forward Network pattern) :

La réadaptation des poids se fait au passage de chaque exemple, ce qui rend le processus sensible à chaque exemple individuel. Par conséquent, elle est utile lors d'un apprentissage en temps réel. En revanche, l'apprentissage est sensible aux bruits.

1.3.2.1.2. Block adaptive learning (batch learning) :

La réadaptation des poids se fait après le passage de tous les exemples. Elle est calculée à partir de leur moyenne et est insensible aux bruits.

La validation des poids adaptés se fait suivant deux approches. La première est basée seulement sur la validité de la décision. Le superviseur doit juger l'exactitude de la décision du réseau. Cette approche permet la classification de chaque exemple de l'espace d'entrée. Elle est nommée : *apprentissage par pénalité ou récompense*. La deuxième approche concerne l'optimisation d'un critère appelé fonction coût. Elle traite en général les problèmes d'approximation de fonctions.

1.3.2.2. L'apprentissage non supervisé :

Ces réseaux sont souvent nommés *auto-organiseurs* ou bien à *apprentissage compétitif*. Certaines architectures de ces modèles sont basées sur les principes des structures vivantes : tels que la théorie de l'évolution et de la génétique. Les travaux ont débuté, depuis les années soixante (60), mais les travaux restaient limités en raison de l'absence d'un modèle mathématique exacte de ces réseaux. Leur comportement mathématique fut déterminé à partir des années quatre-vingts (80) grâce aux travaux de Hopfield, de M. Cohen et de S. Grosberg. Ces résultats ont donné un grand élan à la recherche dans ce domaine. [4]

La réadaptation des poids d'éléments d'entrée est autonome. L'apprentissage qui se fait par calcul de distances [4], nécessite uniquement les valeurs des entrées. Ce type d'apprentissage est très utile dans les applications de classification. Dans la littérature, plusieurs architectures existent et les plus évoquées sont :

1.3.2.2.1. Les réseaux compétitifs :

Ce sont des réseaux fortement couplés, chaque neurone est lié au neurone de la même couche ainsi qu'aux neurones des autres couches. La compétition quasi-permanente des neurones est à l'origine de leur appellation. Ces réseaux servent à classer et rassembler les données dans des catégories.

1.3.2.2.2. Les réseaux auto-organiseurs de Kohonen :

Les réseaux dits de *Kohonen* sont inspirés des constatations biologiques sur ce qu'on appelle les aires de projection primaires du cortex (enveloppe extérieure du cerveau humain) [6]. Il a été prouvé que les connaissances préalables se répartissent sur des zones de cortex, de telle sorte que les perturbations (stimulus) extérieures proches de ces zones provoquent leur connaissance et apprentissage. Kohonen a réussi à projeter des stimulus voisins situés dans un espace géométrique beaucoup plus simple. Généralement mono ou bidimensionnel, en conservant au mieux la topologie. Sa méthode est très connue sous le nom de la carte auto-organisatrice de Kohonen (Self Organisation Maps : SOM). Elle représente un type particulier des réseaux compétitifs qui définit un voisinage spatial pour chaque neurone de sortie. Le réseau de Kohonen est constitué d'une couche compétitive unique : *la couche topologique*.

1.4. LES TECHNIQUES D'APPRENTISSAGE ADAPTEES A LA CLASSIFICATION :

La classification et le diagnostic sont deux domaines où les réseaux de neurones ont connu des progrès phénoménaux. Durant ces dernières années, les logiciels de classification et de diagnostic à base des réseaux neuronaux ont envahi le marché. Ces logiciels utilisent des bases de données assez complètes et significatives, et donnent des résultats très satisfaisants.

L'apprentissage supervisé est majoritairement utilisé dans les problèmes de classification car dans la plupart des cas, les classes sont connues. Le programmeur attribue à chaque exemple d'entrée du système, la classe à laquelle il fait partie. Lorsque les classes ne sont pas connues a priori, l'apprentissage non supervisé est alors utilisé.

Historiquement, c'est les réseaux à apprentissage supervisé qui sont apparus en premier dans le modèle de Adaline : Adaptive Linear Element, (B. Widrow : 1960). L'apprentissage de ce réseau se fait par l'algorithme de la rétro-propagation. Le modèle d'Adaline a montré une incapacité de classification des systèmes non linéaires. En 1985 cette difficulté est surmontée avec l'apparition des algorithmes de rétro-propagation du gradient qui sont adaptés au réseau de neurones multicouches. Cette découverte a permis de réaliser des systèmes de classification non linéaires d'entrée/sortie sur un réseau en décomposant cette fonction en une suite d'étapes linéairement séparables. De nos jours, les réseaux multicouches et la rétro-propagation de gradient reste le modèle le plus étudié et le plus productif au niveau des applications notamment dans le diagnostic et la classification.

1.4.1. La technique de la descente de gradient :

Cette méthode consiste à déterminer les adaptations adéquates du poids avec un minimum d'erreur. Pour cela le gradient de la fonction est calculé par rapport au vecteur poids W . L'erreur quadratique instantanée est donnée par la relation :

$$\xi_t = \frac{1}{2} (e)^2 = \frac{1}{2} (d - y)^2 \quad \text{où : } d \text{ est la sortie désirée et } y \text{ est la sortie du réseau.}$$

La progression dans le sens opposé du gradient se fait suivant la relation :

$$\Delta W = -\alpha \frac{\partial \xi_t}{\partial W} = eX \quad \text{où } X \text{ est le vecteur d'entrée, } W \text{ est le vecteur des poids, } e \text{ est}$$

l'erreur quadratique et α est le taux d'apprentissage.

Les poids adaptés sont donnés par la relation : $W_{ij}(t+1) = W_{ij}(t) + \alpha (d_i - y_i) x_i$

1.4.2. L'algorithme de rétro-propagation (Back-Propagation):

Pour appliquer cet algorithme, la fonction du transfert du neurone doit être continue, non-linéaire et différentiable [8]. L'algorithme comprend sept étapes qui sont :

- 1- initialisation des poids à des valeurs aléatoires de faible grandeur généralement entre -1 et 1;
- 2- lecture d'un premier exemple d'apprentissage $(E, d)^1$ dans la base de données ;
- 3- calcul de la sortie obtenue S^1 et l'erreur Er^1 , en utilisant la relation :

$$\xi_t = \frac{1}{2} (e)^2 = \frac{1}{2} (d - y)^2$$

4- calcul du gradient de l'erreur et des taux de réadaptation des poids inter couches :

$$\frac{\partial \xi_t}{\partial W_{ij}^l} = \alpha \delta_{wi}^l f'(x_i^l(t))$$

où : α est le taux d'apprentissage.

δ_{wi}^l est le taux de réadaptation partielle des poids. Il est défini par : $\delta_{wi}^l = (d_i^l - y_i^l)$

pour les neurones de la couche de sortie et par : $\delta_{wi}^l = \sum_{i=1}^{n+1} \delta_{w_{si}^l} W_{ij}^l$ pour les neurones des autres couches ;

5- si l'apprentissage est de type Data Learning, la réadaptation des poids suit la relation :

$$\Delta W_{ij}^l(t) = [\Delta W_{ij}^l(t)]_p$$

si s'agit d'un Block Learning :

. si $p \neq M$ (M étant le nombre d'exemples d'entraînement) : retourner à 1.

. sinon : $\Delta W_{ij}^l(t) = \sum_{p=1}^{p=M} [\Delta W_{ij}^l(t)]_p$

6- réadaptation des poids ;

7- tant que l'erreur est trop importante : retourner à l'étape 2 (pour traiter l'exemple suivant).

Le taux d'apprentissage α joue un rôle important dans la régulation de la vitesse de convergence de la méthode. En général, sa valeur est prise entre 0.1 et 0.7 [4]. Selon la détermination de ce paramètre, deux variantes améliorent sensiblement l'algorithme d'apprentissage.

1.4.2.1. La rétro-propagation avec Momentum :

C'est une méthode efficace, qui assure une convergence rapide. Elle consiste à utiliser les changements précédents des poids pour la réadaptation des poids actuels. Les nouveaux poids sont donnés par la relation : $W_{ij}^l(t+1) = W_{ij}^l(t) + \alpha \Delta W_{ij}^l(t) + \alpha \Delta W_{ij}^l(t-1)$. Le terme : $\alpha \Delta W_{ij}^l(t-1)$ est appelé *le Momentum*.

Actuellement les algorithmes de la rétro-propagation sont pratiquement tous utilisés avec Momentum.

1.4.2.2. La rétro-propagation à taux d'apprentissage adaptatif :

Si le taux d'apprentissage est fixe le long du programme deux problèmes peuvent surgir. Si α est très petit, la convergence est très lente alors que si elle est très grande l'algorithme risque de diverger (le système oscille). La méthode d'apprentissage adaptative consiste à varier le taux d'apprentissage selon l'erreur calculée pour chaque étape.

Cette technique est très simple. Elle augmente la rapidité de la convergence, et met en œuvre un algorithme qui a pour rôle de contrôler l'erreur d'entraînement à chaque étape.

1.4.3. Les limites de la rétro-propagation :

La mise au point de l'algorithme de la rétro-propagation en 1985, a amplifié de façon phénoménale l'utilisation des réseaux neuronaux multicouches. Cet algorithme est appelé perceptron multicouches. L'algorithme de la rétro-propagation et ses variantes présentent des caractéristiques très intéressantes qui servent de manière efficace dans le domaine de la classification. Néanmoins ces algorithmes ont des limites.

Comme tous les autres réseaux, la convergence n'est pas toujours assurée. De plus, le temps de calcul est important et le nombre d'exemples d'entraînements ne peut pas être prédéterminé. D'autres structures ont été proposées, parmi lesquelles la méthode de Levenberg-Marquardt.

1.4.4. La méthode de Levenberg-Marquardt :

La méthode de Levenberg-Marquardt est très souvent utilisée en raison de sa rapidité de convergence. Elle utilise la fonction énergie donnée par la relation :

$$E(k+1) = E(k) + \dot{E}(k) \Delta W(k) + \frac{1}{2} (\Delta W(k))^T \ddot{E}(k) \Delta W(k)$$

Elle consiste à minimiser cette fonction, ce qui revient à écrire :

$$\frac{\partial E(k+1)}{\partial \Delta W(k)} = 0 \text{ où } \Delta W(k) \text{ est le vecteur global des taux d'apprentissage.}$$

La dérivée de l'erreur par rapport à ses poids est donnée par la relation :

$$\frac{\partial E(k+1)}{\partial \Delta W(k)} = F_i^l = \sum_{p=1}^K \sum_{i=1}^{nL} (d_i^p - y_i^l) \left[-\frac{\partial y_i^p}{\partial W_i^l} \right]$$

où : F_i^l sont les vecteurs des dérivées de l'erreur à la sortie du réseau par rapport à ses vecteurs poids W_i^l . Ils sont rassemblés dans un vecteur Jacobien F ;

L est le nombre de couches cachées et $l = 1 \dots L$;

K est le nombre d'exemples d'entraînement ;

nL est le nombre de neurones de la dernière couche.

Afin de trouver l'optimum, on utilise la relation : $H \Delta W = -F$ où H représente la matrice hessienne de la fonction d'erreur en sortie du réseau.

Pour un système de 'n' variables, cette matrice est symétrique. Elle est donnée par :

$$H = \frac{\partial F}{\partial W} = \begin{bmatrix} \frac{\partial F_{11}}{\partial W_1} & \dots & \frac{\partial F_{1n}}{\partial W_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{n1}}{\partial W_1} & \dots & \frac{\partial F_{nn}}{\partial W_n} \end{bmatrix}$$

$$\text{avec : } \frac{\partial F_{ij}}{\partial W_i} = \frac{\partial^2 E(W)}{\partial W_i \partial W_j} = \frac{\partial}{\partial W_i} \left(\frac{\partial E(W)}{\partial W_j} \right)$$

Cette équation est un système linéaire dont la résolution peut être faite par la méthode de Gauss-Newton [9]. Toutefois, son calcul est très lent. La méthode de Levenberg-Marquardt remplace le calcul du Hessien par une approximation numérique donnée par [9] :

$$H = F F^T + \lambda \Omega \quad \text{où : } F F^T \text{ est une matrice semi-positive ;}$$

$$\lambda \quad \text{est un coefficient positif ;}$$

$$\Omega \quad \text{est une matrice choisie.}$$

Plusieurs possibilités sont envisageables pour les valeurs de Ω et de λ . Le choix doit être fait judicieusement. La matrice Ω est généralement une matrice identité diagonale ou une matrice diagonale dont les éléments sont égaux à ceux de la matrice $F F^T$.

L'algorithme de détermination optimale de ΔW :

- 1- Initialisation des poids synaptiques W_0 et choix de $\lambda < 1$;
- 2- Calcul de l'erreur quadratique $E(k)$ après le passage de tous les exemples d'apprentissages ;
- 3- Calcul de F :
$$\frac{\partial E(k+1)}{\partial \Delta W(k)} = F_i^l = \sum_{p=1}^K \sum_{i=1}^{nL} (d_i^p - y_i^l) \left[-\frac{\partial y_i^p}{\partial W_i^l} \right]$$
- 4- Calcul de H : $H = F F^T + \lambda \Omega$
- 5- Détermination de ΔW par résolution du critère d'optimisation : $H \Delta W = -F$
Sa résolution peut se faire par l'utilisation de la méthode de GAUSS-SIEDEL [9].

D'autres variantes visant à rendre cet algorithme plus efficace existent. L'amélioration se base sur le choix de la valeur de λ qui conduit vers un minimum local.

La méthode de Levenberg-Marquardt converge rapidement. En contre partie, elle est sensible au bruit et par conséquent, elle n'est conseillée que pour un apprentissage en Bloc Learning [7].

1.5. L'APPLICATION DES RESEAUX DE NEURONES ARTIFICIELS EN TECHNIQUE DE LA HAUTE TENSION :

Depuis 1994, les RNA sont appliqués à grande échelle dans le domaine de la haute tension, surtout en ce qui concerne la prévision des tensions disruptives des systèmes électriques. Une nouvelle application consiste à les introduire dans la marche de diagnostic et comme outil d'aide à la décision.

1.5.1. Les décharges partielles et les tensions disruptives :

En 1994, A.A. Mazroua et al avaient utilisé trois sortes de RNA afin de déterminer les tailles d'impulsion des décharges partielles produites par différentes sources. L'application d'un réseau compétitif, d'un réseau learning vector quantization (LVQ) et d'un réseau multicouches dont l'algorithme d'apprentissage est la rétro-propagation avait donné des

résultats proches. Cependant le réseau LVQ avait montré une bonne performance dans la distinction entre les décharges partielles produites par la même cavité, par rapport à celles nées d'une arborescence électrique. Pour reconnaître les tailles d'impulsion, T. Kexioug et Z. Deheng avaient utilisé, en 1997, des RNA pour la reconnaissance des modèles des décharges partielles tridimensionnelles. Ils cherchaient la structure la plus adaptée à cette application. Pour cela ils avaient étudié l'influence du nombre de neurones dans la couche d'entrée et dans la couche de sortie. En 1999, Y. Tian avait présenté une méthode originale pour identifier des décharges partielles en utilisant les RNA. Cette méthode se base sur la mesure des émissions acoustiques, les RNA utilisés étaient : la rétro-propagation, le réseau de Kohonen (SOM) et le réseau LVQ. En 2001, Y. Kawaguchi et T. Shimizu avaient présenté un travail sur l'application des RNA à la reconnaissance des décharges partielles dans les gaz isolants. La même année, à la CEIDP L. Mokhnache et A. Boubakeur [10] ont utilisé les RNA à base radiale à fonction gaussienne (RBGF) entraînés par la ROM, pour la prédiction de la tension disruptive d'intervalle d'air pointe-barrière-plan avec différentes positions de la barrière et différentes tensions appliquées. L'utilisation de ce type de réseaux de neurones semblait très prometteuse. La même équipe a présenté en 2003, d'autres travaux utilisant les RNA dans la prédiction de la tension disruptive du même système en fonction de la longueur des intervalles et de la largeur de la barrière. Le RNA utilisé fut la carte auto-organisatrice de Kohonen [11] et [12].

Les RNA ont également été introduits dans le diagnostic par les décharges partielles. En effet, les premiers travaux remontent à 1995, quand K. Engel et D. Peier utilisaient les décharges partielles et leur développement pour la détermination du type de défauts. Le réseau de neurones utilisé fut la carte auto-organisatrice de Kohonen. Ces mêmes réseaux furent utilisés, en 1997, pour classifier les décharges partielles, dans un système développé par S. Happe et H.G. Kranz qui servait à supprimer en temps réel des signaux d'impulsions sonores qui sont nuisibles et causent une mauvaise interprétation des activités des décharges partielles.

1.5.2. Les systèmes d'isolations électriques :

Les équipements électriques nécessitent des systèmes d'isolation très pointus qui requièrent un haut niveau de spécialisation et ne tolèrent aucune erreur. De plus, les décisions sur l'état des isolations nécessitent une haute qualification et une longue expérience car il y a presque toujours des conflits entre les données, ainsi que des cas d'indécisions. Pour permettre à des ingénieurs inexpérimentés de s'approcher de la qualité de diagnostic des experts, il a fallu trouver un système d'aide à la décision. Les RNA sont des outils particulièrement adaptés. C'est pour cela que de nombreuses études sont entreprises pour évaluer les apports des réseaux de neurones avant leur mise en œuvre de façon opérationnelle.

1.5.2.1. L'isolation des transformateurs :

Le fonctionnement sans défaillance des transformateurs de puissance (de tension ou de courant), des disjoncteurs et des sectionneurs dépend en grande partie des qualités diélectriques des huiles isolantes. En plus de leur rôle principal d'isolation, elles doivent évacuer l'énergie dissipée par les parties actives sous forme de chaleur. Ces deux fonctions sont vitales pour un transformateur car toute défaillance d'ordre électrique se traduit par un claquage dont les conséquences sont toujours graves et irréversibles. Une surveillance des caractéristiques isolantes est donc nécessaire afin de réduire ces risques. Elle s'effectue par extrapolation des caractéristiques de l'huile.

En 1995 M.E. Ghourab a utilisé les RNA pour l'identification de l'effet de moisissure sur le spectre de la polarisation des isolateurs de transformateur. L'algorithme utilisé était celui de la rétro-propagation. En 2000, X.H. Jin a publié un travail dans lequel il utilisait les RNA pour la classification des défauts provoqués par des décharges partielles sur les isolations des transformateurs. Il avait utilisé pour cela deux types de RNA : la : rétro-propagation et le LVQ. Tous les deux avaient montré leur capacité de généralisation. Ce travail ne fut pas le seul, d'autres avaient été publiés auparavant dont celui, en 1995, de T. Nogami. Il a utilisé les RNA dans la pré-détection des anomalies dans l'huile de transformateur. Un autre travail fut publié par L. Tang et M. R.Raghvveer. Il traite de la classification et du vieillissement des transformateurs [4]. Malheureusement le type des RNA adoptés n'avait pas été mentionné. En 2002, L. Mokhnache et A. Boubakeur ont publié un article sur l'application des variantes de la rétro-propagation dans la classification des huiles de transformateurs [13]. En 2003, ils ont publié un autre travail où une carte auto-organisatrice et des réseaux bayesiens furent utilisés dans la classification des huiles de transformateurs [14].

L'application des RNA ne se limite pas au domaine du diagnostic et de la classification. En 2001 L. Mokhnache, A. Boubakeur, B. Ould Noureddine et M.A.R. Bedja ont employé les RNA pour prédire le vieillissement thermique d'huiles de transformateur utilisées par la SONELGAZ. Ils ont utilisé des réseaux RBF entraînés par la ROM. Ceci leur a permis de prédire jusqu'à 3000^h [15]. Ce travail ne fut pas le seul, un ensemble de publications a traité de l'application des RNA dans la prédiction du vieillissement thermique des transformateurs [16], de la comparaison entre différents RNA utilisés dans le diagnostic des transformateurs [17], de l'étude de l'influence des paramètres d'apprentissage sur quelques RNA supervisés utilisés dans le diagnostic des huiles de transformateur [18] ainsi que de l'étude de quelques réseaux neuronaux supervisés et non supervisés utilisés dans la classification et le diagnostic des huiles de transformateurs [19].

1.5.2.2. L'isolation des câbles électriques de haute tension :

Actuellement, le contrôle du bon état des isolations des câbles électriques est une clé pour assurer la permanence du service. Les câbles électriques sont souvent isolés en polychlorure de vinyle PVC dont le nom découle de l'appellation anglaise : « Poly Vinyl Chloride ». Le PVC appartient à la grande famille des polymères, qui sont très utilisés dans l'isolation électrique. Ces matériaux isolants se dégradent sous l'action des contraintes auxquelles ils sont soumis pendant la mise en service des équipements dans lesquels ils sont utilisés. Parmi ces contraintes, nous trouvons : la température, le champ électrique, les radiations,...etc. Vu leur importance dans ce secteur, elle sont concernées par l'application des RNA. De nombreuses publications ont tenté de prédire la durée de vie des PVC soumis à différentes contraintes de fonctionnement. Ainsi, en 2000, L. Mokhnache, A. Boubakeur, A.L. Nemmour et A. Khelifi ont publié, à la conférence internationale Quality, Reability & Maintenance (QRM) un travail où ils utilisaient un réseau à fonction de base radiale de noyau Gaussien (RBF) entraîné par la ROM, pour prédire l'influence du vieillissement du PVC sur ses caractéristiques mécaniques. Ils ont même pu prédire la durée de vie du PVC en fonction du temps de vieillissement [20]. Un autre travail fut publié, en 2003, par L. Mokhnache et A. Boubakeur. Il traitait de l'application des réseaux de neurones de Kohonen pour prédire le vieillissement thermique du PVC [21]. De nombreuses publications ont vu le jour depuis [22], [23] et [24].

1.5.3. Les autres applications :

En réalité, l'application des RNA répond à plusieurs domaines de la haute tension. En 1995, P.S. Ghosh et al [4], avaient modélisé un arc électrique d'un isolant pollué en utilisant un réseau de neurones entraîné par l'algorithme d'apprentissage de type « Resilient Propagation Algorithm (RPROP) ». Cet algorithme prend en considération la topologie locale de la fonction "erreur", de telle sorte que la modification des poids dépend des signes des dérivées de E et non pas de leurs amplitudes. La règle d'apprentissage est celle de Manhattan. En 1999, S.Chowdhary avait présenté un travail dans le Symposium International de Haute tension [4], sur la détermination de la tension de rupture du SF₆ sous un champ non-uniforme. Pour cela il avait utilisé un RNA entraîné par la Project Pursuit Learning Network (PPLN). Ce type de réseaux neuronaux se construit au fur et à mesure que l'apprentissage se fait. Il commence par un neurone caché. En ajoutant un autre neurone, l'apprentissage ne se fait pas pour tout le réseau, mais seulement pour le nouveau neurone. Ce réseau a de bonnes capacités de généralisation [4]. Dans la même année, I. Nunes Da Silvan et A. Da Suza ont déterminé à l'aide des RNA, l'intensité du flux électrique et la valeur critique de la tension disruptive dans les sous stations. La détermination se fit en tenant compte des différents facteurs atmosphériques et géométriques. Le RNA utilisé se composait de deux réseaux en cascade. Le premier calculait la tension disruptive, qui servait comme entrée au second qui à son tour, déterminait l'intensité du champ électrique entre les jeux de barres. Le but était de dimensionner des sous stations de 138 kV [4].

Chapitre II

Introduction à la logique floue

2.1. HISTORIQUE :

Les prémices de la logique floue sont apparues dans les années 20, avec les premières approches de chercheurs américains à partir du concept d'incertitude de Heisenberg. Les paradoxes étant interdits par les axiomes de la logique classique booléenne, les physiciens leur ont attribués la valeur de vérité $\frac{1}{2}$ et ont ainsi introduit une troisième valeur dans le système logique bivalent $\{0,1\}$.

Au début des années 30, le logicien polonais Jan Lukasiewicz a étendu le système logique avec trois valeurs à tous les nombres rationnels entre 0 et 1. Il a défini la logique floue comme « une logique qui utilise la fonction générale de vérité, laquelle associe à une affirmation un niveau de vérité qui peut prendre toutes les valeurs entre 0 (faux) et 1 (vrai) » [25].

Dans les années 1930, Max Black a appliqué la logique floue aux ensembles d'éléments ou de symboles et a dessiné la première fonction d'appartenance d'un ensemble flou.

En 1965, le professeur Lotfi A. Zadeh, de l'université de Berkeley (Californie), déclare : « un contrôleur électromécanique doté d'un raisonnement humain serait plus performant qu'un contrôleur classique » [26]. Il publie un article où il introduit le terme « fuzzy » dans la littérature technique et donne naissance au concept d'ensemble flou [27]. Le professeur tentait de modéliser les phénomènes sous forme floue afin de remédier aux limitations dues aux incertitudes des modèles classiques à équations différentielles. Il a réalisé depuis, des avancées majeures dans le domaine et a été rapidement accompagné par de nombreux chercheurs développant des travaux théoriques.

En 1974, le docteur Mamdani, de l'Université de Londres, a introduit la commande floue dans la régulation d'un processus industriel, en réalisant un contrôleur flou expérimental sur un moteur à vapeur [28]. Le problème sur lequel il s'était penché, avait la réputation d'être difficile et les résultats obtenus étaient très encourageants.

En 1978, la société danoise F.L.Smidth & Co réalise la première véritable application industrielle de la logique floue : le contrôle d'un four à ciment.

A partir des années 80, les recherches se sont arrêtées en Europe et ont été reprises au Japon où plusieurs applications ont commencé à émerger.

En 1985, la logique floue a été introduite au Japon par le chercheur M. Sugeno. Les sociétés japonaises ont très vite compris son avantage à la fois technique et commercial. La recherche n'était pas uniquement théorique mais également appliquée. A la fin des années 80, ce fut une véritable 'explosion du flou'. Les Japonais ont mis sur le marché plus de 2000 produits flous destinés au grand public, notamment des appareils électroménagers et du matériel audio-visuel. Dans l'industrie, l'application de la logique floue s'est étendue au traitement des eaux, aux grues portuaires, aux métros, aux systèmes de ventilation et de climatisation ainsi qu'à d'autres domaines très différents tels que la finance ou le diagnostic médical.

A partir de 1990, apparaissent de nombreuses applications en Allemagne et à moindre échelle aux Etats Unis d'Amérique.

De nos jours, une intense activité de recherche a débuté afin de combiner la logique floue aux algorithmes génétiques et aux réseaux de neurones.

2.2. LES BASES GENERALES DE LA LOGIQUE FLOUE :

2.2.1. La logique booléenne :

La logique booléenne se distingue un univers de discours réduit à l'ensemble $\{0, 1\}$. Ses constantes et ses variables ne peuvent prendre que ces deux valeurs. Les valeurs booléennes ne représentent pas des nombres réels mais plutôt l'état d'une variable. La valeur 0 est synonyme de faux et 1 de vrai. Les quatre opérations élémentaires de la logique binaire sont :

- la négation ou l'inversion logique (NON) qui correspond à la complémentation ;
- la multiplication logique (ET) qui correspond à l'intersection ;
- l'addition logique (OU) qui correspond à l'union ;
- l'implication (\Rightarrow) définie par : $(A \Rightarrow B) = (\text{NON}(A) \text{ OU } B)$.

En logique classique, le processus de raisonnement est basé sur les tautologies. Nous n'en mentionnerons que deux :

- *le modus tollens* : $(A \text{ ET } (A \Rightarrow B)) \Rightarrow B$.

Il permet d'affirmer à partir des prémisses (x est A) et $(A \Rightarrow B)$ que (x est B). Par exemple : (Socrate est un homme ET (être homme \Rightarrow être mortel)) \Rightarrow Socrate est mortel.

- *le modus ponens* : $((A \Rightarrow B) \text{ ET } \bar{B}) \Rightarrow \bar{A}$.

2.2.2. La théorie des ensembles flous :

La notion d'ensemble est une des notions fondamentales des mathématiques mais elle ne permet cependant pas de rendre compte de situations simples et assez fréquentes car dans la théorie des ensembles classiques, un élément appartient ou n'appartient pas à un ensemble donné.

Par exemple, il est aisé de définir parmi les personnes, l'ensemble des hommes. Par contre, il est impossible de définir l'ensemble des hommes vieux puisque les hommes ne deviennent pas vieux du jour au lendemain. Ils vieillissent graduellement.

C'est pour tenir compte de ce genre de situations que Zadeh a introduit la notion d'ensemble flou. Elle repose sur la notion d'appartenance partielle. Chaque élément x peut :

- ne pas appartenir à l'ensemble flou A ;
- appartenir un peu à A ;
- appartenir assez à A ;
- appartenir fortement à A ;
- appartenir totalement à A .

Le contour de chaque ensemble flou n'est pas "rigide" ou "net" mais "graduel" ou "flou". (Fig.11)

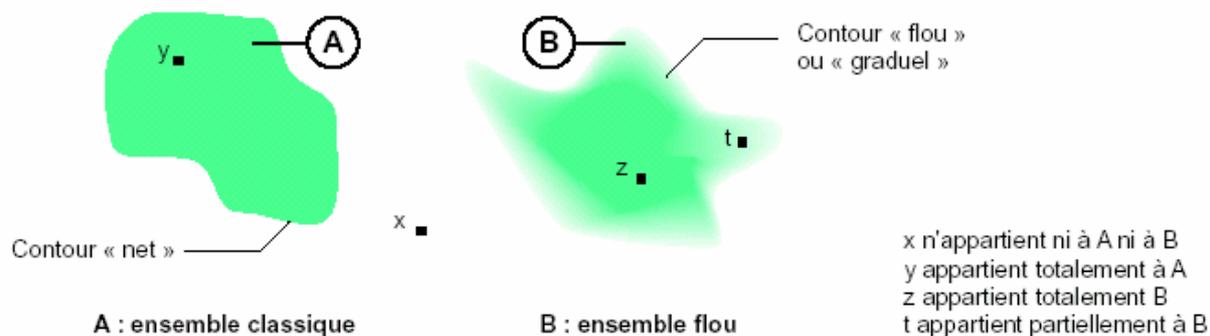


Fig.11 : comparaison entre un ensemble classique et un ensemble flou [29].

2.2.2.1. La fonction d'appartenance :

La fonction d'appartenance correspond à la notion de fonction caractéristique en logique classique. Elle associe à chaque élément de l'ensemble flou son degré de compatibilité avec le concept qui est représenté par cet ensemble [25]. Dans la théorie des ensembles classiques, l'appartenance à un ensemble ne peut avoir que la valeur 0 (n'appartient pas) ou la valeur 1 (appartient) alors qu'en logique floue, elle varie entre ces deux valeurs.

Pour bien comprendre cette notion, considérons l'exemple classique de la classification d'une population donnée suivant sa taille en trois catégories : petite, moyenne et grande.

En logique classique, nous définissons le premier ensemble comme étant celui des personnes qui font moins de 1,50 m, le second comme étant celui des personnes qui font entre 1,50 m et 1,90 m et le dernier comme celui des personnes qui font plus de 1,90 m. La fonction caractéristique de l'un des ensembles est donc égale à 1 pour toutes les tailles qui sont dans l'intervalle qui le définit et 0 hors de cet intervalle. (Fig. 12)

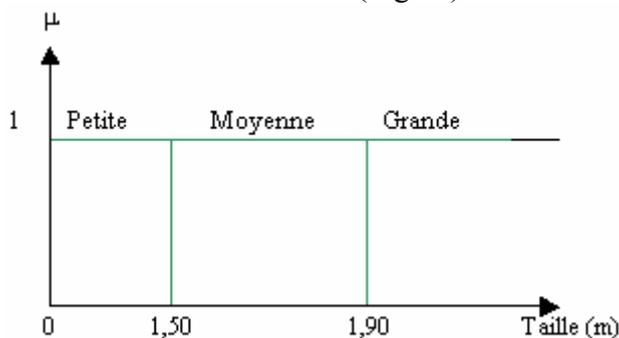


Fig.12 : fonctions caractéristiques des ensembles : "petite", "moyenne" et "grande".

Cette logique est très éloignée du raisonnement humain qui évalue la taille sans faire une distinction aussi tranchée entre deux catégories. Il est, en effet, insensé de dire qu'une personne de 1,89 appartient à l'ensemble des personnes de "taille moyenne" alors qu'une autre qui fait à peine 2 cm de plus appartiendrait à celui des personnes de "grande taille".

Les ensembles flous des personnes "petites", "moyennes" et "grandes" seront par contre définis par des fonctions d'appartenance qui font correspondre à chaque taille possible un degré d'appartenance à l'ensemble flou compris entre 0 et 1. (Fig.13)

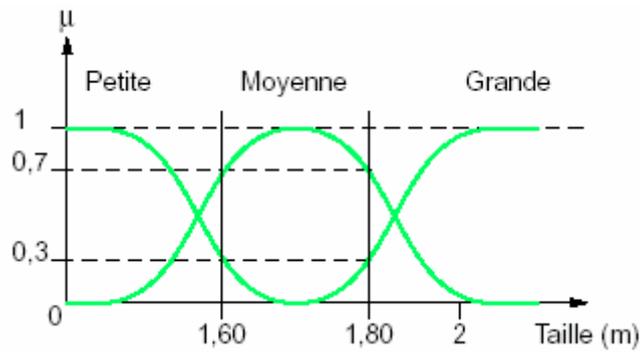


Fig.13 : fonctions d'appartenance des ensembles flous : "petite", "moyenne" et "grande" [29].

Cette représentation est beaucoup plus proche du raisonnement humain, car elle permet de faire intervenir des notions telles que : "assez petite", "plutôt grande" ou "très grande". Une personne de 1,80 m appartient à l'ensemble "taille grande" avec un degré 0,3 et à l'ensemble « taille moyenne » avec un degré de 0,7.

2.2.2.2. Les différentes formes de fonctions d'appartenance :

Les fonctions d'appartenance peuvent théoriquement prendre n'importe quelle forme, toutefois, les plus usuelles sont les suivantes [30] :

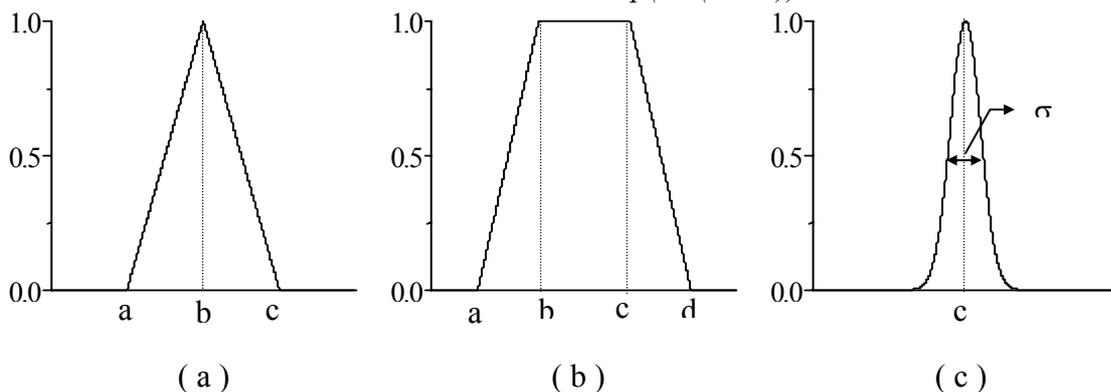
2.2.2.2.1. La fonction triangulaire : $\mu(x) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$; (Fig.14a)

2.2.2.2.2. La fonction trapézoïdale : $\mu(x) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right)$; (Fig.14b)

2.2.2.2.3. La fonction gaussienne : $\mu(x) = \exp\left(-\frac{(x-c)^2}{2\sigma}\right)$; (Fig.14c)

2.2.2.2.4. La fonction en forme de cloche : $\mu(x) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}$. (Fig.14d)

2.2.2.2.5. La fonction sigmoïde : $\mu(x) = \frac{1}{1 + \exp(-a(x-c))}$; (Fig.14e et f)



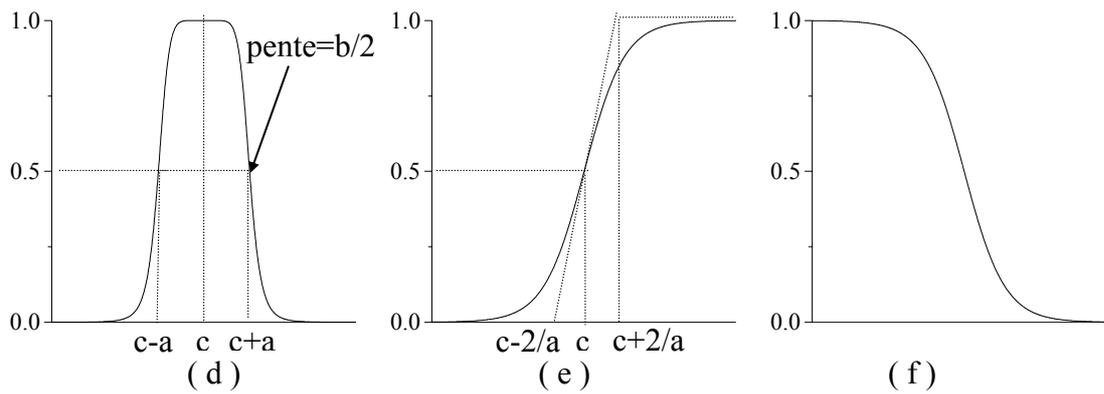


Fig.14 : les formes usuelles des fonctions d'appartenance [31].

2.2.2.3. Les ensembles flous :

Un ensemble flou est une extension de l'ensemble classique. C'est "une collection telle que l'appartenance d'un élément quelconque peut prendre toutes les valeurs comprises entre 0 et 1" [27]. Il est défini par des couples ordonnés : $A = \{ (x, \mu_A(x)) \mid x \in U \}$ où $\mu_A(x)$ la fonction d'appartenance et U l'univers de parcours.

Nous pouvons également utiliser la notation :

$$A = \sum_{i=1}^n \frac{\mu_A(x_i)}{x} \quad \text{si } U \text{ est discret}$$

$$A = \int_x \frac{\mu_A(x_i)}{x} \quad \text{si } U \text{ est continu}$$

2.2.2.4. Les opérateurs de la logique floue :

2.2.2.4.1. Les opérations algébriques :

2.2.2.4.1.1. La multiplication cartésienne :

Le produit cartésien de deux ensembles flous A et B est l'ensemble flou C défini par : $C = A \times B = \{ \mu_C(c) / c \mid c = (a, b) \text{ où } a \in A, b \in B \text{ et } \mu_C(c) = \min \{ \mu_A(x), \mu_B(x) \} \}$.

2.2.2.4.1.2. Le produit algébrique :

La multiplication algébrique des deux ensembles flous A et B est l'ensemble flou C tel que : $C = AB = \{ \mu_A(x)\mu_B(x) / x \mid x \in A, x \in B \}$.

2.2.2.4.1.3. L'exposant :

Elever un ensemble flou à la puissance α est un cas spécial de la multiplication algébrique. Le résultat est l'ensemble flou : $A^\alpha = \{ (\mu_C(x))^\alpha / x \mid x \in A \}$.

2.2.2.4.1.4. La concentration et la dilatation :

La concentration d'un ensemble flou A est exprimée par : $CON(A) = A^2$ et sa dilatation est donnée par : $DIL(A) = A^{0,5}$.

Ces deux opérations sont propres aux ensembles flous. Elles n'ont pas d'équivalent dans la théorie des ensembles classiques. Elles ont pour effet la réduction (pour la concentration) et l'augmentation (pour la dilatation) de la fonction d'appartenance. (Fig.16)

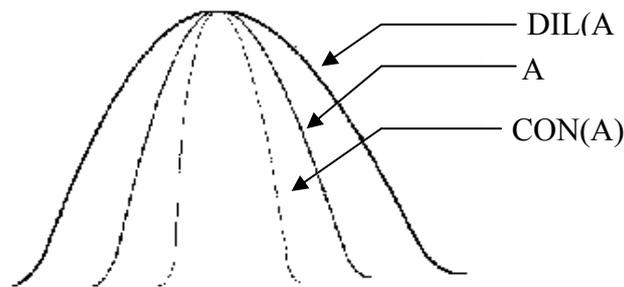


Fig.16 : Effet de la concentration et de la dilatation sur la fonction d'appartenance [7].

2.2.2.4.1.5. La somme algébrique :

La somme algébrique des deux ensembles flous A et B donne l'ensemble flou C définie par : $C = A + B = \{ \mu_C(x) / x \mid x \in A, x \in B, \mu_C(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x) \}$.

2.2.2.4.2. Les opérateurs linguistiques :

Egalement appelés *modificateurs des ensembles flous*, ces opérateurs sont ceux utilisés dans le langage humain (très, près de, autour de, presque, plus ou moins, ...). Leur définition est assez subjective. Il existe donc une infinité d'opérateurs proposés par différents auteurs (Fig.17) [25].

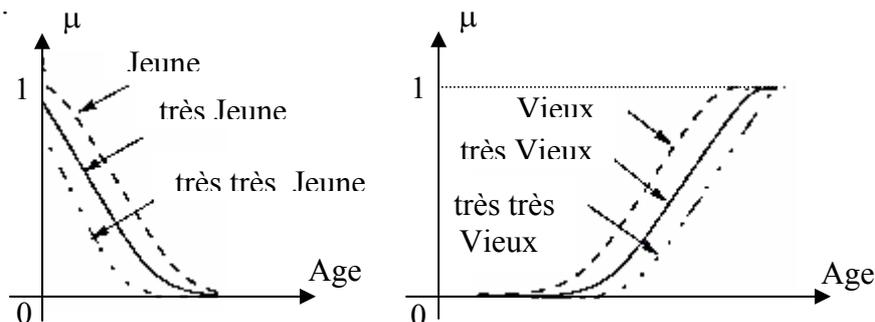


Fig.17 : l'opérateur linguistique 'très' réalisé par la concentration [25].

2.2.2.4.3. Les opérateurs logiques :

Ils permettent d'écrire des combinaisons logiques entre les notions floues et de faire des calculs sur leurs degrés de vérité. Les trois opérations de base utilisées pour la manipulation des ensembles classiques peuvent être généralisées aux ensembles flous. Cette généralisation n'est pas unique. Il existe de nombreuses variantes dans ces opérateurs.

2.2.2.4.3.1. La complémentation :

La complémentation d'un ensemble flou A dans un ensemble I est définie de la même manière que celle d'un ensemble classique. Sa fonction d'appartenance est définie par : $\text{non}(\mu_A(x)) = \mu_{\bar{A}}(x) = 1 - \mu_A(x)$. (Fig.18)

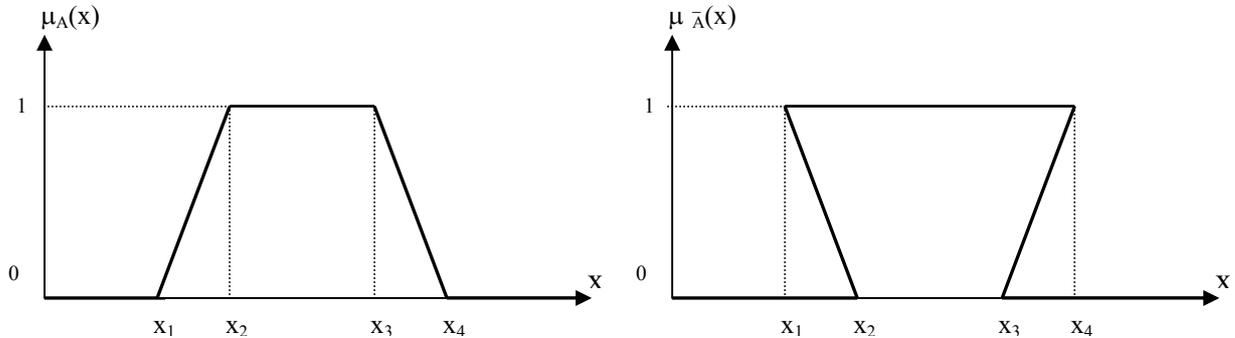


Fig.18 : l'opérateur NON.

2.2.2.4.3.2. L'intersection :

L'intersection des ensembles flous A et B est définie par une norme triangulaire (t-norme) :

$$\begin{aligned} ET : [0, 1] \times [0, 1] &\rightarrow [0, 1] \\ (x, y) &\rightarrow z = x \text{ ET } y \end{aligned}$$

La norme ET possède les propriétés suivantes :

- commutativité ($xTy = yTx$) ;
- associativité ($xT(yTz) = (xTy)Tz = xTyTz$) ;
- monotonie ($xTz < yTz$ si $x < y$) ;
- admet 0 comme élément absorbant ($0Tx = 0$) et 1 comme élément neutre ($1Tx = x$).

Cet opérateur peut être réalisé de plusieurs façons. Cependant les solutions les plus simples et les plus utilisées pour caractériser la satisfaction simultanée de deux propriétés sont :

- opérateur minimum (d'après Zadeh) : $\mu_{A \cap B}(x) = \min \{ \mu_A(x), \mu_B(x) \}$; (Fig.19)
- produit algébrique (représentation probabilistique) : $\mu_{A \cap B}(x) = \mu_A(x) \times \mu_B(x)$;
- produit borné (d'après Lukasiewicz) : $\mu_{A \cap B}(x) = \max \{ \mu_A(x) + \mu_B(x) - 1, 0 \}$;
- produit d'Hamasher : $\mu_{A \cap B}(x) = \frac{\mu_A(x)\mu_B(x)}{[\beta + (1-\beta)(\mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x))]}$ où $\beta \in [0,1]$;
- produit d'Einstein : $\mu_{A \cap B}(x) = \frac{\mu_A(x)\mu_B(x)}{2 - [\mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)]}$;
- produit drastique (d'après Weber) : $\mu_{A \cap B}(x) = \begin{cases} \mu_A(x) & \text{si } \mu_B(x) = 1 \\ \mu_B(x) & \text{si } \mu_A(x) = 1 \\ 0 & \text{sinon} \end{cases}$

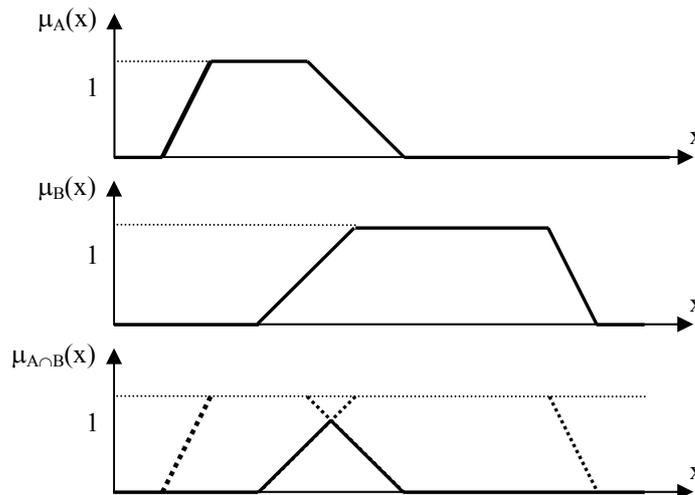


Fig.19 : l'opérateur ET réalisé par la formation du minimum.

2.2.2.4.3.3. L'union :

L'union des deux ensembles A et B est définie par une co-norme triangulaire (s-norme) :

$$\begin{aligned} \text{OU} : [0, 1] \times [0, 1] &\rightarrow [0, 1] \\ (x, y) &\rightarrow z = x \text{ OU } y \end{aligned}$$

La co-norme OU possède les propriétés suivantes :

- commutativité ($xTy = yTx$) ;
- associativité ($xT(yTz) = (xTy)Tz = xTyTz$) ;
- monotonie ($xTz < yTz$ si $x < y$) ;
- admet 1 comme élément absorbant ($1Tx = 1$) et 0 comme élément neutre ($0Tx = x$).

Tout comme pour l'opérateur ET, il existe de nombreuses possibilités pour représenter l'opérateur OU. Les plus répandus sont :

- opérateur maximum (d'après Zadeh) : $\mu_{A\cup B}(x) = \max \{ \mu_A(x), \mu_B(x) \}$; (Fig.20)
- somme algébrique : $\mu_{A\cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$;
- somme bornée (d'après Lukasiewicz) : $\mu_{A\cup B}(x) = \min \{ \mu_A(x) + \mu_B(x), 1 \}$;
- somme d'Hamasher : $\mu_{A\cup B}(x) = \frac{[\mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)] - (1-\beta)\mu_A(x)\mu_B(x)}{[1 - (1-\beta)\mu_A(x)\mu_B(x)]}$ où $\beta \in [0, 1]$;
- produit d'Einstein : $\mu_{A\cup B}(x) = \frac{\mu_A(x) + \mu_B(x)}{1 + \mu_A(x)\mu_B(x)}$;
- somme drastique (d'après Weber) : $\mu_{A\cup B}(x) = \begin{cases} \mu_A(x) & \text{si } \mu_B(x) = 0 \\ \mu_B(x) & \text{si } \mu_A(x) = 0 \\ 1 & \text{sinon} \end{cases}$
- somme disjonctive : $\max \{ \min \{ \mu_A(x), \bar{\mu}_B(x) \}, \min \{ \bar{\mu}_A(x), \mu_B(x) \} \}$.

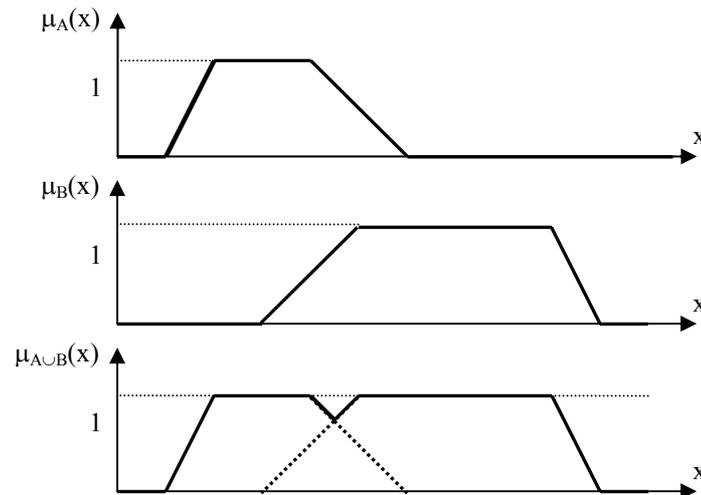


Fig.20 : l'opérateur OU réalisé par la formation du maximum.

2.3. LA LOGIQUE FLOUE ET LES PROBABILITES :

La théorie des probabilités est très différente de la théorie de la logique floue bien que toutes deux décrivent une notion de doute, d'incertain et ce à l'aide de nombres compris entre 0 et 1 [7].

La fonction d'appartenance est une notion subjective qui quantifie une variable linguistique de manière floue alors que "la probabilité est une notion objective, qui est le résultat de mesures convenablement faites ou de raisonnement sur la répétition des états" [25]. La logique floue décrit l'ambiguïté d'un événement alors que les probabilités décrivent le doute quant à l'occurrence d'un événement.

Mathématiquement, la différence essentielle est la violation du principe du tiers-exclus par la logique floue qui permet les contradictions. Cela se traduit par :

- $P(A \cap \bar{A}) = P(\emptyset) = 0$ alors que $\mu_{A \cap \bar{A}}(x)$ pas nécessairement égal $\mu_{\emptyset}(x) = 0$;
- $P(A \cup \bar{A}) = P(U) = 1$ alors que $\mu_{A \cup \bar{A}}(x)$ pas nécessairement égal $\mu_U(x) = 1$.

2.4. LES SYSTEMES BASES SUR LA LOGIQUE FLOUE :

Depuis les années 80, les systèmes basés sur la logique floue sont devenus l'un des domaines de recherche les plus fructueux en informatique. Les travaux contemporains se sont inspirés des recherches de Mamdani sur le control flou [28] et ont été motivés par les articles de Zadeh sur l'approche linguistique dans le réglage des systèmes [27] [33] [34]. Depuis les années 85, la logique floue est appliquée dans les systèmes de réglage ainsi que dans les systèmes experts, dans les systèmes de décision et pour la reconnaissance des formes. Le domaine d'utilisation le plus important est celui du réglage et de la commande des systèmes industriels où le japon détient la plus grande part : appareils photo, commandes d'ascenseur, contrôleur de réacteurs nucléaires, aspirateurs, produits électroménagers ...

2.4.1. L'univers de discours :

L'univers de discours est l'ensemble de référence d'un mot du langage naturel. Il contient des termes qui évoquent à différents degrés le même concept que la variable floue et couvre l'ensemble des valeurs prises par cette variable.

2.4.2. Les termes et les variables linguistiques :

Pour décrire les procédés, les phénomènes ou les situations auxquelles nous sommes confrontés, nous avons recours à des expressions telles que: quelque, beaucoup, souvent, petit, lent, jeune ... etc. De telles expressions portent le nom de *termes linguistiques*. Ils forment les valeurs des *variables linguistiques* de la logique floue. En général, une variable linguistique est décomposée en un ensemble de termes qui couvrent tout son domaine de variation. Elle est caractérisée par le triplet $(x, T(x), U)$ dans lequel x est le nom de la variable, $T(x)$ est l'ensemble des termes linguistiques que peut prendre la variable x et U est l'univers de discours associé. Par exemple, la variable linguistique *température* peut avoir comme valeur les termes linguistiques qui appartiennent à l'ensemble $T(x) = \{\text{très froid, froid, tiède, chaud, très chaud}\}$. L'univers de discours est $U = [0^{\circ}\text{C}, 40^{\circ}\text{C}]$. (Fig.21)

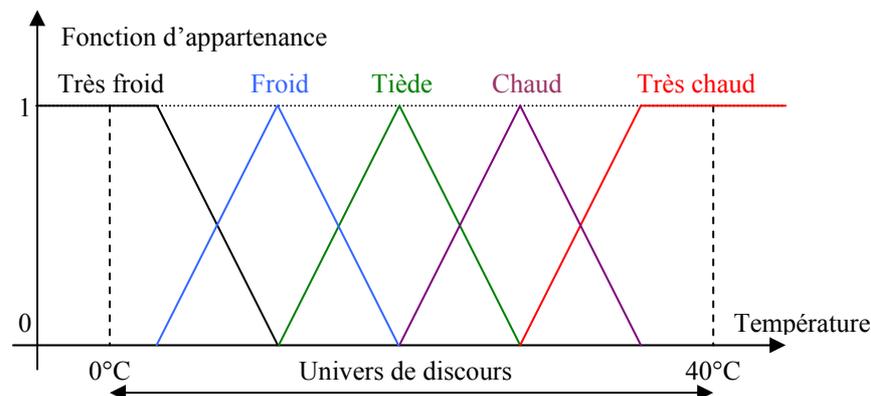


Fig.21 : Variable linguistique, termes linguistiques et univers de discours.

2.4.3. Les règles linguistiques :

La logique floue a pour objectif de formaliser et de mettre en œuvre le raisonnement d'un être humain. L'idée principale pour cela est de l'exprimer sous forme de *règles linguistiques ou floues* de type : *si* (un ensemble de conditions est satisfait) *alors* (un ensemble de conséquences est validé).

Les règles floues sont basées sur l'implication floue. Elles relient les variables floues d'entrée d'un système aux variables floues de sortie de ce système. La partie antécédente exprimée par "si" (prédicat ou prémisse) décrit l'état du système. C'est une combinaison de propositions par des opérateurs ET, OU, NON.

La partie conséquente exprimée par "alors" (conclusion) correspond à l'action que l'opérateur doit effectuer. C'est une combinaison de propositions liées par des opérateurs ET uniquement. L'opérateur OU n'est pas admis dans les conclusions car il signifierait que l'expertise ne permet pas de déterminer quelle décision prendre [32]. La négation n'est également pas permise car une conclusion du type "alors accélération non moyenne" est insuffisante. Elle ne tranche pas entre les décisions "accélération forte" et "accélération nulle".

Les règles linguistiques peuvent être exprimées de plusieurs façons. Celle qui a été proposée en premier est la forme [34] : si (x est A) alors (y est B).

Une autre forme ne fait apparaître les sous-ensembles flous que dans les prédicats. La sortie est une fonction analytique de l'entrée [35] : si (x est A) alors (y est f(x)).

2.4.4. L'implication floue :

L'implication floue, notée " \rightarrow " est un opérateur qui permet d'évaluer le degré de vérité d'une règle floue. Les normes d'implication les plus utilisées sont :

- l'implication matérielle : $\mu_{A \rightarrow B}(x) = \text{NON}(\mu_A(x)) \text{ OU } \mu_B(x)$;
- le calcul propositionnel : $\mu_{A \rightarrow B}(x) = \text{NON}(\mu_A(x)) \text{ OU } (\mu_A(x) \text{ ET } \mu_B(x))$.

En logique floue, le raisonnement est basé sur les principes du *modus tollens généralisé* et *modus ponens généralisé* :

- a) le *modus tollens généralisé* : prémisses : x est A'
implication : si (x est A) alors (y est B)
conséquence : y est B'
- b) le *modus ponens généralisé* : prémisses y est B'
implication : si (x est A) alors (y est B)
conclusion : x est A'

Exemple :

Prémisse : cette tomate est très rouge.

Implication : si une tomate est rouge alors elle est mûre.

Conclusion : cette tomate est très mûre.

2.4.5. La structure générale d'un système flou :

La mise en oeuvre d'un système basé sur la logique floue fait apparaître les quatre modules principaux suivants (Fig.22) :

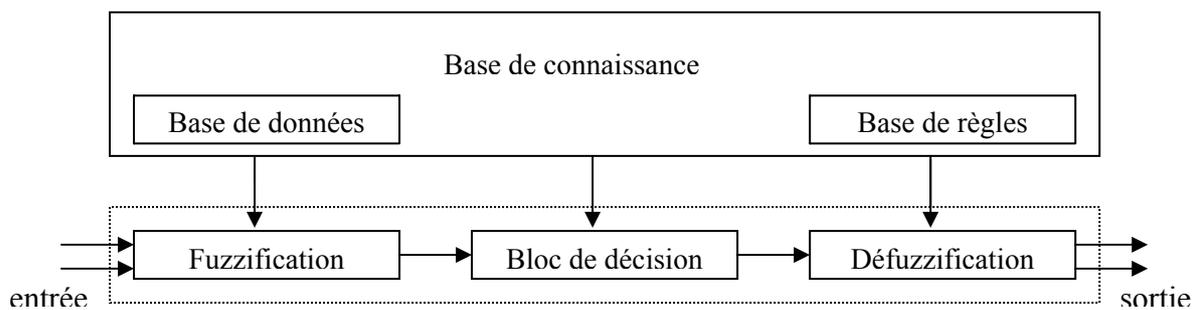


Fig.22 : Structure générale d'un système basé sur la logique floue.

2.4.5.1. L'interface de fuzzification : Elle permet de passer du "monde réel" au "monde flou" en transformant les entrées exactes du système en degrés d'appartenance.

2.4.5.2. La base de connaissance : Elle est scindée en deux parties : la base des règles floues qui contient les relations entre les variables d'entrée (variables linguistiques) et les variables de sortie (variables linguistiques aussi) et la base de données qui contient la définition des ensembles flous, les facteurs d'échelle pour la normalisation des ensembles de référence et la partition de l'espace flou d'entrée et de sortie.

2.4.5.3. Le bloc de décision ou le moteur d'inférence : Il contient les inférences des opérations sur les règles.

2.4.5.4. L'interface de défuzzification : C'est l'opération inverse de la fuzzification. Elle fait correspondre à chaque résultat flou calculé, la sortie précise adéquate.

2.4.6. La procédure de raisonnement flou :

2.4.6.1. La fuzzification :

La fuzzification consiste à attribuer à la valeur réelle de chaque entrée, au temps t , son degré d'appartenance à chacune des classes préalablement définies. L'objectif de cette étape est de définir les fonctions d'appartenance pour les différentes variables. On réalise ainsi un passage des grandeurs physiques aux variables linguistiques à traiter par la suite.

Le choix du nombre d'ensembles flous, de la forme des fonctions d'appartenance, du recouvrement de ces fonctions et de leur répartition sur l'univers de discours est primordiale et un ensemble de règles doit être suivi.

Une subdivision très fine de l'univers de discours sur plus de sept ensembles flous n'apporte en général aucune amélioration du comportement dynamique du système à réguler [32]. Par contre, la manière dont les fonctions d'appartenance des ensembles flous sont disposées sur l'univers de discours influe énormément sur ce dernier [32].

Les fonctions triangulaires et trapézoïdales sont les plus répandues car leur implémentation est facile et l'étape de fuzzification ne requiert alors que peu de temps de calcul lors de son évolution en temps réel. Elles sont généralement choisies symétriques par rapport à $x = 0$ et distribuées de manière équidistante (les maxima des fonctions d'appartenance des différents ensembles sont séparés par la même distance) (Fig.23)

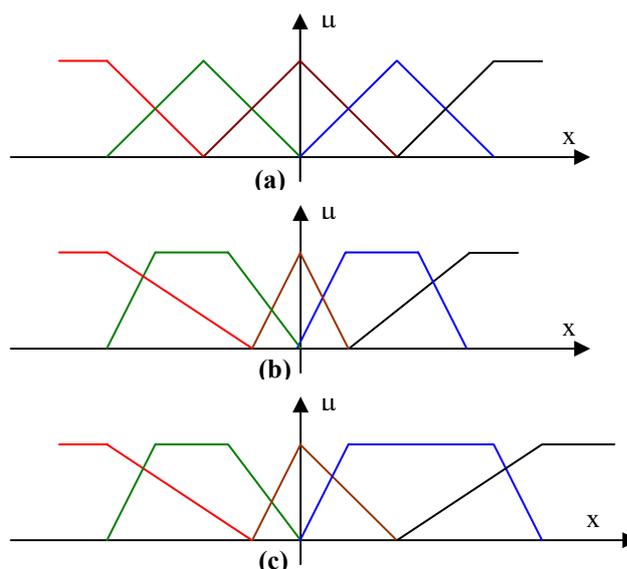


Fig.23 : a) Fonctions d'appartenance symétrique et équidistante.
 b) Fonctions d'appartenance symétrique et non équidistante.
 c) Fonctions d'appartenance non symétrique et non équidistante.

Le passage d'une fonction d'appartenance à sa voisine doit s'effectuer en douceur. Des chevauchements insuffisants ou trop importants risquent de mener à des résultats insensés. Il est donc préférable que le recouvrement de deux fonctions d'appartenance voisines soit de 1, c'est-à-dire que leur croisement s'effectue à $\mu = 0.5$ (Fig.24).

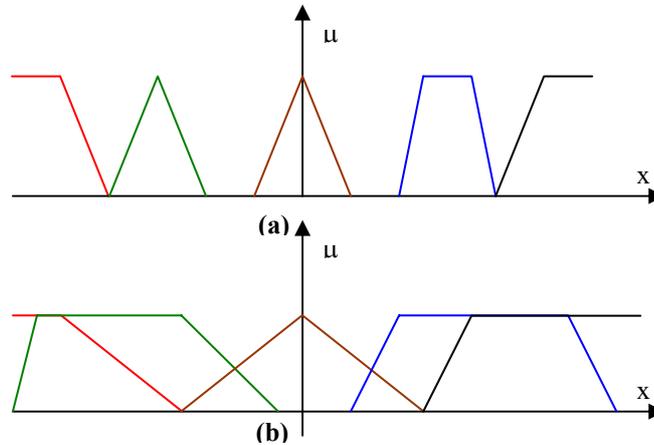


Fig.24 : a) Fonctions d'appartenance avec lacunes ou chevauchement insuffisant.
 b) Fonctions d'appartenance avec chevauchement trop important.

2.4.6.2. Les inférences floues :

Les inférences floues relient les variables d'entrée à une variable de sortie par des règles linguistiques combinées en utilisant les opérateurs logiques ET et OU. L'opérateur ET est utilisé pour les variables à l'intérieur d'une même règle, alors que l'opérateur OU relie les différentes règles entre elles.

Les règles d'inférence peuvent être décrites de plusieurs façons :

- Linguistiquement : en les explicitant : si la formation est bonne et l'expérience est longue alors la qualification est très bonne.
- Symboliquement : en utilisant des abréviations pour désigner les ensembles flous et les variables flous : si F est B et E est L alors Q est TB.
- Par une matrice d'inférence : qui rassemble toutes les règles d'inférence sous forme de tableau où les colonnes correspondent à l'une des entrées et les lignes à l'autre. L'intersection d'une colonne et d'une ligne donne l'ensemble flou de la variable de sortie défini par la règle (Tab.1).

Qualification		Formation		
		Mauvaise	Moyenne	Bonne
Expérience	Courte	Aucune	Mauvaise	Moyenne
	Correcte	Mauvaise	Moyenne	Bonne
	Longue	Moyenne	Bonne	Très bonne

Tab.1. Des règles d'inférence décrites par une matrice d'inférence.

Il n'est pas nécessaire que toutes les cases du tableau soient remplies, on parle dans ce cas de *règles d'inférence incomplètes*. Cela ne signifie pas que la sortie n'existe pas, mais plutôt que le degré d'appartenance est nul pour la règle en question.

2.4.6.3. L'agrégation des règles linguistiques :

Lorsque la base de connaissance comporte plusieurs règles, l'ensemble flou inféré est obtenu après une opération appelée agrégation ou activation des règles. Il s'agit de pondérer la contribution de toutes les règles activées pour en déduire la sortie. Généralement, les règles sont en parallèle, liées par l'opérateur OU. C'est comme si chaque règle floue donnait un avis sur la valeur à attribuer à la sortie et la combinaison des poids de chaque avis détermine le degré de vérité de la conclusion.

2.4.6.4. Les types d'inférences floues :

Il existe plusieurs méthodes pour définir les degrés d'appartenance de la variable de sortie à ses sous-ensembles flous. Elles se différencient essentiellement par la manière dont vont être réalisés les opérateurs Et et OU utilisés dans les règles d'inférence.

Les trois méthodes d'inférence les plus usuelles sont : Max-min, Max-produit et Somme-produit.

2.4.6.4.1. La méthode d'inférence max-min :

L'implication floue et l'opérateur ET sont réalisés par la fonction minimum et l'opérateur logique OU est réalisé par la fonction maximum. Elle présente l'inconvénient de nécessiter un temps de calcul trop élevé qui la rend inutilisable pour les applications en temps réel.

Cette méthode est également dite "implication de Mamdani". La fonction d'appartenance résultante correspond au maximum des deux fonctions d'appartenance partielles puisque les règles sont liées par l'opérateur OU.

2.4.6.4.2. La méthode d'inférence max-prod :

La seule différence avec la méthode précédente est la réalisation de l'implication floue, le produit est utilisé. Les opérateurs logiques ET et OU sont toujours réalisés respectivement par les fonctions minimum et maximum. La méthode max-prod est également dite implication de Larsen [32].

2.4.6.4.3. La méthode d'inférence somme-prod :

La méthode d'inférence somme-prod réalise, au niveau de la condition, l'opérateur OU par la formation somme, et l'opérateur ET par la formation produit. Au niveau de la conclusion, elle réalise également l'opérateur ALORS par la formation produit.

2.4.6.5. La défuzzification :

A la fin de l'inférence, le résultat est un sous-ensemble flou qui n'est pas directement utilisable pour donner une information précise à l'opérateur ou commander un actionneur. La défuzzification permet de transformer ces fonctions d'appartenance en informations déterministes et donc de définir des variables de sortie numériques précises.

Plusieurs méthodes de fuzzification sont proposées dans la littérature et aucune ne peut être jugée meilleure que les autres. Le choix de la méthode à utiliser dépend essentiellement de l'expérience du développeur et du cahier des charges. La priorité peut donc être donnée à la réduction du temps de calcul ou à la diminution de la complexité du système ou encore à

l'universalité de l'application (applicabilité à toutes les situations considérées) ou bien alors à la plausibilité des résultats du point de vue expert [7].

2.4.6.5.1. La méthode du centre de gravité :

La sortie défuzzifiée x^* correspond à l'abscisse du centre de gravité de la surface de la fonction d'appartenance résultante $\mu(x)$ (Fig.25). Mathématiquement, cela se traduit par :

$$x^* = \frac{\int_U \mu(x) x dx}{\int_U \mu(x) dx}$$

L'intégrale au dénominateur correspond à la surface tandis que celle au numérateur donne le moment d'inertie de cette surface.

Si la sortie du système est discrète (c'est souvent le cas), l'intégrale devient une sommation :

$$x^* = \frac{\sum_{i=1}^n \mu(x_i) x_i}{\sum_{i=1}^n \mu(x_i)}$$

La défuzzification par Centre de Gravité est la méthode la plus répandue. Ceci dit, le calcul de l'intégrale (ou de la somme) nécessite beaucoup de calculs. Le temps d'exécution de cette méthode est donc assez important et ceci la rend inadéquate pour un traitement en temps réel.

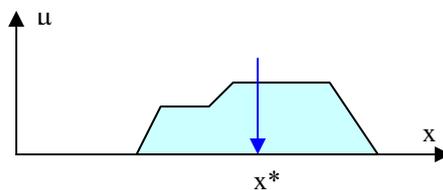


Fig.25 : Défuzzification par centre de gravité.

2.4.6.5.2. La méthode de la moyenne des maxima :

Elle a été introduite afin de remédier à la lourdeur des calculs de la méthode du centre de gravité. La sortie défuzzifiée est obtenue par le calcul de la moyenne des abscisses pour lesquelles la fonction d'appartenance est maximale :

$$x^* = \frac{\sum_{i=1}^n u(x_i)}{n} \quad \text{où les } x_i \text{ sont définis par : } \mu(x_i) = \max(\mu(x)).$$

Chapitre III

Les réseaux hybrides

3.1. LA COMBINAISON DES RESEAUX NEURONAUX ET DE LA LOGIQUE FLOUE :

Les réseaux de neurones et la logique floue sont deux concepts totalement inspirés du raisonnement humain. Bien que ces deux modèles numériques soient différents du point de vue structure, ils possèdent de nombreux points communs. En effet, l'utilisation de ces deux techniques ne nécessite pas de modèle mathématique bien défini [7] car la résolution se base sur les valeurs numériques des entrées et des sorties pour les réseaux neuronaux et sur la logique du système dans le cas de la logique floue. De plus, les résultats sont généralement incertains. En revanche ces deux techniques peuvent accomplir des tâches complémentaires. La logique floue fournit des connaissances avec un certain degré d'incertitude (ou d'exactitude) et les réseaux de neurones peuvent modéliser et reproduire l'apprentissage humain, d'où l'idée de combiner ces deux techniques pour créer un système artificiel intelligent qui résout les problèmes de manière proche de celle de l'être humain.

Les méthodes des réseaux de neurones peuvent être fusionnées à l'intérieur des systèmes flous. Ceci permet de les adapter à de nouvelles situations et de nouvelles règles, lors du fonctionnement en temps réel. De même, les techniques de la logique floue peuvent être introduites dans les réseaux neuronaux afin de perfectionner et d'améliorer l'expressivité et la flexibilité du réseau. Ces combinaisons sont appelées *les systèmes hybrides : neuro-flous (fuzzy-neuro systems)* [7]. De telles combinaisons permettent non seulement de s'approcher davantage de la structure de la pensée humaine mais aussi d'obtenir un champ d'utilisation plus large et d'utiliser les avantages de chaque technique, à savoir la simplicité des connaissances floues d'une part et l'habileté d'apprentissage des réseaux de neurones de l'autre. L'utilisation des réseaux hybrides permet, en plus, de diminuer le temps d'apprentissage par la possibilité d'introduire de la connaissance a priori dans le système d'aide au diagnostic. Le choix de la façon d'intégration de ces deux techniques est crucial. Différentes architectures sont ainsi proposées.

3.2. LES DIFFERENTES STRUCTURES DES SYSTEMES HYBRIDES :

La manière de combiner les réseaux neuronaux avec la logique floue dépend essentiellement du domaine d'application. Globalement, deux grandes catégories prédominent [36] : les réseaux de neurones flous où les opérateurs de la logique floue sont introduits dans le neurone lui-même et les systèmes neuro-flous où soit des réseaux neuronaux interviennent localement dans un système flou et inversement (systèmes séparés), soit que l'architecture des réseaux de neurones est inspirée du raisonnement flou (systèmes fusionnés).

3.2.1 Les réseaux de neurones flous :

Ces réseaux sont des réseaux neuronaux dont l'élément de base est un opérateur flou qui est souvent appelé *le neurone flou (fuzzy-neuron)* [7].

3.2.1.1. Le neurone flou :

Un neurone flou (NF) est un système mathématique flou qui possède la même architecture que son homologue le neurone formel. La différence est que quelques-uns ou tous ses paramètres sont décrits par la logique floue. Selon la fonction d'entrée (h) du neurone ou sa fonction de sortie (g), plusieurs variétés existent. Chacune possède ses propres particularités. Suivant la nature des entrées et leurs poids, les neurones flous peuvent être classés en [37] :

- entrées non floues (booléennes) et poids flous ;
- entrées floues et poids non flous ;
- entrées floues et poids flous.

L'étape la plus importante dans la modélisation d'un neurone flou est la fuzzification des synapses (dendrites) d'entrée et la définition des fonctions d'appartenance de chacun des sous-ensembles de l'espace de travail. Dans une fuzzification neuronale 'Neuronal fuzzification', chaque neurone peut être vu comme la représentation d'une valeur linguistique. Sa sortie peut être associée à une fonction d'appartenance. En d'autres termes, la sortie exprime le degré d'appartenance de l'entrée à une certaine catégorie linguistique. La sortie du neurone possède une valeur comprise entre 0 et 1.

Le comportement du neurone flou est déterminé par les opérateurs d'agrégation choisis, et par la fonction d'entrée utilisée.

3.2.1.1.1. Le neurone flou de type 'max.' (Max fuzzy neuron) :

Le neurone flou est de type 'max' lorsque sa fonction d'agrégation floue d'entrée (h) 'aggregation function' prend le maximum des valeurs d'entrées pondérées. C'est une implémentation de la fonction logique OU. Sa fonction d'entrée est donnée par la relation : $h(x_1 \dots x_n) = \max (x_i w_i)$ où x_i est une entrée et w_i son poids [36].

Ce neurone est également appelé *le neurone flou max-min* lorsque, la fonction minimum remplace la multiplication de l'entrée et de son poids. Son entrée totale 'E' est donnée par : $E = \max (\min (w_i , x_i))$ pour $i = 1 \dots n$ (n étant le nombre d'entrées) [37].

3.2.1.1.2. Le neurone flou de type 'min' (Min fuzzy neuron) :

Le neurone flou est de type 'min' lorsque sa fonction d'entrée (h) prend la valeur minimale des entrées pondérées. C'est une implémentation de la fonction logique ET. Elle est décrite par l'équation : $h(x_1 \dots x_n) = \min (x_i w_i)$ [36].

Ce neurone est également appelé *le neurone flou min-max* lorsque, la fonction maximum remplace la multiplication de l'entrée par son poids. Son entrée totale 'E' est donnée par : $E = \max (\min (w_i , x_i))$ pour $i = 1 \dots n$ (n étant le nombre d'entrées) [37].

3.2.1.1.3. Le neurone flou de type produit-somme (product-sum fuzzy neuron) :

La fonction d'entrée (h) de ce neurone flou est la somme algébrique des produits arithmétiques des entrées et de leurs poids respectifs. La valeur de l'entrée totale du neurone est alors donnée par la relation de récurrence suivante : [37]

$$\begin{cases} v_0 = x_0 w_0 \\ v_i = x_i w_i + v_{i-1} - x_i w_i v_{i-1} \quad \text{pour } i = 1 \dots n \\ E = v_n \end{cases}$$

3.2.1.1.4. Le neurone flou de Kwan et Cai :

Les entrées sont introduites avec leur poids. Le neurone possède plusieurs sorties (y_i) dont les valeurs varient entre 0 et 1. Elle représente le degré d'appartenance du vecteur d'entrée à chaque sous-ensemble flou. Les valeurs d'entrées sont une représentation des variables linguistiques alors que les sorties expriment le degré d'appartenance à des descriptions telles que grand, petit, moyen. Les sorties sont ensuite transférées vers d'autres neurones (Fig.26).

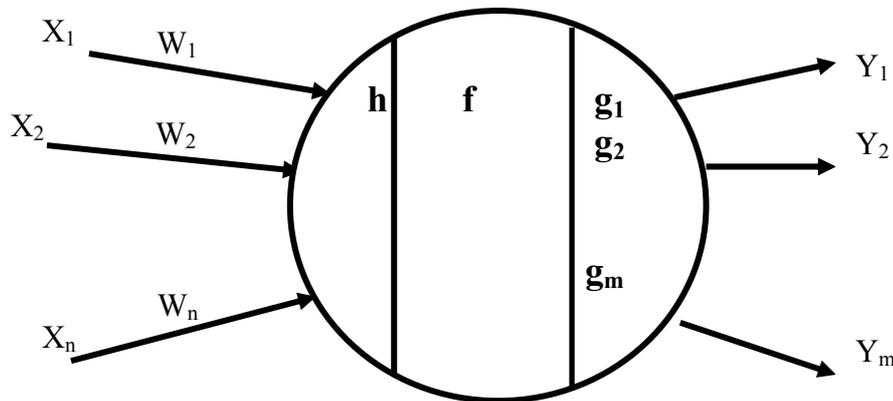


Fig.26 : Architecture du neurone flou de Kwan et Cai.

L'expression mathématique du neurone flou est donnée par la relation [7]:

$$y_j = g_j [f (h_{i=1}^n (w_i - x_i) - \theta)]$$

où : h est un opérateur flou. Il remplace la somme dans un neurone formel ;
 f est une fonction d'activation ;
 θ est le seuil d'activation ;
 g_j ($j = 1 \dots m$) sont les m fonctions de sorties.

3.2.1.1.5. Le neurone flou compétitif :

Dans ce neurone le seuil d'activation θ de la fonction de sortie (g) est variable. La sortie de ce neurone est unique, elle est donnée par la relation [7] :

$$y = g (A - \theta) = \begin{cases} 0 & \text{si } s < \theta \\ 1 & \text{si } s > \theta \end{cases}$$

où $\theta = f (c_1, c_2 \dots c_k)$, t est la fonction de seuillage et les c_i sont des variables compétitives du neurone flou ;
 g est la fonction de sortie du neurone ;
 $A = f (h (x_i w_{ij}))$, A est la valeur de sortie de la fonction d'activation f et h est une fonction d'agrégation floue d'entrée.

3.2.1.1.6. Le neurone flou de Eiji Uchinot et Takeshi Yamakawat :

Ce modèle a été proposé en 1992. Les poids des entrées sont déterminés à partir des fonctions d'inférences floues (Fig.27 et 28) [38].

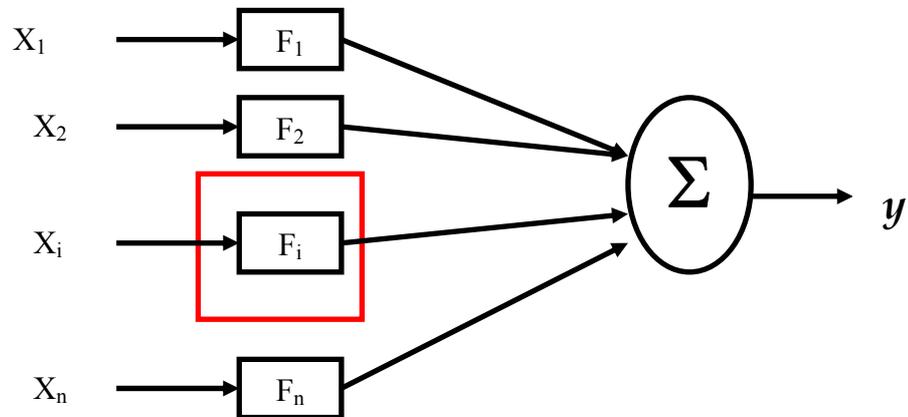


Fig.27 : Architecture du neurone flou de E.Uchinot et T.Yamakawat.

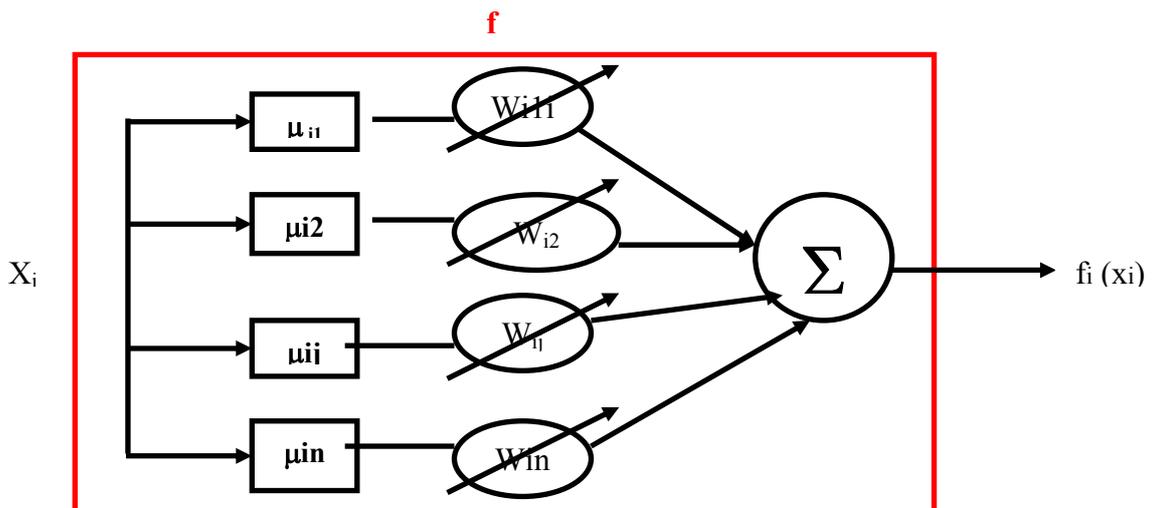


Fig.28 : Fonction d'activation d'une entrée synaptique.

La sortie du neurone est obtenue par la somme des différentes fonctions de synapses :

$$Y = \sum_{i=1}^n f_i(x_i)$$

L'adaptation des poids subit elle aussi l'inférence floue : si x_i est A_{ij} alors le poids est W_{ij} où A_{ij} est le sous-ensemble flou dont la fonction d'appartenance est $\mu_{ij}(x_i)$ et W_{ij} est le poids attribué à l'entrée.

3.2.1.2. La connexion des neurones flous :

Les neurones flous sont organisés en réseaux qui accomplissent des tâches. Plusieurs architectures ont été proposées dans la littérature. Parmi ces architectures, il y a :

3.2.1.2.1. La structure de Pedrycz :

C'est un réseau de neurones flous de trois couches. Le type de neurones flous change d'une couche à l'autre. Deux possibilités sont envisageables. Soit la couche cachée est constituée de neurones de type 'max' et la couche de sortie est formée à partir d'un seul neurone flou de type 'min' (Fig.29) ou bien la couche cachée contient des neurones flous de type 'min' et le neurone flou de la couche de sortie est de type 'max'.(Fig.30)

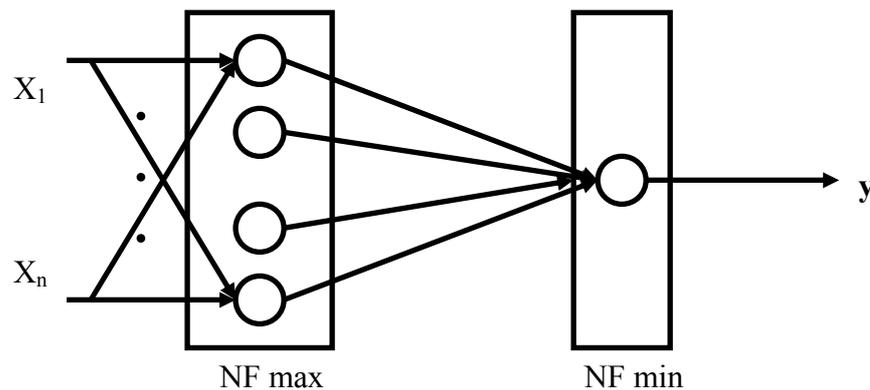


Fig.29 : Réseaux de Pedrycz de type max-min.

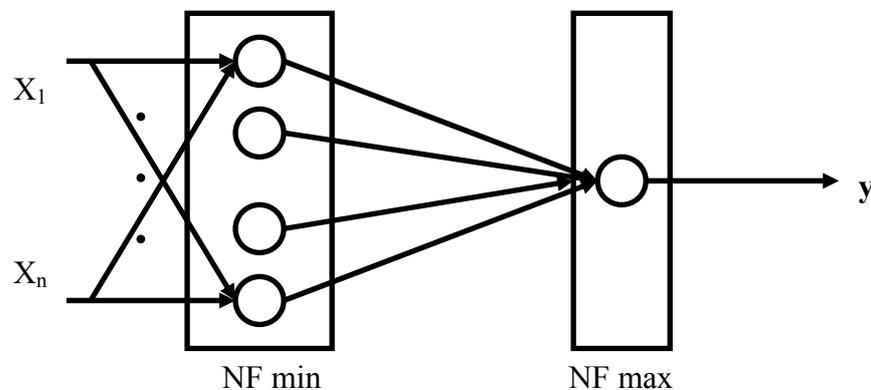


Fig. 30 : Réseaux de Pedrycz de type min-max.

3.2.1.2.2. L'apprentissage et l'adaptation des réseaux de Pedrycz :

L'apprentissage des neurones flous implique le changement des relations floues entre les synapses d'entrées et la sortie en agissant sur le poids des entrées. L'adaptation suit en général, la méthode de descente de gradient (Fig.31).

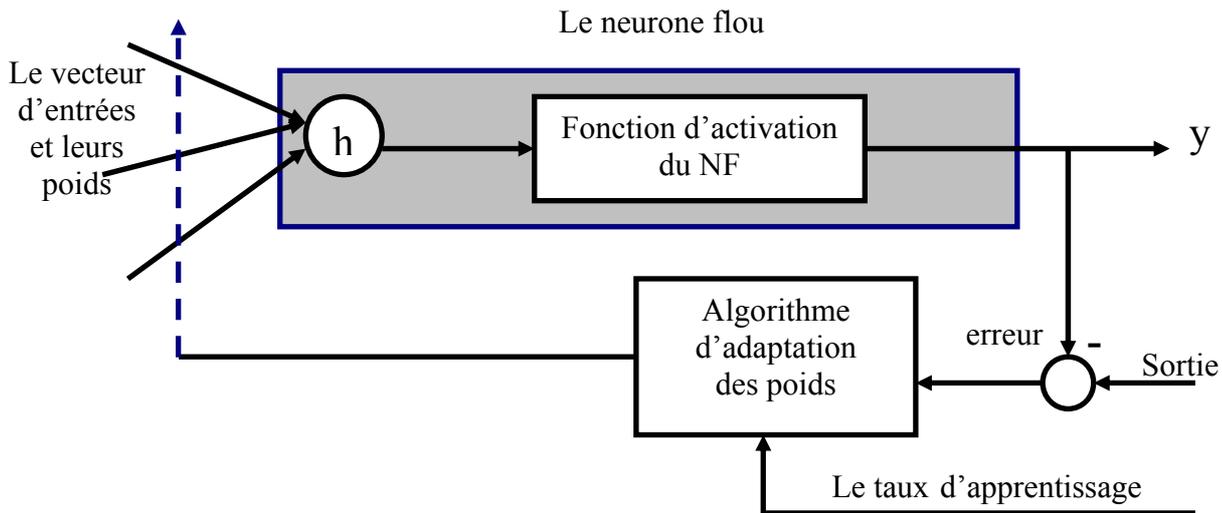


Fig.31 : Schéma d'apprentissage et d'adaptation d'un réseau Pedrycz.

Théoriquement, les neurones flous peuvent être introduits dans n'importe quelle structure de réseaux neuronaux (les réseaux multicouches à apprentissage supervisé, les réseaux bayésiens, les réseaux compétitifs à apprentissage non supervisé, la carte auto-organisatrice de Kohonen ...). D'où les structures de Kwan-Cai, les réseaux ART flous (Fuzzy ART), et les cartes floues (Fuzzy ARTMap) [7].

3.2.2. Les systèmes neuro-flous :

L'apport des méthodes neuronales dans des systèmes flous, a pour but d'extraire ou d'identifier les règles floues et les fonctions d'appartenance. Cette introduction permet en plus d'adapter les systèmes flous aux changements physiques des systèmes et de leur environnement. La stratégie d'appui neuronal consiste à identifier certains paramètres des systèmes flous et à utiliser les réseaux neuronaux pour les générer et les ajuster.

L'introduction de la logique floue dans les réseaux neuronaux permet de résoudre les problèmes liés à la taille de la base de données, à l'initialisation des poids et à leur adaptation.

Deux grandes approches de la combinaison neuro-floue sont proposées : les systèmes séparés et les systèmes fusionnés.

3.2.2.1. Les systèmes séparés :

3.2.2.1.1. Les modèles coopératifs (cooperative models) :

Dans les systèmes coopératifs, l'un des deux systèmes (flou ou neuronal) intervient une seule fois pour effectuer une tâche spécifique. Les résultats obtenus sont injectés et utilisés par le second système.

Réseaux de neurones pour la détermination des fonctions d'appartenance :

La détermination des fonctions d'appartenance des systèmes flous est d'une importance primordiale pour leur conception. La démarche de la fuzzification est souvent subjective et dépend en grande partie de l'expérience requise. Donc cette fonction peut être vue comme étant une récolte de données et de classification de problèmes. De là, les RNA peuvent être utilisés pour générer les fonctions d'appartenance des sous-ensembles flous. Une fois

déterminés, les résultats sont transférés au système flou pour une éventuelle résolution du problème (Fig.32).

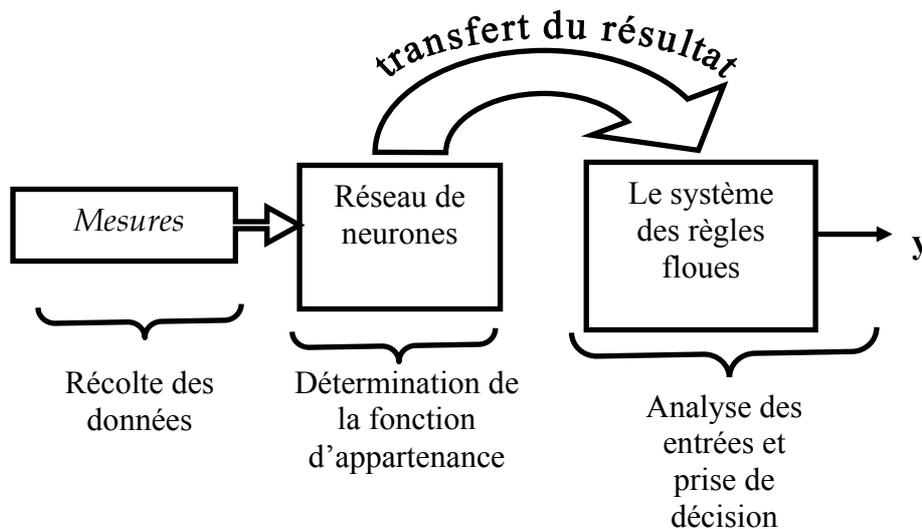


Fig.32 : Schéma des modèles coopératifs.

L'algorithme proposé par ADELI et HUNG et qui porte leur nom reste le plus connu chez l'ensemble des chercheurs du domaine [7].

Cet algorithme comporte deux étapes. La première est celle de la classification. Un réseau de neurone est utilisé pour classer les données en différentes catégories. Les réseaux utilisés sont les réseaux non supervisés (généralement une carte auto-organisatrice de KOHONEN de topologie plane). Le nombre de neurones d'entrée est égal au nombre d'exemples d'entrée dans chaque phase d'apprentissage alors que le nombre de neurones de sortie est égal au nombre des catégories déterminées.

La deuxième étape est l'étape de la fuzzification. Les fonctions d'appartenance sont évaluées pour chaque exemple d'entraînement à partir des valeurs d'entrées classées sous différentes catégories.

Le principe de l'algorithme :

Soit à identifier un système de 'M' entrées et de 'N' sous-ensembles de sortie. L'algorithme commence par un réseau (noté $RN(M,1)$) de 'M' entrées et d'une seule sortie. A la fin de l'apprentissage les 'N' sorties doivent être déterminées. C'est-à-dire le réseau s'écrit $RN(M,N)$. Le premier exemple d'apprentissage appartient à la première catégorie. Ensuite, le second exemple d'apprentissage est soumis au test suivant :

- si la sortie est classée dans la première catégorie, alors le neurone qui représente cette catégorie est actif ;
- sinon, cette sortie est classée comme étant une nouvelle catégorie et un nouveau neurone sera ajouté à la couche de sortie.

Cette procédure est reproduite pour tous les exemples de la base de données.

3.2.2.1.2. Les modèles compétitifs (competitive models) :

Dans les systèmes compétitifs, l'un des deux systèmes (flou ou neuronal) intervient continuellement pour effectuer une tâche spécifique. Les résultats obtenus sont injectés et utilisés par le second système.

Réseaux de neurones pour la détermination des règles floues :

Il est extrêmement difficile de formuler les règles floues "si alors" ou de les extraire des systèmes physiques lorsque les variables d'entrée sont soit supérieures à trois ou bien indirectement mesurées. La difficulté est aussi rencontrée lorsque leur environnement change de manière permanente. Les modèles compétitifs présentent une bonne solution à ce genre de problèmes. Dans ce type de structures, les RNA assistent et aident continuellement les systèmes flous pour déterminer les fonctions d'appartenance et les règles floues (Fig.33).

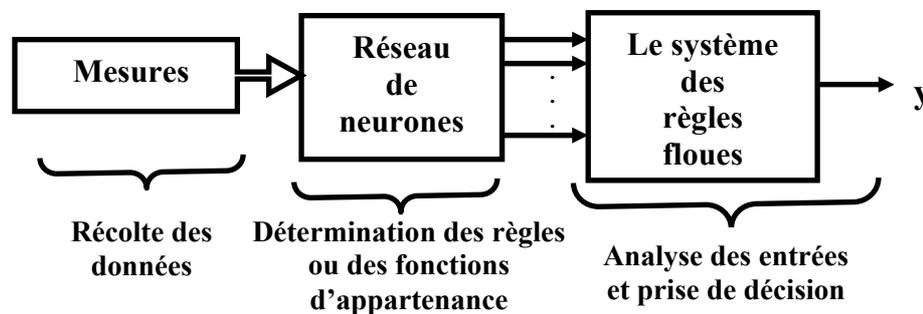


Fig.33 : Schéma des modèles compétitifs.

L'une des méthodes proposées pour introduire et ajuster ces règles est la méthode de Hayashi, Nomura et Wakami [39]. Elle a été proposée lors de la conférence internationale IFSA'91 qui s'est déroulée à Bruxelles. Cette méthode utilise exclusivement une fonction d'appartenance triangulaire et se base sur l'apprentissage supervisé en utilisant la méthode de la descente du gradient.

3.2.2.2. Les structures fusionnées (fused systems) :

Dans ces modèles, les opérations concernant le raisonnement flou sont réalisées à partir des architectures neuronales adaptées. Ces architectures sont les systèmes neuro-flous les plus utilisés par l'ensemble des chercheurs. Un système neuro-flou fusionné est un réseau de neurones topographiquement (ou topologiquement) équivalent à un système flou.

Il existe un grand nombre de structures. Chacune possède des caractéristiques bien déterminées et est orientée vers un domaine d'application spécifique. Par exemple, l'architecture Falcon, proposée en 1991 par C.T Lin et C.S.G Lee, est utilisée pour le contrôle, l'architecture ANFIS, proposée en 1993 par J.S.R Jang, est spécifique à l'approximation des fonctions et l'architecture NEFClass est dédiée à la classification [40].

Les systèmes neuro-flous de classifications (NEuro Fuzzy CLASSification) NEFClass :

Les systèmes NEFClass sont souvent appelés les réseaux perceptrons flous de trois couches (three layers fuzzy perceptron networks). Les neurones de la couche d'entrées représentent les caractéristiques des entrées, ceux de la couche cachée représentent les règles floues alors que les neurones de la couche de sortie donnent les différentes classes. Dans un système NEFClass la connexion entre la couche d'entrée et la couche cachée se fait par le biais des poids flous alors que la couche cachée et la couche de sortie sont reliées par des poids booléens (Fig.34) [41].

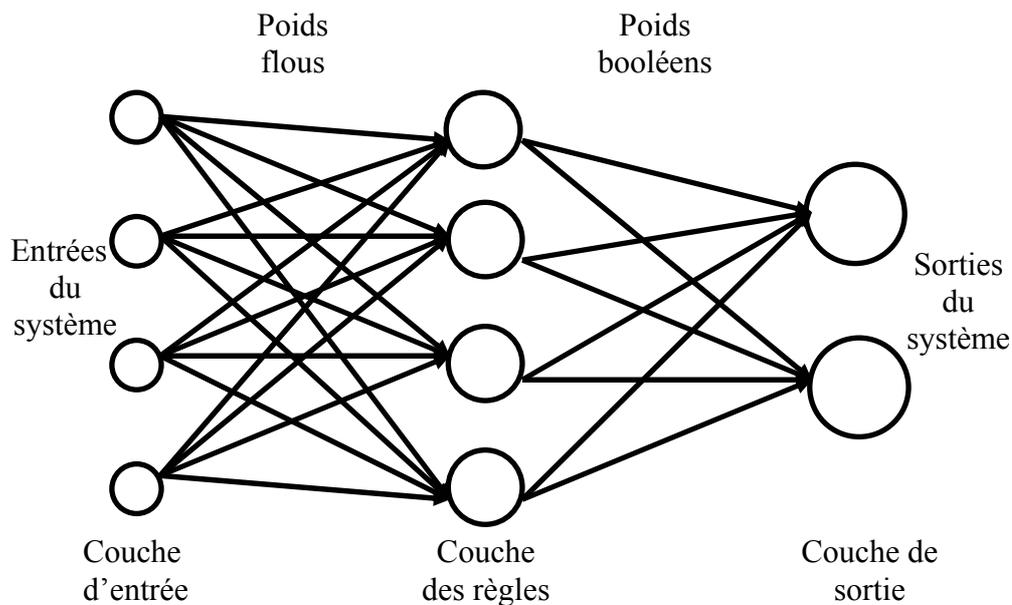


Fig.34 : Architecture d'un système NEFClass.

Ce modèle a surtout été utilisé par des chercheurs allemands pour le diagnostic et le suivi des différentes structures industrielles comme les turbines des centrales électriques [42], [43], [44] et [45].

Chapitre IV :

Diagnostic des huiles de transformateurs

4.1. GENERALITES :

Les isolants diélectriques sont indispensables dans tous les équipements électriques. Leur performance doit s'accroître en fonction de la contrainte électrique à laquelle ils sont soumis. Leur rigidité diélectrique doit être d'autant plus élevée que la tension appliquée l'est. Les diélectriques utilisés en haute tension appartiennent aux trois grandes catégories suivantes :

- les isolants gazeux : l'air, certains gaz électronégatifs (SF_6 , N_2 ... etc) ;
- les isolants liquides : la grande gamme des huiles diélectriques : minérales ... etc ;
- les isolants solides : les matériaux synthétiques (PVC), le papier ... etc.

La perte de la rigidité d'un isolant électrique se traduit par son claquage. Pour l'éviter, il faut assurer un bon suivi de la performance de ces isolants. Les normes indiquent les procédures à suivre ainsi que les caractéristiques tolérées pour chaque type d'isolant.

4.2. LES CARACTERISTIQUES DES HUILES MINERALES ISOLANTES :

Les isolants liquides englobent une gamme assez large d'huiles électriques ainsi qu'une petite partie de gaz liquéfié.

Du point de vue historique, les premières bobines d'induction ou les transformateurs primitifs (1884) étaient isolés à sec. Il a fallu attendre 1891 pour que l'huile de pétrole soit utilisée à titre expérimental dans l'isolation de transformateurs triphasés. Elle a réellement été introduite à partir de l'année 1905, à cause de l'accroissement du réseau électrique et de l'augmentation de la puissance installée. Elle a également été adoptée, en 1917, pour l'isolation du premier câble à conducteurs creux et en 1932, dans l'isolation des condensateurs imprégnés d'huile.

Les huiles diélectriques sont classées selon leurs origines en huiles minérales, huiles végétales et huiles de synthèse. Les huiles minérales sont des produits "naturels" obtenus directement par raffinage de bruts pétroliers. Elles sont les plus répandues en industrie d'isolation électrique en raison de leur moindre coût de revient. Elles sont utilisées comme imprégnant d'isolation solide ou comme produits de remplissage de matériels électriques très divers : transformateurs, condensateurs, câbles, disjoncteurs ... etc. Elles sont généralement employées avec des isolants solides et qu'en cas de développement de produits nouveaux, la compatibilité liquide-solide doit être vérifiée par de nombreux essais.

Le rôle principal des huiles isolantes utilisées dans les transformateurs est d'éliminer l'air et les autres gaz qui peuvent être emprisonnés à l'intérieur des cavités dues aux défauts de construction. Leur introduction améliore de façon considérable la tenue de l'isolant solide utilisé. De plus, l'huile sert à évacuer la chaleur (agent caloporteur). Elle est également utilisée pour son pouvoir extincteur de l'arc électrique. Ces performances peuvent être appréciées d'après les propriétés physico-chimiques de l'huile. Ces caractéristiques permettent la détermination de l'état de l'huile. La norme CEI 60296 spécifie les propriétés requises pour les huiles minérales utilisées dans les transformateurs et la norme CEI 60465 celle des huiles pour câbles [46].

L'étude et la prédiction du comportement des huiles diélectriques dans les équipements, fait appel à plusieurs sciences : la chimie, l'électricité, la mécanique des fluides ... etc. La

connaissance globale de ces sciences est nécessaire afin de mieux comprendre les phénomènes qui s'y passent.

Le choix d'une huile isolante repose sur un nombre important de propriétés telles que :

4.2.1. Les propriétés physiques :

4.2.1.1. L'aspect et l'indice de couleur :

L'aspect est un test visuel de l'huile qui permet de détecter la présence des corps en suspension (poussière, eau) et d'évaluer la couleur et la viscosité. Un bon état visuel de l'huile, signifie un aspect limpide.

La couleur est une propriété intrinsèque de l'huile neuve, elle a une relation avec les hydrocarbures qui constituent l'huile. Elle permet d'apprécier la qualité des huiles neuves et constitue un moyen efficace pour surveiller l'acidité des huiles en service. Elle renseigne également sur le vieillissement de l'huile puisque celle-ci devient plus foncée avec l'âge.

Une huile neuve possède un indice de couleur inférieur à 0.5. Pour une huile régénérée, la norme ASTM D1500, tolère un indice de couleur inférieur à 2.

L'aspect, la couleur ainsi que l'odeur, nous permettent la détermination rapide de l'état général de l'huile. Par exemple : une mauvaise odeur renseigne sur la présence d'amorçages d'arcs électriques dans l'huile, une couleur sombre indique une dégradation de l'huile et l'aspect trouble signifie la présence d'eau et d'impuretés.

4.2.1.2. La viscosité :

Elle traduit les forces qu'opposent les molécules d'huile à une force tendant à les déplacer. En d'autres termes, elle mesure les frottements dus aux mouvements internes entre les molécules du fluide.

Il existe deux types de viscosité : la viscosité dynamique (en poise) qui visualise l'aspect microscopique et la viscosité cinématique (en centi stockes : cst) qui est la plus importante car elle traduit l'aspect macroscopique.

Le choix d'une huile dépend essentiellement de la valeur de sa viscosité à une température d'exploitation. C'est l'une des caractéristiques les plus importantes d'une huile.

Pour les transformateurs, le choix des huiles fluides est recommandé pour faciliter un échange thermique efficace, et l'évacuation rapide de la chaleur dégagée par les éléments actifs du transformateur.

Sa détermination se fait par l'écoulement capillaire dans des tubes normalisés nommés « Ubelhode ». Elle est donnée par la relation : $Cst = \Delta t \times coef$ où t est le temps d'écoulement en secondes et $coef$ est un coefficient qui dépend du tube utilisé.

La viscosité varie très peu avec la durée du service et elle dépend essentiellement de la température pour laquelle elle est mesurée. Généralement, elle est déterminée pour deux températures 20°C et 40°C. Des études recommandent d'effectuer la mesure de la viscosité pour une température plus élevée (90°C par exemple) [47]. Si la valeur de la viscosité cinématique est hors normes pour les deux températures (basse et élevée), l'état de contamination de l'huile est considérable et un traitement ne permet pas de la remettre à un niveau acceptable. Si par contre elle n'est insuffisante qu'à haute température, il est possible

de déduire la présence d'eau à concentrations tolérables et de particules solides précipitables à froid. La norme CEI 296 définit une bonne viscosité comme étant inférieure à 11 Cst (à 40°C)

4.2.1.3. La densité :

C'est une caractéristique propre à l'huile, elle permet d'apprécier sa fluidité et donc son efficacité pour la dissipation de la chaleur. Elle ne change que très peu durant l'exploitation et permet au côté de l'indice de réfraction et de la masse volumique de classer l'huile [46].

4.2.1.4. Le point éclair :

Cette propriété renseigne sur la présence de certains Hydrocarbures gazeux dissous dans l'huile suite à un défaut électrique ou thermique. Elle correspond à la température pour laquelle l'huile émet un flash, c'est à dire, la température critique maximale à ne pas dépasser au risque de provoquer un incendie dans le transformateur.

D'après la norme CEI 296, le point éclair d'une huile diélectrique doit rester supérieur à la valeur de 130°C.

4.2.2. Les propriétés chimiques :

4.2.2.1. L'oxydation :

C'est un ensemble de réactions chimiques lentes au sein du liquide qui sont dues à une température de service élevée. Dans ces conditions de température, le fer et le cuivre jouent un rôle de catalyseurs favorisant ainsi le processus d'oxydation de l'huile. Ces réactions en chaîne donnent naissance à des produits acides et à des produits de décomposition dissous dans l'huile (alcools, aldéhydes et acides). Leur propriété polaire a pour conséquence directe l'augmentation des pertes diélectriques et de l'acidité de l'huile [48]. L'oxydation des huiles engendre également un accroissement de la viscosité et défavorise donc l'évacuation de la chaleur. Elle conduit également à la décomposition du papier isolant des enroulements et par conséquent à la formation de boues.

4.2.2.2. L'acidité :

La détermination de l'acidité a une grande importance. Elle permet de suivre l'altération de l'huile, de donner une idée claire sur son degré d'oxydation et de décider du moment où elle doit être mise hors service. Elle est due à la présence de produits hydrocarbures acides tels que les Carboxyliques et les Hydroxyliques dont la formation est favorisée sous l'effet de la température, de la pression et de l'oxygène.

Sa mesure se fait par un titrage acido-basique. Elle est déterminée en mg KOH/g. D'après la norme CEI 296, une bonne huile possède une acidité inférieure à 0.03 mg KOH/g.

4.2.2.3. La teneur en eau :

L'eau est un puissant catalyseur qui favorise l'altération de l'huile. Elle influe sensiblement sur ses autres caractéristiques (la rigidité diélectrique, le facteur de dissipation et la résistivité). C'est un polluant indésirable qu'il faut éliminer. Les huiles isolantes doivent être anhydres et leur humidité doit être contrôlée en permanence. La norme CEI 296 tolère une teneur en eau maximale de 30 p.p.m. (partie par million d'unité). Cependant, quelque soient les précautions prises, de l'eau réussit toujours à s'infiltrer à travers l'appareillage. Pour

minimiser ces infiltrations, la plupart des transformateurs sont munis de tamis moléculaires (Silica Gel). Un diélectrique très chargé en humidité peut assurer un fonctionnement électrique (même s'il est mauvais) tant que sa température assure une solubilité suffisante car ce qui constitue un danger réel pour l'huile c'est la présence d'eau sous sa forme liquide [46].

4.2.3. Les propriétés électriques :

4.2.3.1. La permittivité :

C'est une caractéristique électrique liée à la formation du liquide. Pour une huile diélectrique, la permittivité exprime la possibilité de libérer des charges sous l'effet d'un champ électrique. Elle est définie comme étant le rapport de la capacité (C) d'une cellule remplie de liquide et la capacité (C_v) de la même cellule ayant le vide comme diélectrique : $\xi = C/C_v$. Elle est déterminée pour une température et une fréquence données (en général : 90°C et 50Hz).

Une bonne huile minérale possède une constante électrique comprise entre 2 et 2.2 [49].

4.2.3.2. Le coefficient de dissipation :

Tout diélectrique soumis à une tension continue ou alternative est toujours le siège de pertes électriques qui se traduisent par un échauffement plus au moins important du liquide. La cause principale de ces pertes est la présence d'un courant de fuite qui traverse le liquide sous l'effet du champ électrique.

Le facteur de dissipation diélectrique "tan δ " d'un matériau isolant est la tangente de l'angle complémentaire du déphasage entre la tension appliquée et ce courant de fuite (l'angle de perte). Elle permet d'apprécier la pureté d'une huile et dépend de la température et de la fréquence. D'après la norme CEI 296, une huile neuve possède un facteur de dissipation au plus égal à 0.005.

4.2.3.3. La rigidité diélectrique :

Elle correspond à la valeur maximale du champ électrique qui peut être appliqué à un isolant sans qu'il y ait passage d'un arc électrique entre les deux électrodes. Toutes les huiles diélectriques sont de bons isolants mais leur résistance au claquage dépend de la nature de l'huile et de sa qualité. Cependant, quelque soit la qualité d'une huile, sa rigidité est influencée par la présence de polluants extérieurs qui la réduisent fortement.

La détermination de la rigidité n'est en rien un critère de la qualité de fabrication de l'échantillon mais un essai conventionnel destiné à déceler son état.

La norme CEI 296 recommande, pour les huiles traitées, une tension de claquage qui soit supérieure à 50kV.

4.3. L'ALTERATION DES HUILES DE TRANSFORMATEURS :

Malgré les avantages et les performances des huiles minérales, elles finissent par se détériorer sous l'effet de l'action combinée ou séparée des contraintes électriques, chimiques ou thermiques auxquelles elles sont soumises. Ce phénomène est connu sous le nom de vieillissement. Il se traduit par une lente dégradation des propriétés du matériau.

L'huile de transformateur subit trois sortes de vieillissement :

- un vieillissement électrique qui provient de l'action des décharges partielles et des phénomènes de polarisation dont l'origine est l'exposition prolongée au champ électrique ;
- un vieillissement électrochimique qui résulte de l'action de certains agents chimiques soumis à un champ prolongé. La présence de ces agents dans l'huile est soit accidentelle, soit due à sa propre dégradation ;
- un vieillissement thermique dû aux températures élevées aux quelles est soumise l'huile de façon continue ou périodique pendant sa mise en service.

Le diagnostic permet de déterminer l'état interne du transformateur et de détecter les causes de ses défaillances. La décision est prise en se basant sur les recommandations des normes [Tab.2], les connaissances préalablement acquises et l'expérience.

Propriété	Valeur normalisée
Indice de couleur	≤ 2
Pertes diélectriques	$< 0,9$
Indice d'acidité [mgKOH/g]	$< 0,1$
Viscosité cinématique [CST]	$\geq 12,5$
Rigidité diélectrique [kV/cm]	≥ 50
Teneur en eau [ppm]	< 40

Tab.2 : Les normes recommandées par CEI pour les huiles des transformateurs ayant une tension comprise entre 72kV et 170kV [1]

Les méthodes de diagnostic sont nombreuses. Les méthodes internes cherchent à modéliser les systèmes tandis que les méthodes externes se basent sur le facteur d'expérience. Pour la détermination des avaries des transformateurs et l'établissement d'un diagnostic des huiles, deux approches existent. La première, assez largement utilisée, se base sur l'analyse des gaz dissous dans l'huile. Elle est efficace pour la prévention des accidents car elle renseigne sur l'état de la partie active. Elle comporte trois phases successives : l'échantillonnage de l'huile, l'extraction des gaz et leur analyse. La deuxième approche se base sur les tests physico-chimiques. Elle permet de classer les huiles selon quatre catégories :

4.3.1. Les huiles neuves :

Leurs caractéristiques sont données comme suit [49] :

- | | | |
|--------------------------|-----------------|------|
| - permittivité | 2,1 ÷ 2,5 ; | 90°C |
| - résistivité | 20 ÷ 2000 GΩm ; | 90°C |
| - rigidité diélectrique | 30 ÷ 50 kV/cm ; | |
| - facteur de dissipation | 0,001 ÷ 0,005 . | 90°C |

4.3.2. Les huiles faiblement polluées :

Elles se caractérisent par de bonnes valeurs des propriétés physico-chimiques et une tension de claquage faible. Leur contamination est généralement due à la présence d'eau et de particules solides en suspension. Elles nécessitent un traitement physique (filtrage et séchage).

4.3.3. Les huiles moyennement contaminées :

Ces huiles nécessitent un traitement chimique (régénération) suivi d'un traitement physique. Leurs caractéristiques sont les suivantes :

- indice de couleur $2 \div 4$;
- indice d'acidité $0.03 \div 0.04$ g KOH/g ;
- facteur de dissipation $0.005 \div 0.05$.

4.3.4. Les huiles en dégradation avancée :

Elles sont fortement contaminées et possèdent des caractéristiques très éloignées des valeurs recommandées par les normes. Leur récupération est possible mais elle serait plus coûteuse que leur remplacement.

4.4. DIAGNOSTIC ET TRAITEMENT DES HUILES DE TRANSFORMATEURS :

Vu l'importance de l'huile dans un transformateur, elle doit faire l'objet d'une attention particulière. Les méthodes de prélèvement des échantillons sont décrites dans les normes CEI 475 (1975) et CEI 296 (1982).

La norme CEI 422 spécifie les essais de contrôle d'une huile neuve. Ces mêmes essais sont appliqués à l'huile de transformateur en service. Ces essais sont effectués dans des laboratoires et l'interprétation des résultats doit être faite par un spécialiste. La décision est primordiale car un mauvais jugement risque d'engendrer des dégâts matériels exposant des travailleurs à des accidents mortels, des pertes économiques colossales ainsi que l'arrêt du réseau et le retrait du transformateur.

Notre étude s'intéresse à la classification et au diagnostic de l'huile de transformateur utilisée par SONELGAZ : la BORAK22.

Il est possible de déterminer l'état interne cette huile à partir des résultats de six tests normalisés : l'indice de couleur, la viscosité cinématique, l'indice d'acidité, le facteur de dissipation, la rigidité diélectrique et la teneur en eau. A partir de leurs valeurs, l'huile est classée en une des quatre catégories et une décision du type de traitement est prise (garder l'huile, la filtrer, la régénérer et la jeter).

Toute huile contaminée peut être traitée mais le coût de revient du traitement et les performances de huile après traitement rendent cette solution non rentable. La décision de réformer l'huile est donc principalement économique.

Chapitre V

Diagnostic des huiles par les systèmes adoptés

5.1. LA COMBINAISON DES RESEAUX NEURONAUX ET DE LA LOGIQUE FLOUE POUR LE DIAGNOSTIC :

Durant cette dernière décennie, un grand nombre de chercheurs s'est intéressé à l'application de méthodes intelligentes au diagnostic. Leurs recherches ont conduit au développement de plusieurs techniques de diagnostic automatique, aussi bien dans le domaine médical que dans le domaine technique.

Dans un système industriel, les pannes et les défauts proviennent de trois grandes sources. La première est l'ensemble des composants du système, la seconde est le système de commande et de régulation et la dernière est les capteurs d'entrées et de sorties et leurs systèmes d'affichage. Dans un transformateur, la cause principale des pannes est le claquage du fluide diélectrique à cause de sa détérioration sous l'effet des conditions de fonctionnement.

Beaucoup de chercheurs ont consacré leurs recherches à l'application des RNA au diagnostic (Marcu et al. : 1999 ; Korbicz et al. : 1999 ; Mokhnache et Boubakeur. : 2000-2003). Leur but était de réaliser des systèmes experts intelligents qui diagnostiquent et déterminent les défauts des systèmes industriels. La logique floue a également été appliquée pour la détermination des défauts des systèmes d'isolation (Koscielny et al. : 1999).

Les dernières années ont connu un accroissement des travaux qui tentent de combiner la logique floue et les architectures neuronales (Patton et al. : 1999; Calado et Sa da Costa : 1999), pour créer des outils de diagnostic performant bénéficiant ainsi des avantages fournis par chaque technique.

Les techniques de la logique floue peuvent intervenir à plusieurs échelles dans les réseaux de neurones, qui eux aussi peuvent être introduits à différents niveaux dans le calcul des systèmes flous. Cette partie nous permettra de voir l'influence de chaque technique sur l'autre et d'étudier l'influence de certains paramètres d'apprentissage sur le comportement du système neuro-flou.

5.2. LES ENTREES ET LES SORTIES DES SYSTEMES :

La base de données utilisées est obtenue à partir des tests physico-chimiques appliqués aux transformateurs de puissance dont la tension maximale du réseau est limitée entre 72.5kV et 170kV (les transformateurs de catégories "B").

Le vecteur d'entrée est constitué des six caractéristiques physico-chimiques de l'huile : $X = [\text{Indice de couleur} , \text{Viscosité cinématique} , \text{Indice d'acidité} , \text{Rigidité diélectrique} , \text{Facteur de dissipation} , \text{Teneur en eau}]$. Le vecteur de sortie contient les quatre décisions qui peuvent être prises : $T = [\text{Changer} , \text{Régénérer} , \text{filtrer} , \text{Garder}]$.

Tous les réseaux à apprentissage neuronal étudiés ont été entraînés par une base de données de 96 exemples (annexe 1) qui ont été obtenus à partir de tests faits dans le cadre d'un stage effectué au sein de l'Unité de Traitement des Huiles de la Société de Réparation des Transformateurs Electriques du Centre (Transfo-Centre : unité de Blida) et ainsi que de la thèse de doctorat de M^{elle} Mokhnache [4].

5.3. LES ARCHITECTURES DES SYSTEMES NEURO-FLOUS ADOPTES :

5.3.1. Le système flou :

Les variables linguistiques d'entrée sont les six propriétés physico-chimiques de l'huile et la variable de sortie est la décision à prendre. Les fonctions d'appartenance utilisées pour les étapes de fuzzification (Fig.27) et de la défuzzification (Fig.28) ont été déterminées en se basant sur les recommandations de la norme CEI 296.

Le système flou utilise la méthode de Mamdani où l'implication floue et l'opérateur logique ET sont réalisés par la formation du minimum et l'opérateur logique OU est réalisé par la formation du maximum. Les règles inférées sont en parallèle, c'est-à-dire, liées par l'opérateur OU.

Les règles floues ont été extraites à partir de la base de données utilisée. Elles sont au nombre de 43 (annexe 2).

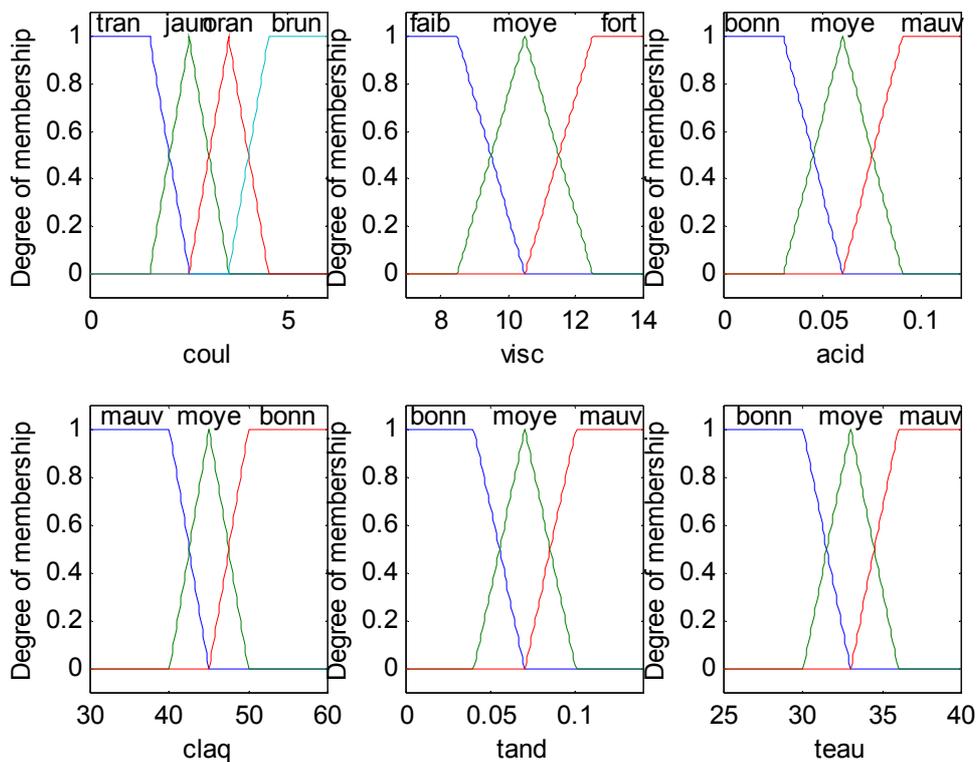


Fig.27 : Les fonctions d'appartenance des variables d'entrée du système.

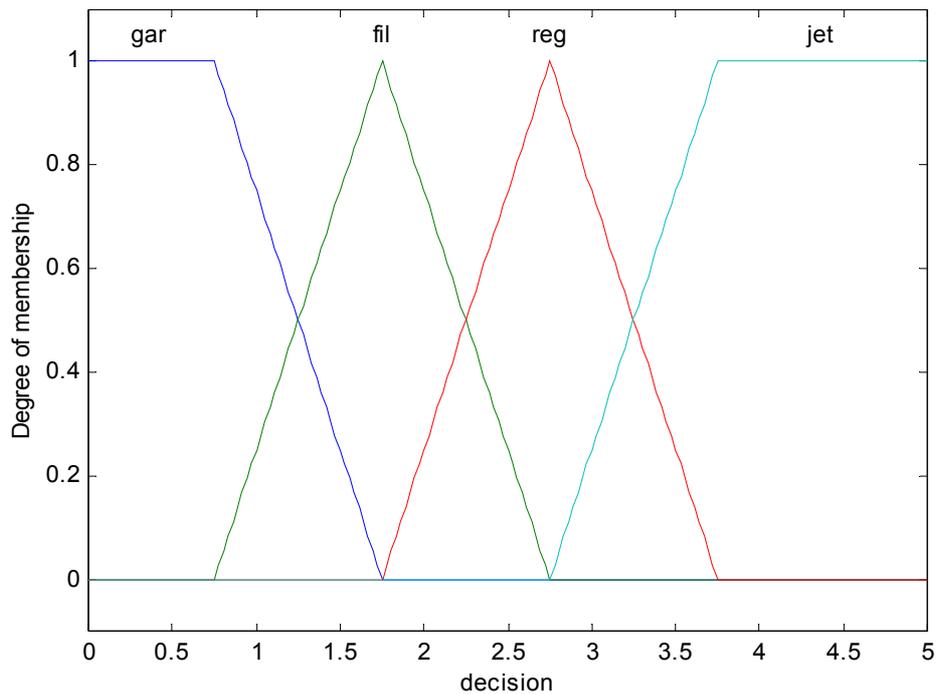


Fig.28 : La fonction d'appartenance de la variable de sortie du système.

5.3.2. Le réseau de neurones artificiels :

Le réseau de neurones artificiels utilisé comporte trois couches : une couche d'entrée de six neurones, une couche cachée de cinq neurones et une couche de sortie de quatre neurones (Fig.29).

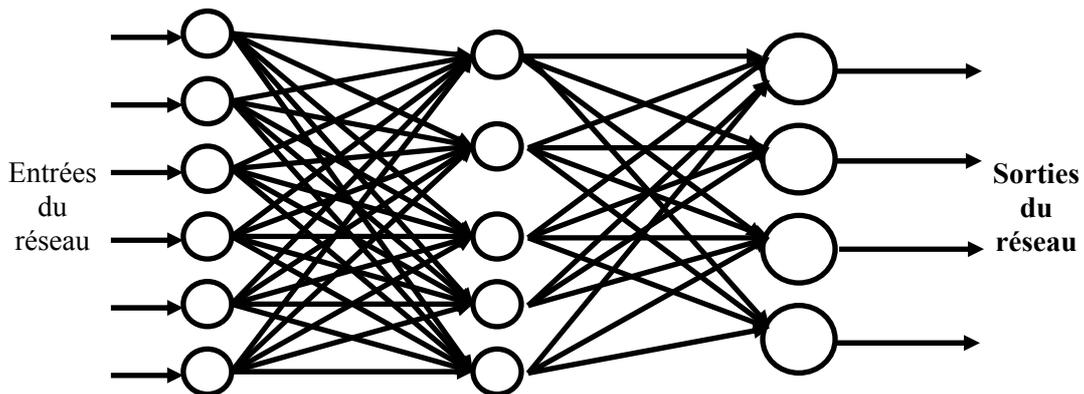


Fig.29 : Architecture du réseau de neurones artificiels.

5.3.3. Le réseau de neurones flous :

Le réseau de neurones flous a la même architecture que le réseau de neurones artificiels. C'est-à-dire, il est formé d'une couche d'entrée de six neurones, une cachée de cinq neurones et une couche de sortie de quatre neurones (Fig.30).

Le réseau est de type "min-max" (les neurones de la seconde couche sont de type "min" et les neurones de la dernière couche sont de type "max") et l'initialisation des poids et des biais est aléatoire.

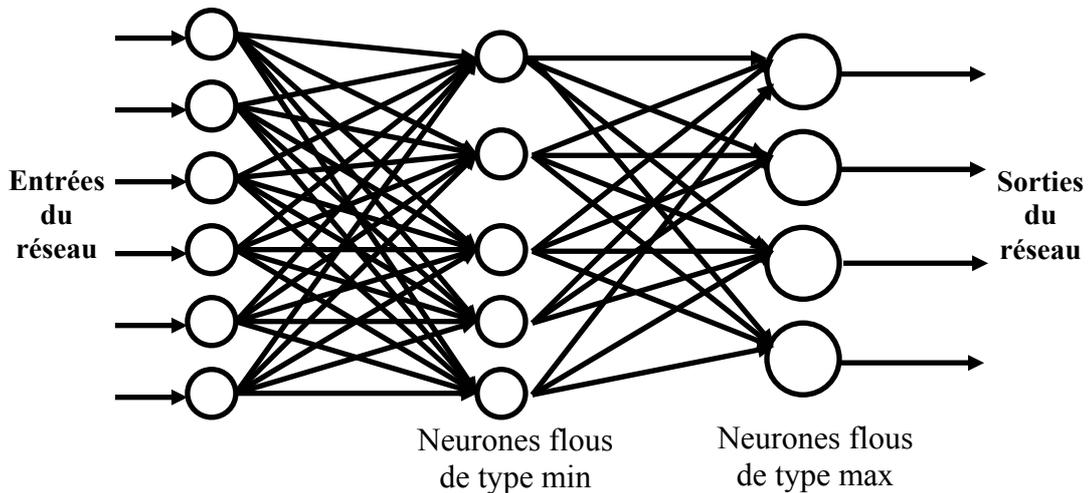


Fig.30 : Architecture du réseau de neurones flous de type max-min.

5.3.4. Le réseau de Pedrycz :

Le système comporte une partie de fuzzification des entrées suivie du réseau neuronal qui détermine les décisions à prendre. Ce réseau comporte trois couches. La couche d'entrée qui possède 19 neurones formels qui représentent les 19 termes flous (couleur transparente, viscosité élevée, acidité bonne ... etc) (Fig.31). La couche cachée comporte 5 neurones flous de type "min" (réalisation de l'opérateur logique ET) et la couche de sortie 4 neurones flous de type "max" (réalisation de l'opérateur logique OU).

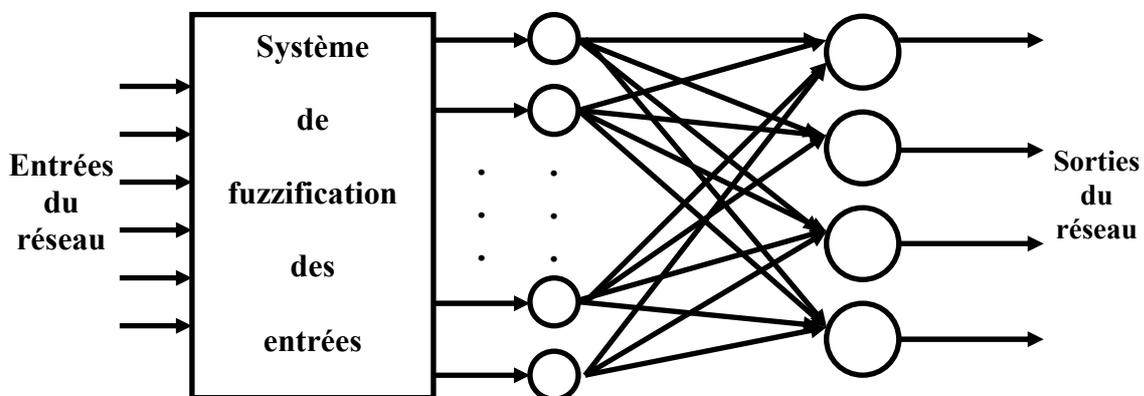


Fig.31 : Architecture du réseau de Pedrycz.

5.3.5. Le réseau de neurones à initialisation floue :

Le comportement d'un réseau de neurone est très lié à l'initialisation des poids et des biais. Une initialisation aléatoire n'est pas toujours efficace. Elle peut mener à des cas d'indécision.

Pour réduire ces cas d'indécision, nous avons pensé à choisir ces valeurs initiales de façon à guider et orienter le comportement du réseau.

Les poids initiaux à attribuer sont déterminés par un système flou qui les calcule à partir d'un vecteur d'initialisation et les injecte dans le réseau neuronal.

Le réseau contient deux couches : une couche d'entrée de six neurones et une couche de sortie de quatre neurones (Fig.32).

L'initialisation des poids prend en charge une seule mesure (les valeurs moyennes des exemples d'entrée pour chaque test) et lui attribue, selon sa valeur, un poids d'entrée et un poids par rapport à chacune des décisions. Ces valeurs numériques sont calculées par un ensemble de six systèmes flous. Ces systèmes ont pour entrée l'une des caractéristiques et pour sortie les poids synaptiques du neurone qui la représente.

Par exemple si la mesure de l'acidité est mauvaise, le poids qui relie le neurone 3 (qui représente la mesure de l'acidité) de la couche d'entrée et le neurone 1 de la couche de sortie (c'est le neurone qui représente la décision : jeter) est alors important. Si, par contre, l'acidité est moyenne alors le poids qui la relie au neurone 'jeter' est faible.

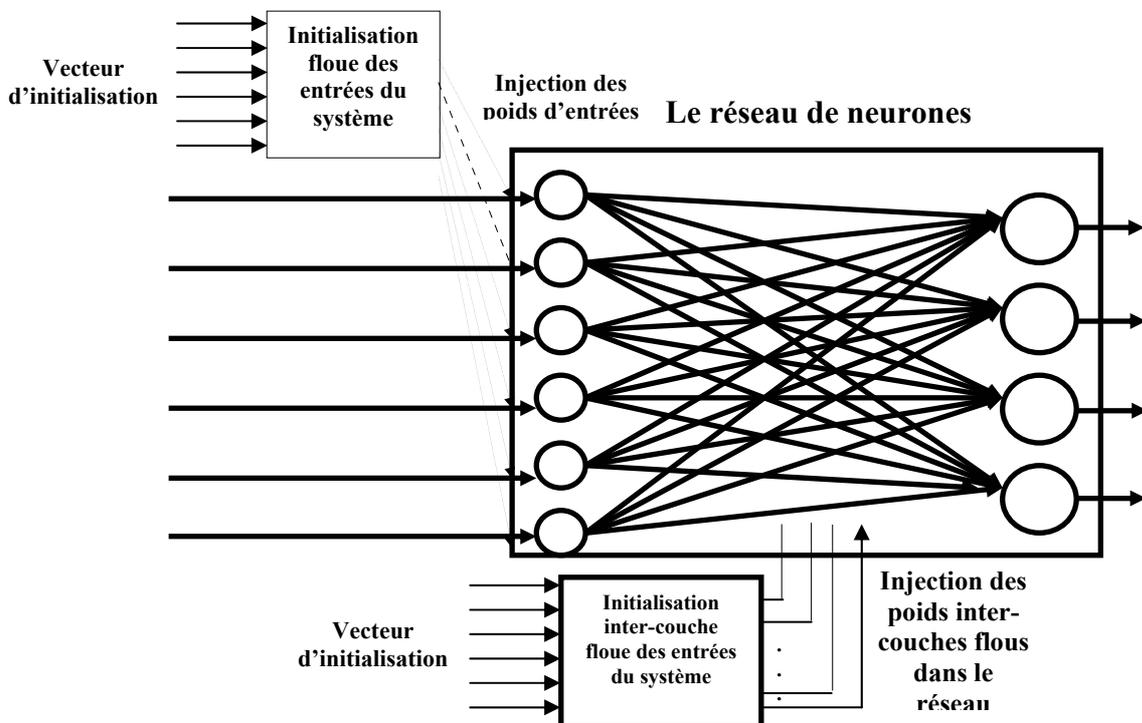


Fig.32 : Architecture du réseau à initialisation floue.

5.3.6. Les systèmes fusionnés :

Une structure fusionnée neuro-floue est un système flou dont les différentes opérations sont effectuées par un réseau de neurones. Ainsi les phases de : fuzzification, agrégation des règles floues et la prise de décision sont toutes réalisées par une architecture neuronale. Ce type de structures permet un apprentissage et une réadaptation du système pour toute variation de l'environnement du travail, tout en gardant la structure du raisonnement flou.

Deux structures ont été réalisées. La première est un système où toutes les opérations floues sont réalisées par un seul réseau de neurones (structure fusionnée compacte) et la seconde est un ensemble de plusieurs réseaux de neurones qui se partagent les différentes tâches (structure fusionnée dissociée).

5.3.6.1. La structure fusionnée compacte :

Cette structure est un réseau de neurones de quatre couches. Chacune correspond à une étape du raisonnement flou (Fig.33).

La couche d'entrée possède six neurones. Chacun représente l'une des six variables linguistiques (couleur, viscosité ...).

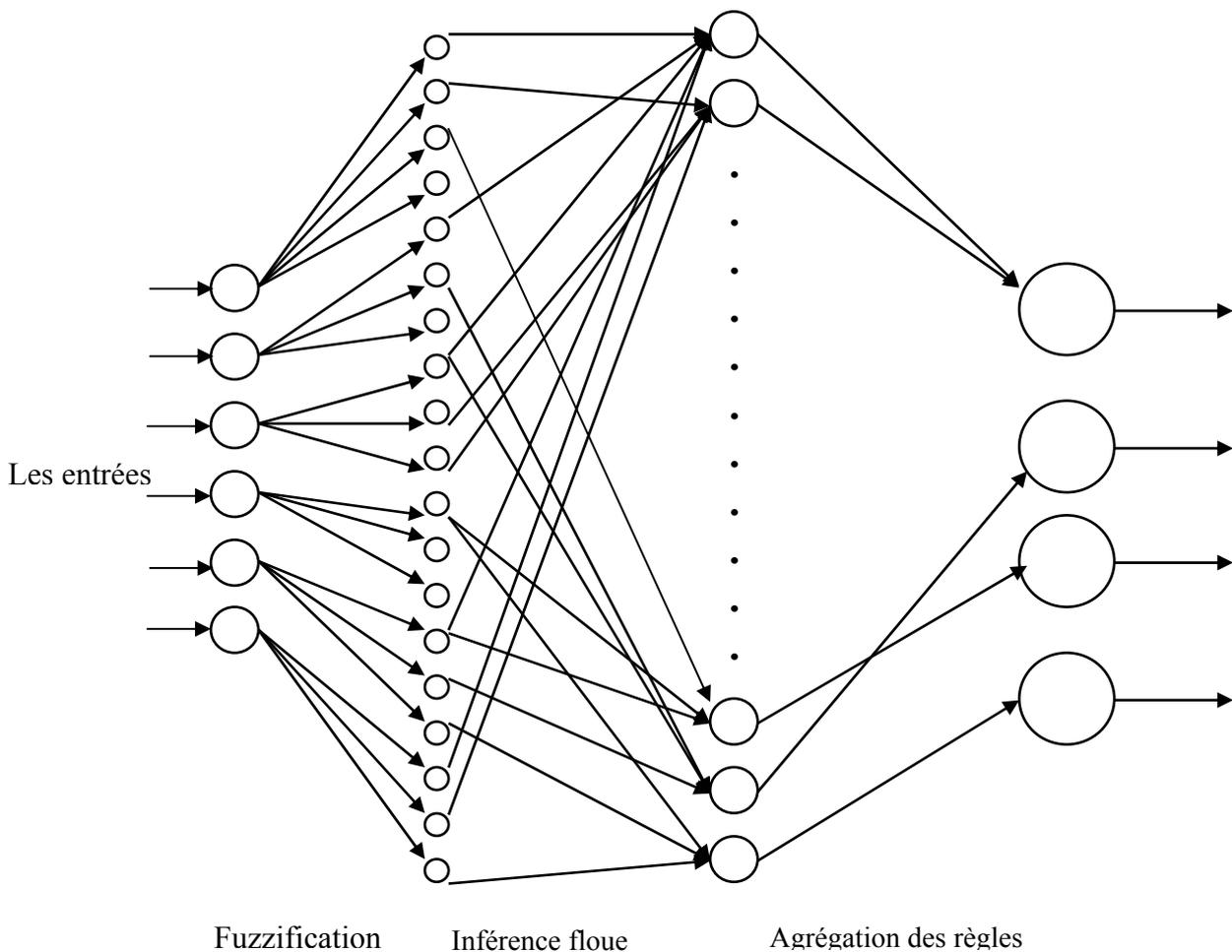


Fig.33 : Architecture du réseau fusionné compact.

La seconde couche contient 19 neurones qui correspondent à l'ensemble des sous-ensembles flous définis pour toutes les entrées (ceux définis pour le système flou). Ainsi, les quatre premiers neurones de cette couche représentent les quatre sous-ensembles flous de la couleur : transparente, jaune, orange et brune, les trois suivants, ceux de la viscosité à savoir : basse, moyenne, élevée et ainsi de suite.

L'initialisation des poids de cette première couche cachée découle de ce raisonnement. Les poids qui relient le neurone : couleur aux neurones transparente, jaune, orange et brune sont égaux à 1 et ceux qui le relient aux autres termes linguistiques (ceux des autres tests) sont nuls. Le même principe est utilisé pour les autres variables linguistiques.

La troisième couche modélise les inférences floues. Elle est constituée de 43 neurones qui correspondent aux 43 règles floues définies.

Chaque neurone est lié à certains neurones de la couche précédente de telle sorte que la combinaison des sous-ensembles flous représentés par chaque neurone de la seconde couche donne la règle représentée par le neurone de cette couche. Ainsi, la combinaison du 1^{ier} neurone (la couleur est transparente), du 7^{ème} neurone (la viscosité est bonne), du 8^{ème} neurone (l'acidité est bonne), du 13^{ème} (la rigidité est bonne), du 14^{ème} neurone (tan(\square) est bonne) et du 17^{ème} neurone (la teneur en eau est bonne) constitue la première règle représentée par le premier neurone de cette couche. Le poids initial de ces liaisons est donc 1 et le poids initial de toutes les autres liaisons est 0.

La couche de sortie comporte les quatre neurones qui correspondent aux quatre décisions possibles.

Chaque neurone de cette couche est relié aux neurones des règles dont la partie conséquente correspond à la décision qu'il représente.

Le premier neurone de la couche des règles sera donc relié au quatrième neurone de sortie (qui correspond à la décision : garder). Le poids initial de cette liaison sera donc 1 et tous les autres seront nuls.

L'étape de défuzzification n'est pas modélisée car le système réalisé n'est pas un système de commande dont la sortie doit être une valeur numérique mais un système de diagnostic dont la sortie est une décision. La réponse que nous désirons obtenir du système est le degré d'appartenance de la variable linguistique 'décision' à chacun des sous-ensembles flous de sortie (jeter, régénérer, filtrer et garder).

5.3.6.2. La structure fusionnée dissociée :

Afin de remédier à au problème de la lenteur de la structure fusionnée compacte, nous avons pensé à subdiviser ce réseau en un ensemble de sous-réseaux. Chaque réseau de neurones artificiels réalise séparément, une seule opération.

Cette structure fusionnée dissociée comporte sept réseaux neuronaux élémentaires qui se complètent en accomplissant chacun une tâche unique.

5.3.6.2.1. Les réseaux de fuzzification :

La fuzzification des entrées est réalisée à partir de six réseaux de neurones. Chacun effectue la fuzzification de l'une des entrées.

Ces réseaux sont formés d'une couche cachée de cinq neurones et d'une couche de sortie dont le nombre de neurones est égal au nombre de sous-ensembles flous de la caractéristique à fuzzifier.

5.3.6.2.2. Le réseau d'inférences floues :

Les sorties des réseaux de fuzzification sont ensuite introduites dans un système d'inférence flou formé d'un réseau de neurones d'une couche cachée de cinq neurones et d'une couche de sortie de quatre neurones.

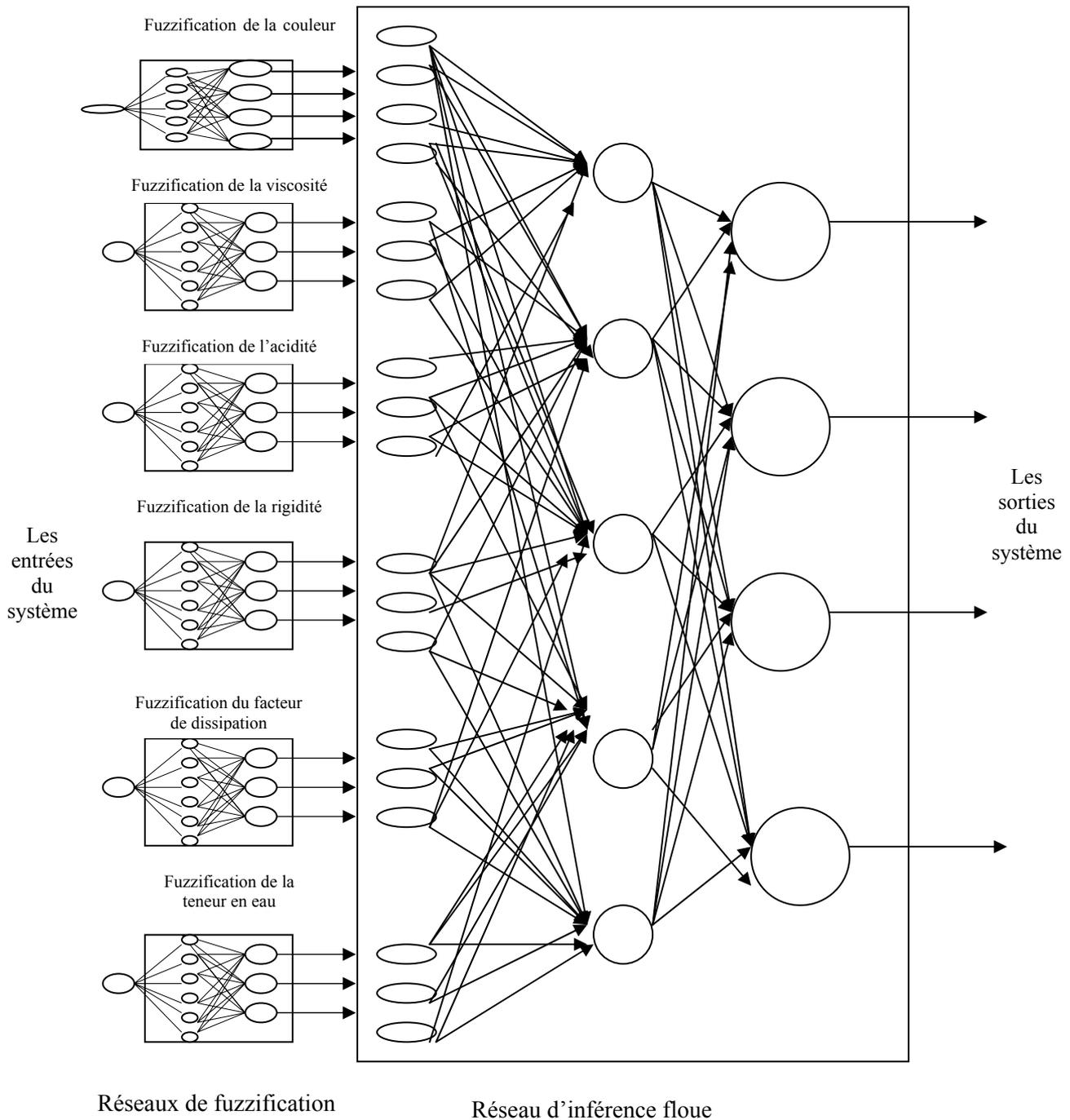


Fig.34 : Architecture du réseau fusionné dissocié.

5.4. LES RESULTATS DU DIAGNOSTIC DES HUILES PAR LES SYSTEMES ADOPTES :

Pour tous les réseaux de neurones utilisés, la fonction d'activation des neurones des couches est la fonction sigmoïde et leur fonction de sortie est la fonction linéaire (l'identité). L'entraînement des systèmes se fait par la méthode de Levenberg-Marquardt. Le taux d'apprentissage est égal à 0.7 et le facteur λ de la matrice hessienne est égal à 0.001.

Pour déterminer les capacités de généralisation et de prise de décision, cinq échantillons sont présentés à chacune des structures adoptées (Tab.3).

Echantillon n°	1	2	3	4	5
L'indice de couleur	0.7	2.3	4.5	2.9	3.4
La viscosité à 40°C (Cts)	10.23	10.87	11.19	10.97	11.5
L'indice d'acidité (mgKOH/g)	0.012	0.091	0.42	0.021	0.07
La rigidité diélectrique (kV/cm)	57	22	30	65	57
Le facteur de dissipation	0.072	0.019	0.55	0.018	0.105
La teneur en eau (p.p.m)	16	40	42	7	23

Tab.3 : Les échantillons proposés aux des structures étudiées.

Ces exemples sont d'abord analysés par un expert. Ses décisions sont :

Echantillon n°1 : l'huile est à garder car tous ces paramètres répondent aux exigences de la norme.

Echantillon n°2 : mis à part la rigidité et la teneur en eau, la mesure des autres paramètres est acceptable, cette huile nécessite un filtrage pour améliorer sa rigidité.

Echantillon n°3 : toutes les caractéristiques de cette huile sont détériorées, le traitement de cette huile n'est pas intéressant. Elle est à jeter.

Echantillon n°4 : les caractéristiques sont excellentes, l'huile est à garder.

Echantillon n°5 : le facteur de dissipation et la teneur en eau de cette huile sont hors normes mais les autres tests sont acceptables. Elle devra être régénérée.

5.4.1. Le diagnostic par la logique floue :

Les décisions que le système propose pour les deux premiers et le dernier échantillon sont correctes et celles proposée pour les deux autres sont fausses (Tab.4). Ceci est dû au fait qu'aucune règle n'est prévue pour ces cas deux de figure. Le système flou prend comme valeur de sortie, la valeur moyenne de l'intervalle où elle est définie ($D = 2.5$).

Echantillon n°	1	2	3	4	5
Jeter	0	0	0	0	0
Régénérer	0	0	0.75	0.75	1
Filtrer	0	1	0.25	0.25	0
Garder	1	0	0	0	0

Tab.4 : Les décisions prises par le système flou.

La logique floue est très intéressante pour le diagnostic si et seulement si, toutes les règles sont connues et introduites dans le système de diagnostic flou. Lorsque la base des règles n'est pas complète le système n'est pas capable de diagnostiquer n'importe quel échantillon.

5.4.2. Le diagnostic par le réseau de neurones artificiels :

Au bout d'une vingtaine d'itérations, la variation de l'erreur n'est plus perceptible. Elle reste au voisinage de la valeur 0.06.

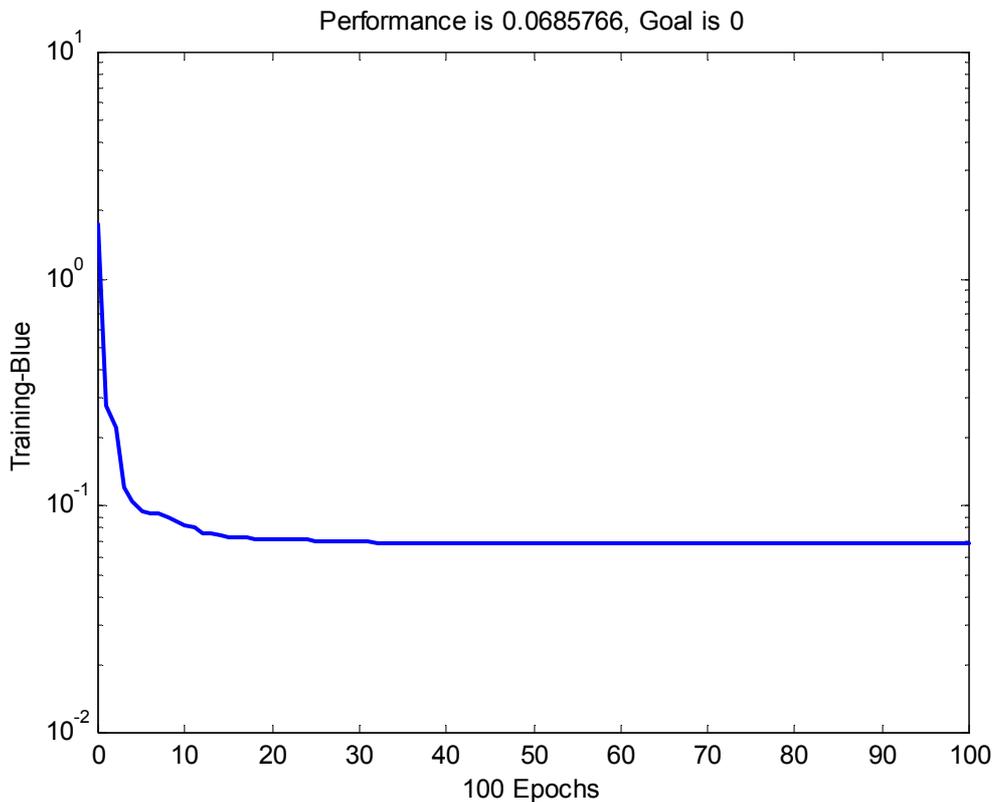


Fig.35 : L'erreur d'apprentissage du réseau de neurones artificiels.

Pour les échantillons 1, 3 et 4, les décisions prises par le réseau sont correctes. Par contre, les échantillons 2 et 5 présentent des indécisions (Tab.5).

Echantillon n°	1	2	3	4	5
Jeter	-0.0223	-0.0367	1.0344	-0.0010	0.2259
Régénérer	0.2037	0.6271	0.1251	0.0177	0.5236
Filtrer	-0.0563	0.6601	-0.1992	0.0735	-0.2800
Garder	0.8749	-0.2505	0.0396	0.9098	0.5304

Tab.5 : La génération du réseau de neurones artificiels.

Les réseaux de neurones sont très sensibles. L'augmentation de la taille de la base de données réduit le nombre de mauvaises décisions. Ceci-dit, elle engendre le problème de la saturation du calcul de l'erreur. De plus, l'initialisation aléatoire des poids et des biais peut mener à des cas d'indécisions.

5.4.3. Le diagnostic par le réseau de neurones flous :

Le remplacement des neurones formel par des neurones flous (structure max-min) engendre une stagnation de l'apprentissage au bout de moins de dix d'itérations à une erreur toujours supérieure 10^{-1} (Fig.36). Pour la structure min-max, la stagnation persiste mais elle apparaît plus tard et l'erreur est moindre.

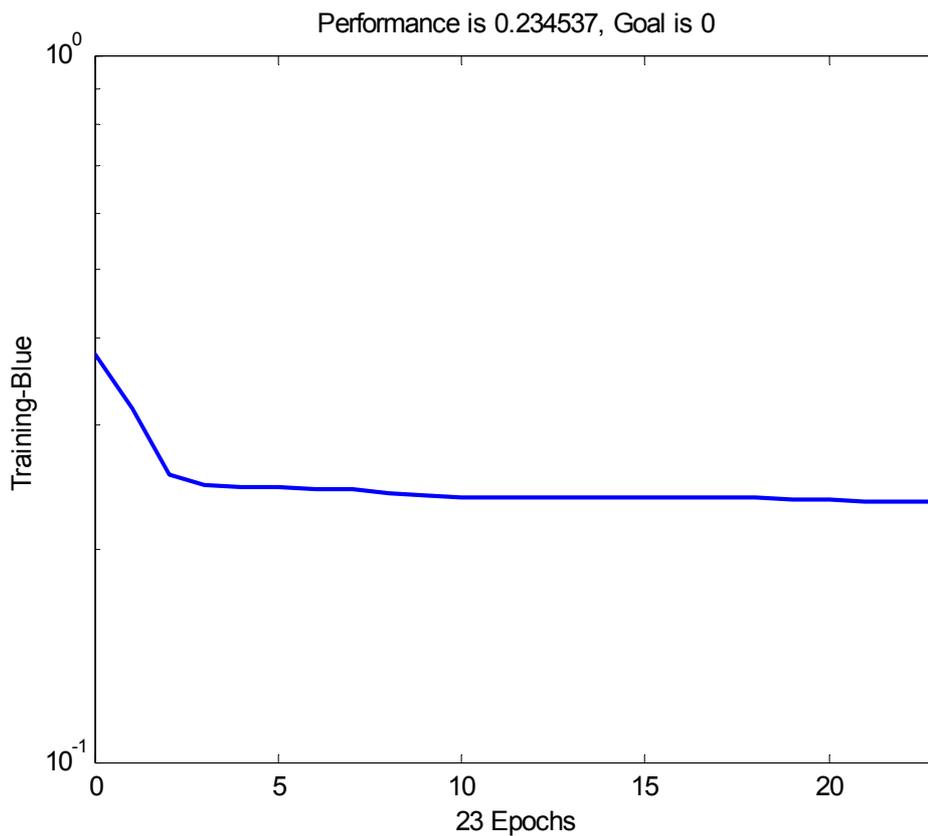


Fig.36 : L'erreur d'apprentissage du réseau de neurones flous (max-min).

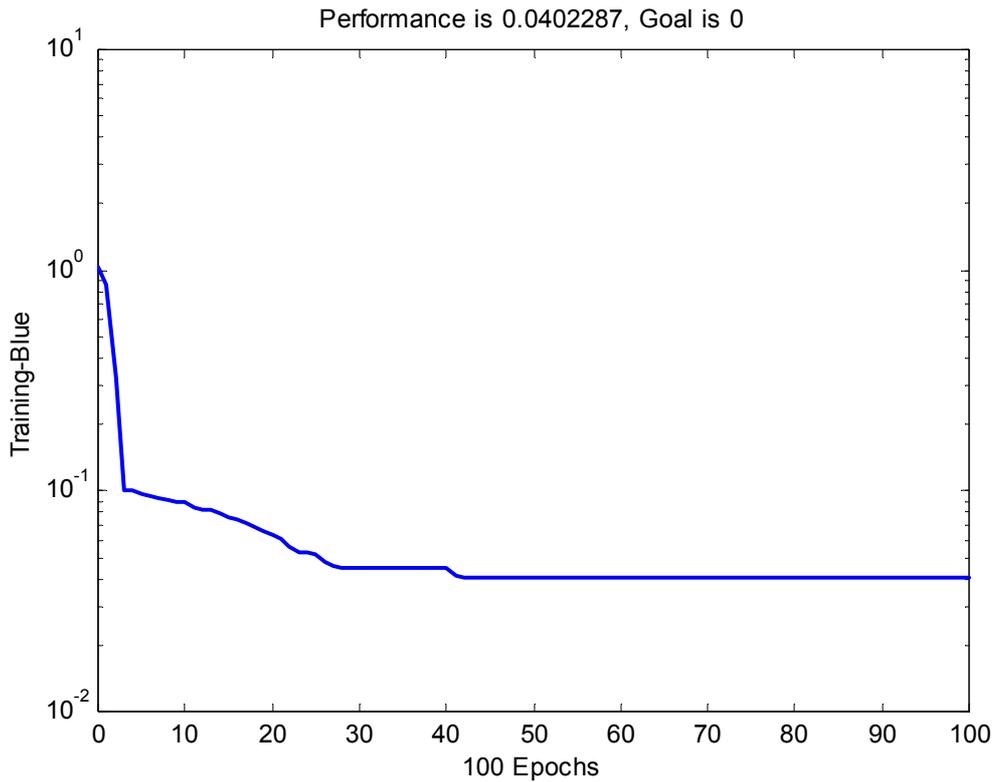


Fig.37 : L'erreur d'apprentissage du réseau de neurones flous (min-max).

La capacité de génération du réseau max-min est médiocre. Les décisions ne sont correctes que pour trois échantillons et dans ces cas, l'écart n'est pas assez élevé. Pour les deux autres échantillons, le réseau a fait un mauvais diagnostic. Il n'est pas fiable (Tab.6)

La génération de la structure min-max est meilleure mais elle n'est pas fiable pour autant. Le réseau a pris trois bonnes décisions mais présente des résultats obsolètes pour les deux autres échantillons (Tab.7).

Echantillon n°	1	2	3	4	5
Jeter	0.5758	0.5758	0.5758	0.5758	0.5758
Traiter	0.5281	0.4455	0.1856	0.5282	0.5259
Filtrer	0.2705	0.2705	0.2705	0.2705	0.2705
Garder	0.8759	0.0078	0.0078	0.8767	0.8449

Tab.6 : La génération du réseau de neurones flous (max-min).

Echantillon n°	1	2	3	4	5
Jeter	0.0372	-0.0202	0.6696	-0.0008	0.5972
Traiter	-0.0414	0.0169	0.3423	0.0632	0.2983
Filtrer	0.0257	0.9950	-0.9589	0.0145	-0.8461
Garder	0.9786	0.0083	0.9470	0.9231	0.9506

Tab.7 : La génération du réseau de neurones flous (min-max).

L'augmentation de la base de données améliore les capacités de génération du réseau. Néanmoins, elle favorise la stagnation du calcul de l'erreur qui influe considérablement sur l'apprentissage du réseau et diminue nettement sa capacité de génération. L'initialisation des biais et des poids n'engendre aucune amélioration des résultats du réseau.

Ces résultats incohérents sont dus au fait que nous avons utilisé les opérateurs de la logique floue, sans pour autant, orienter la structure du réseau vers l'architecture d'un système flou.

5.4.4 Le diagnostic par le réseau de neurones à initialisation floue :

Les courbes de l'erreur des réseaux de neurones à initialisation floue ont les mêmes allures que celles des réseaux à initialisation aléatoire. Cependant, l'erreur est légèrement plus faible et la stagnation est retardée. Elle intervient au bout d'une trentaine d'itérations pour le premier réseau et au bout de quatre-vingt itérations pour le second réseau (Fig.38 et 39).

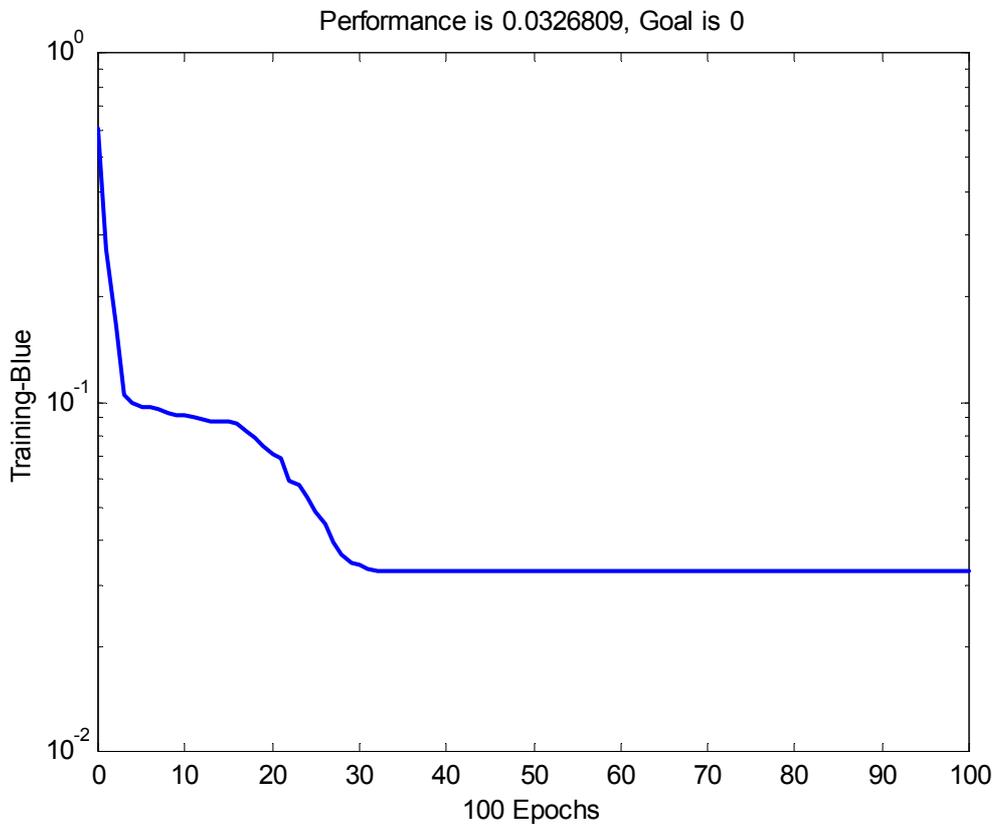


Fig.38 : L'erreur d'apprentissage du réseau de neurones à initialisation inter-couche floue.

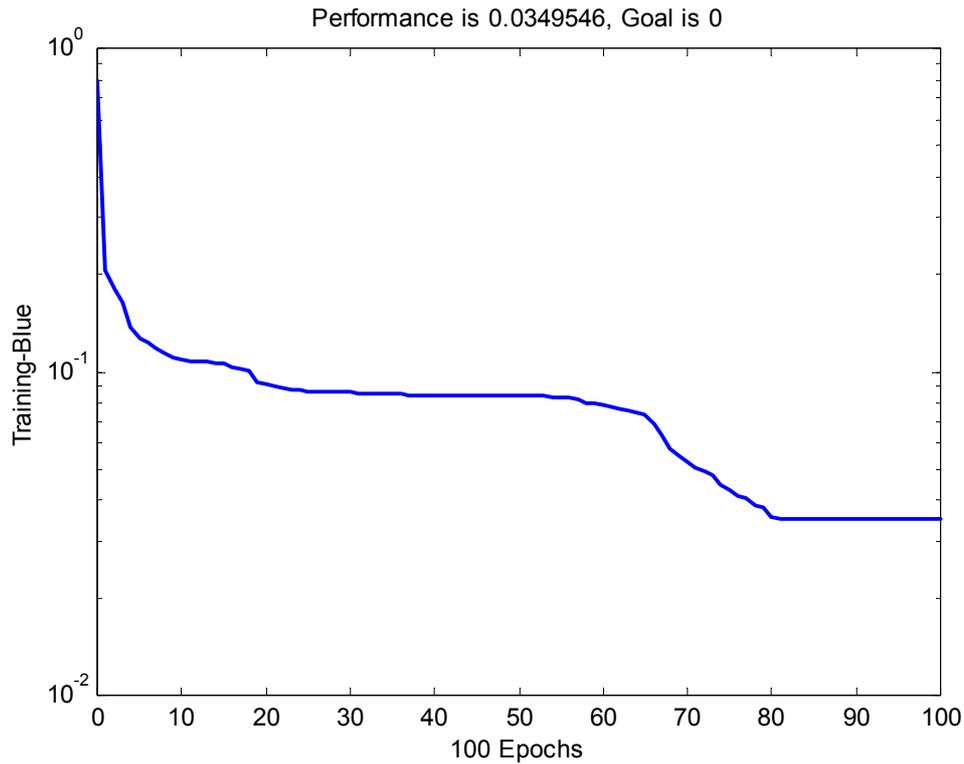


Fig.39 : L’erreur d’apprentissage du réseau de neurones à initialisation floue de tous les poids.

L’initialisation floue des poids a éliminé les cas d’indécisions. Toutes les décisions prises par ces deux réseaux sont correctes. Notons cependant que celles données par le réseau à initialisation floue de tous les poids sont plus démarquées (Tab.8 et 9) .

Echantillon n°	1	2	3	4	5
Jeter	-0.1155	-0.1838	0.7029	-0.0088	0.0535
Traiter	0.3905	0.2678	0.2379	0.0595	0.9758
Filtrer	0.0235	0.9831	-0.0526	0.0034	-0.0198
Garder	0.7015	-0.0671	0.1119	0.9458	-0.0096

Tab.8 : La génération du réseau de neurones à initialisation inter-couches floue.

Echantillon n°	1	2	3	4	5
Jeter	0.0054	-0.0306	0.9714	0.0822	-0.0044
Traiter	-0.0172	0.1757	0.1250	-0.0211	0.8805
Filtrer	0.0080	0.8082	-0.0761	-0.0598	0.1142
Garder	1.0038	0.0467	-0.0203	0.9987	0.0097

Tab.9 : La génération du réseau de neurones à initialisation floue de tous les poids.

Plus le réseau est assisté, plus sa capacité de génération s’améliore. Néanmoins l’initialisation floue des réseaux neuronaux ne permet pas de régler le problème de saturation de l’apprentissage. Les valeurs des biais sont aléatoires. Une choix plus étudié peut être envisagée. Il devrait améliorer davantage les résultats.

5.4.5. Le diagnostic par le réseau de Pedrycz :

L'apprentissage de ce réseau est rapide mais en revanche, le calcul de l'erreur stagne toujours pour une base de données de taille importante (Fig.40, 41 et 42).

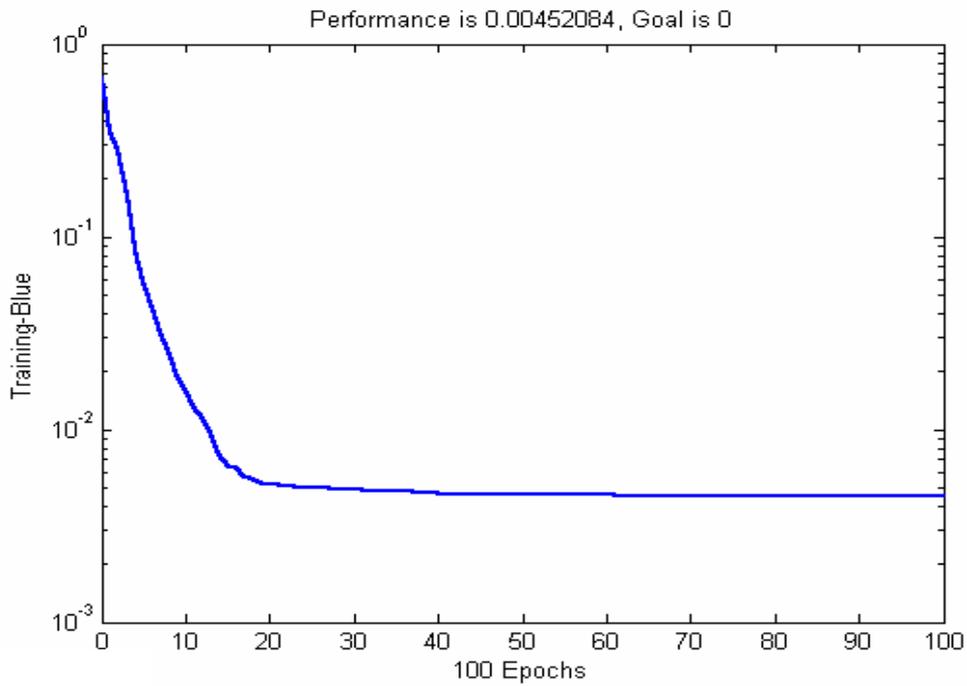


Fig.40 : L'erreur d'apprentissage du réseau de Pedrycz à neurones artificiels.

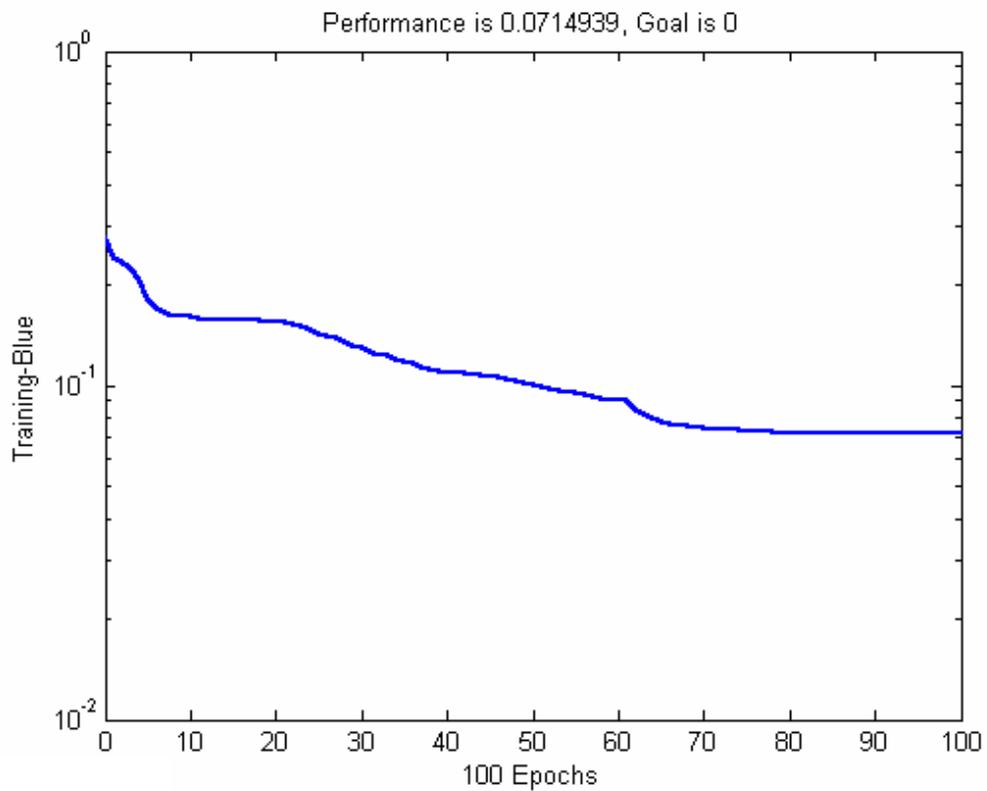


Fig.41 : L'erreur d'apprentissage du réseau de Pedrycz à neurones flous (min-max).

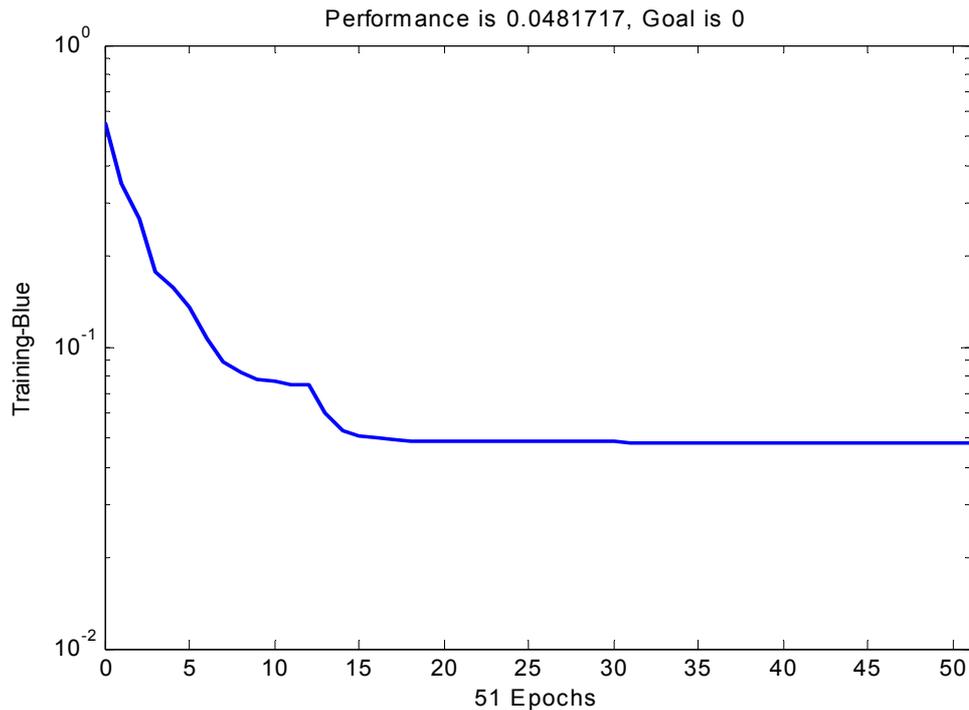


Fig.42 : L'erreur d'apprentissage du réseau de Pedrycz à neurones flous (max-min).

Les résultats obtenus lors du diagnostic par la structure de Pedrycz sont plus intéressants que ceux des réseaux neuronaux. Ceci est dû au fait que l'architecture du réseau n'a été choisie arbitrairement mais de manière à modéliser le raisonnement flou. Toutes les décisions du réseau de neurones artificiels sont bonnes (Tab.10). Ceci dit, l'utilisation de neurones flous de type min-max améliore nettement la génération du réseau (Tab.11).

Les réseaux de neurones flous de type max-min n'ont proposé que deux diagnostics corrects. Les autres sont soit des cas d'indécision ou carrément erronées (Tab.12). Ce réseau génère mal car le système flou que nous désirons modéliser réalise l'opérateur logique ET (respectivement OU) par la fonction minimum (respectivement maximum) et non le contraire.

Echantillon n°	1	2	3	4	5
Jeter	-0.0001	-0.0056	0.8900	0.0000	0.0048
Traiter	-0.0002	-0.0103	-0.0381	0.0578	1.3078
Filtrer	0.0006	0.9972	-0.0131	0.0010	0.0492
Garder	0.8996	0.0186	0.0612	0.9421	-0.4175

Tab.10 : La génération d'un réseau de Pedrycz à neurones artificiels.

Echantillon n°	1	2	3	4	5
Jeter	0.0000	0.0000	0.9832	0.0001	0.0030
Traiter	-0.0007	-0.0081	0.0001	0.0831	0.9745
Filtrer	-0.0070	0.9901	-0.2079	-0.0008	-0.0059
Garder	-0.9989	0.0005	0.0237	0.9985	0.0315

Tab.11 : La génération du réseau d'un Pedrycz à neurones flous de type min-max.

Echantillon n°	1	2	3	4	5
Jeter	-0.0077	0.1462	0.6342	-0.0373	0.4758
Traiter	0.0218	0.5967	0.5967	0.0001	0.5967
Filtrer	-0.0186	0.5066	0.1054	0.3116	-0.9328
Garder	0.9520	-0.0166	-0.0219	0.9520	0.9520

Tab.12 : La génération d'un réseau de Pedrycz à neurones flous max-min.

Le nombre d'itérations n'influe pas sur la stagnation du calcul de l'erreur car elle est due à la nature des neurones et à moindre degré, à la base de données.

L'initialisation des poids et des biais influe sensiblement sur le comportement du réseau, elle peut même engendrer sa divergence et donc son instabilité.

5.4.6. Le diagnostic par le réseau fusionné compact :

L'apprentissage de ce réseau est très lent (environ 30 minutes) à cause de sa structure encombrante et de la taille importante de la base de données. L'erreur stagne aux toutes premières itérations mais cette saturation est surmontée après 300 itérations. Au-delà, elle ne cesse de diminuer et atteint la valeur 10^{-7} à 1000 itérations (Fig.43).

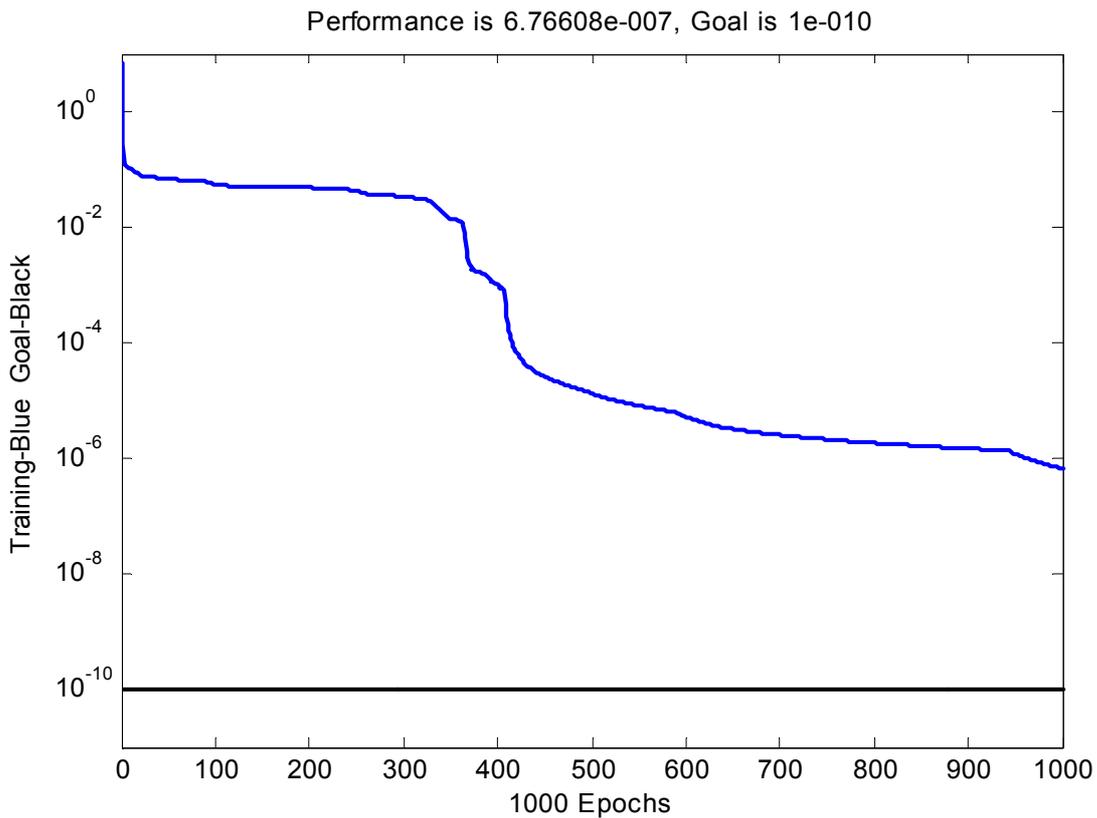


Fig.43 : L'erreur d'apprentissage du réseau fusionné compact.

Les résultats générés par ce réseau sont excellents. Ils correspondent tous à la décision de l'expert et ne présentent aucun cas d'indécision (Tab.13).

Echantillon n°	1	2	3	4	5
Jeter	0.0000	0.1469	1.0012	0.0000	-0.0002
Traiter	-0.0004	-0.3126	0.0002	0.0139	1.0005
Filtrer	-0.0002	1.1703	-0.0004	0.0004	0.0013
Garder	1.0005	-0.0046	-0.0010	0.9856	-0.0016

Tab.13 : La génération du réseau fusionné compact.

L'architecture du réseau et l'initialisation des poids sont préalablement fixées. Le seul paramètre qui peut influencer l'apprentissage et la génération est l'initialisation des biais. Il joue un rôle décisif dans la convergence mais n'influe pas sur la vitesse d'apprentissage et sur la saturation. Une initialisation aléatoire peut mener à des résultats erronés. Pour remédier à ce problème, nous avons cherché les meilleures valeurs initiales des biais.

La lenteur de l'apprentissage n'est pas un réel inconvénient car pour le diagnostic, l'entraînement du réseau ne se fait qu'une seule fois. Ensuite les poids finaux sont récupérés et la génération se fait à partir de leurs valeurs.

Cette structure serait inutilisable si l'entraînement du réseau devait se faire en temps réel.

5.4.7. Le diagnostic par la structure fusionnée dissociée :

Contre tout attente, l'entraînement de cette structure est rapide malgré le nombre important de réseaux qui le constituent. Ceci est dû à la taille réduite de ces réseaux élémentaires.

L'erreur d'apprentissage est considérablement faible par rapport aux autres structures. Elle est pour tous les réseaux inférieure à 10^{-3} et atteint en 81 itérations la valeur de 10^{-23} pour le dernier réseau (Fig.44, 45, 46, 47, 48, 49 et 50).

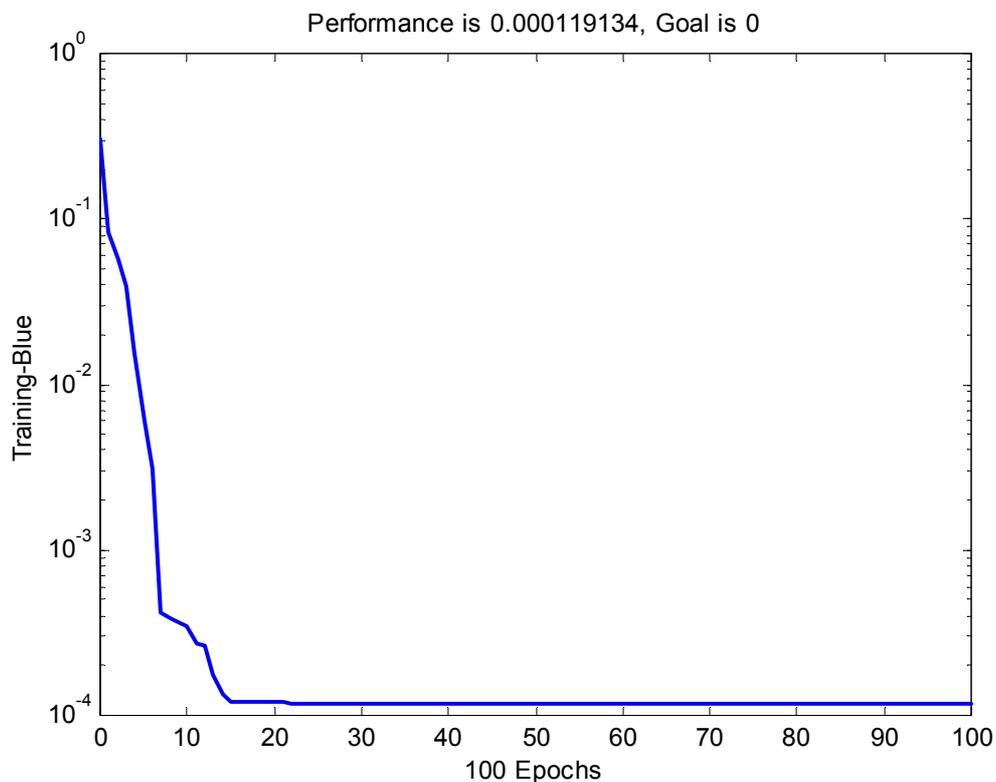


Fig.44 : L'erreur d'apprentissage du réseau de fuzzification de l'indice de couleur.

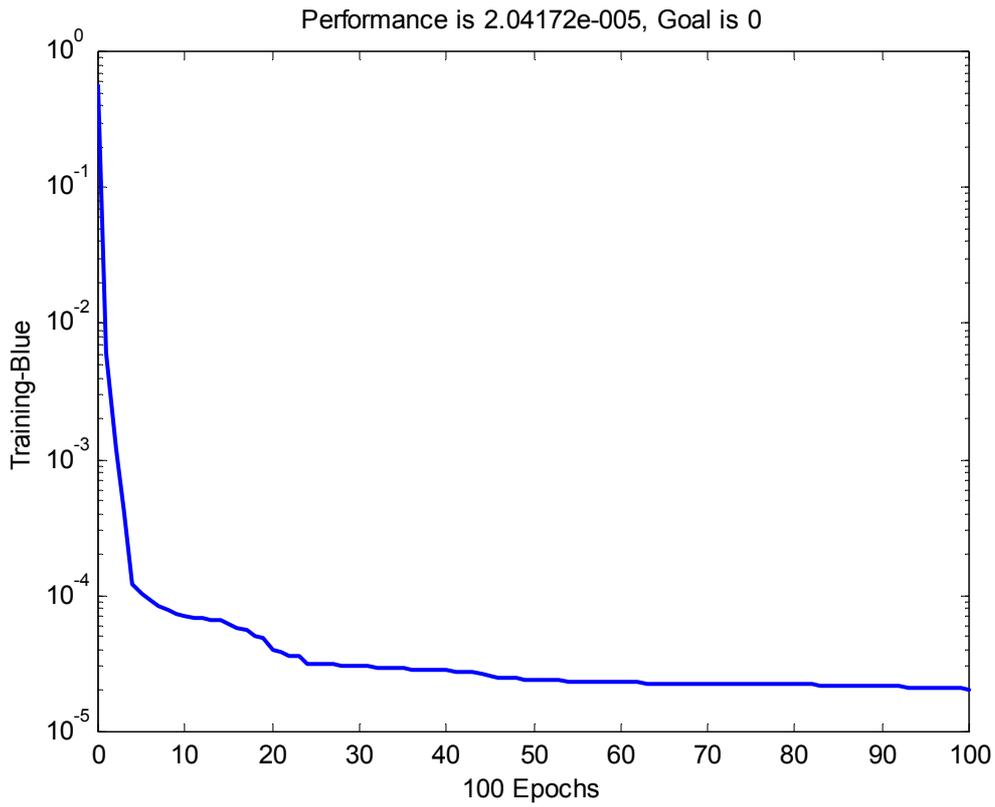


Fig.45 : L'erreur d'apprentissage du réseau de fuzzification de la viscosité cinématique.

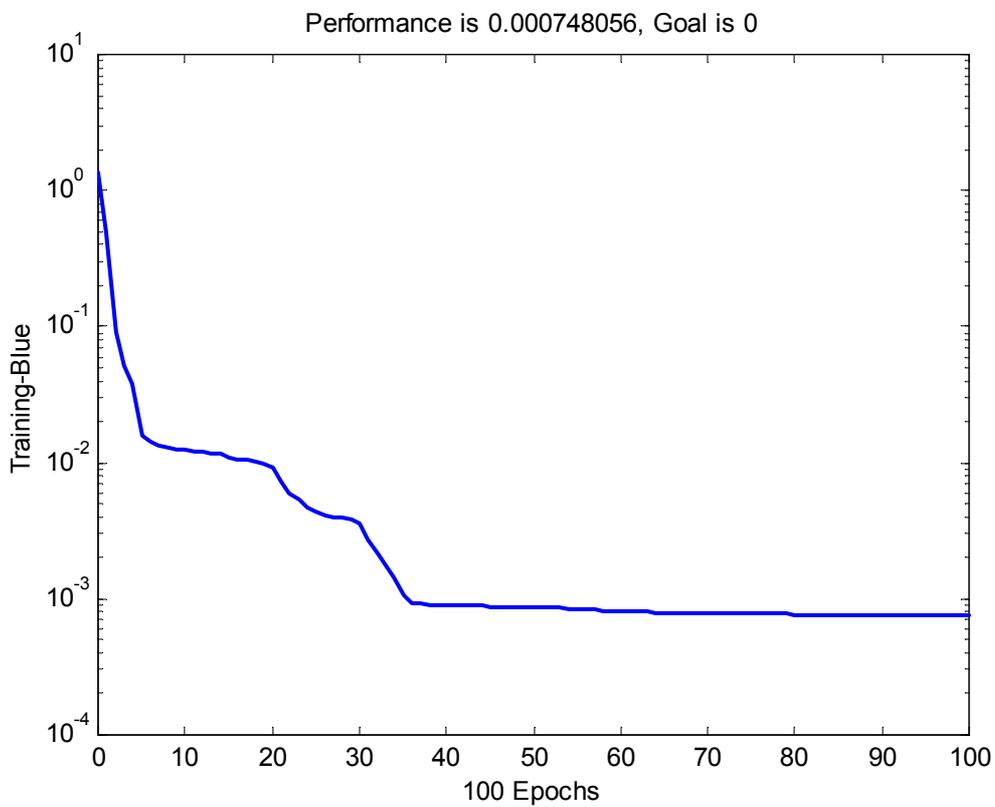


Fig.46 : L'erreur d'apprentissage du réseau de fuzzification de l'indice d'acidité.

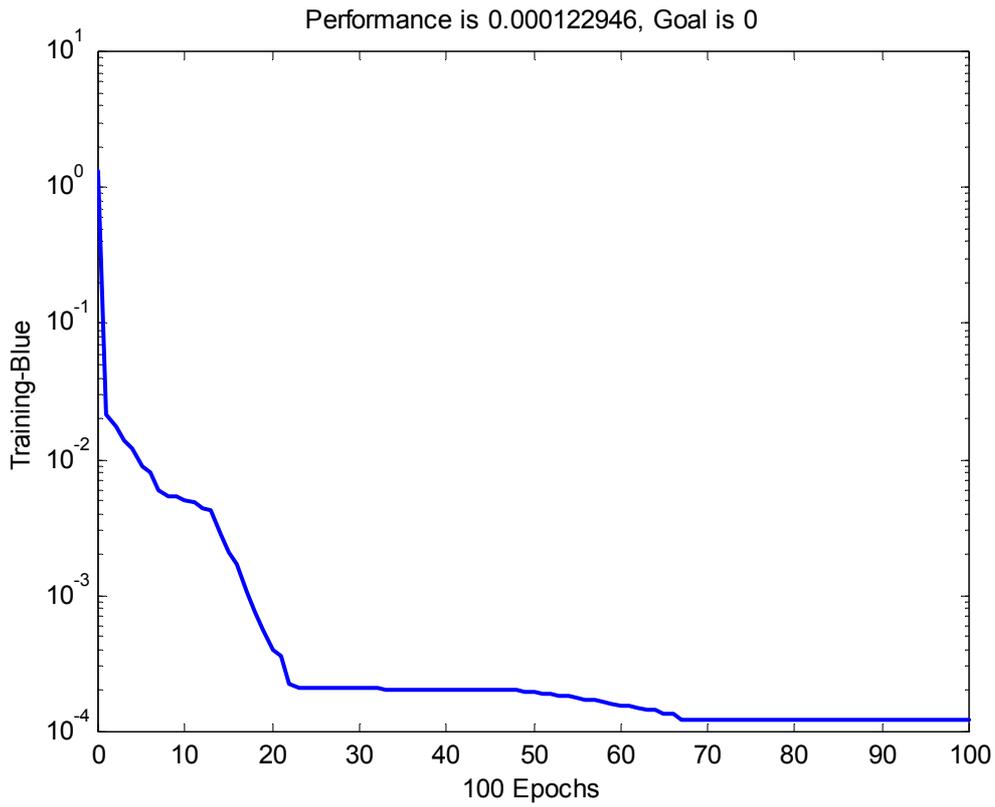


Fig.47 : L'erreur d'apprentissage du réseau de fuzzification de la rigidité diélectrique.

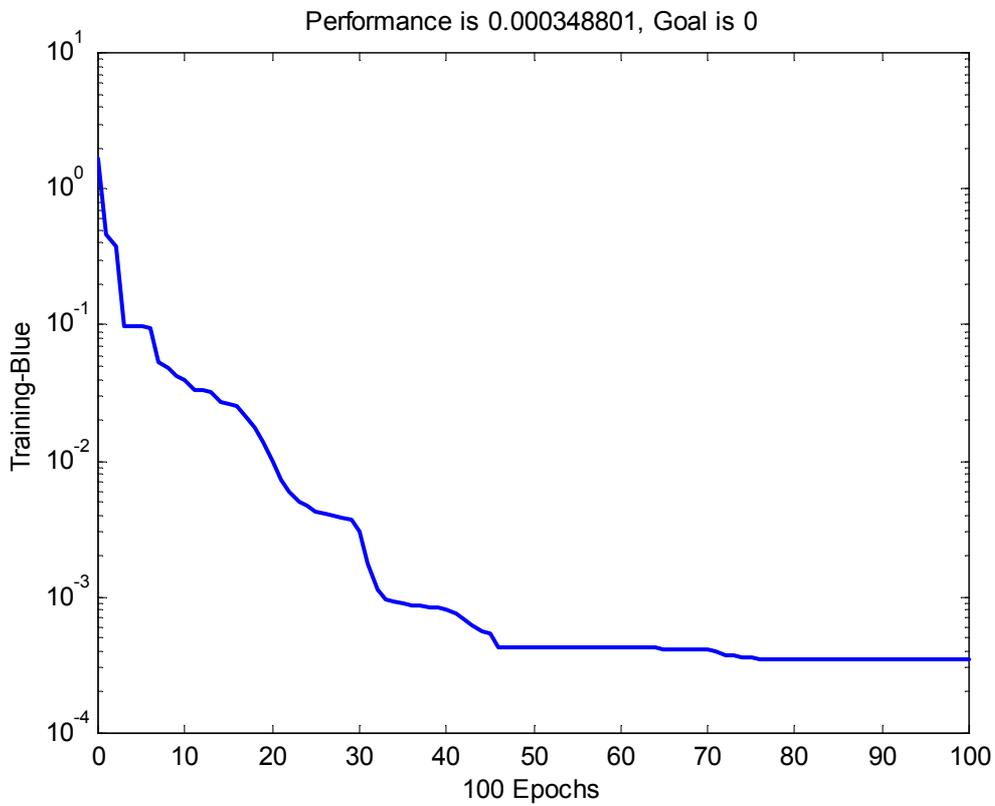


Fig.48 : L'erreur d'apprentissage du réseau de fuzzification du facteur de dissipation.

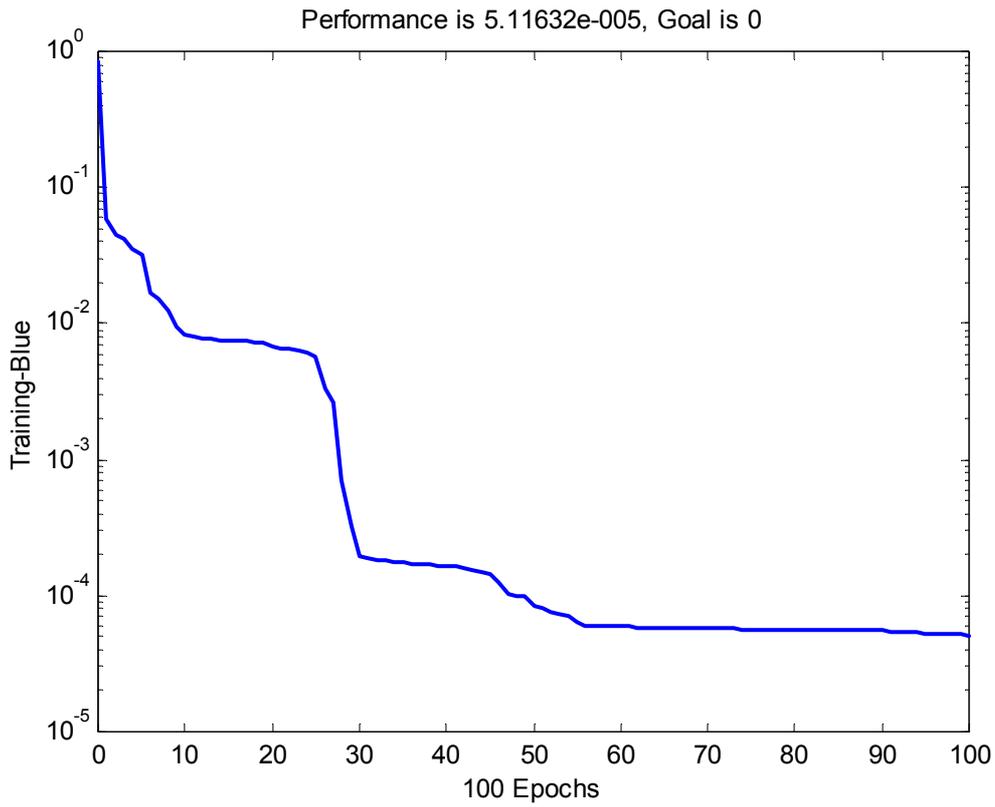


Fig.49 : L'erreur d'apprentissage du réseau de fuzzification de la teneur en eau.

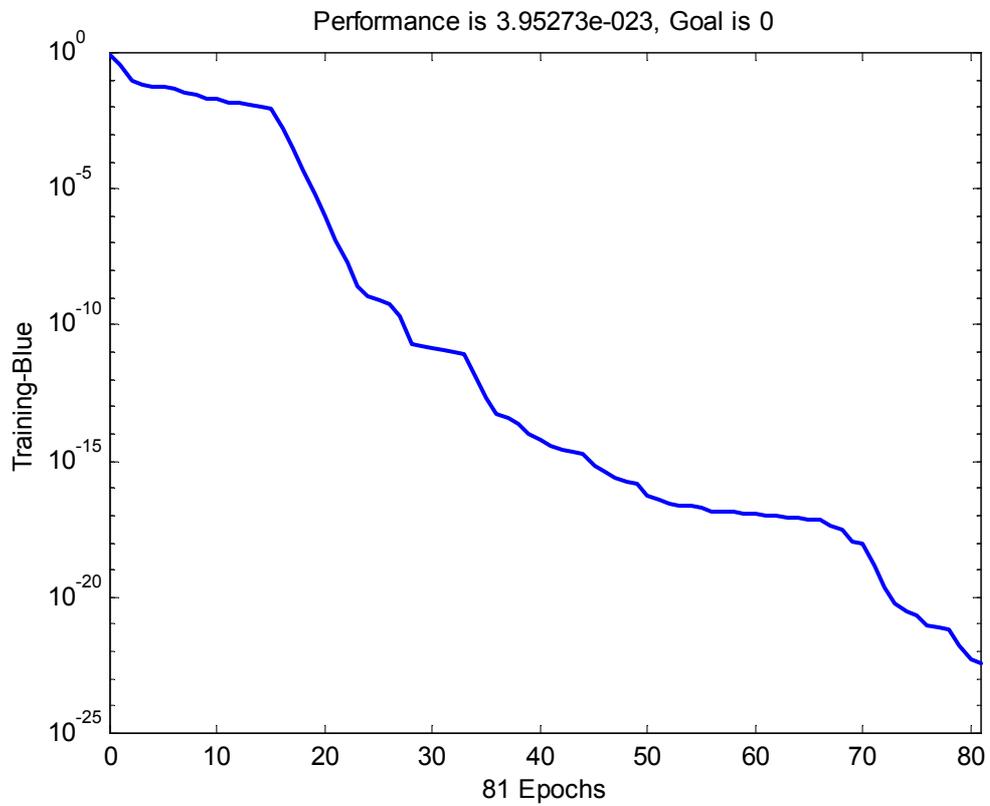


Fig.50 : L'erreur d'apprentissage du réseau d'inférence.

Chaque étape du raisonnement flou est traitée séparément. Les réseaux élémentaires de taille réduite, sont entraînés individuellement par un grand nombre d'exemples. Ceci mène à des résultats précis. Les décisions que cette structure propose correspondent parfaitement à celles données par l'expert (Tab.14).

Echantillon n°	1	2	3	4	5
Jeter	-0.0000	0.0000	1.0000	-0.0000	-0.0000
Traiter	-0.0000	-0.0000	-0.0000	-0.0000	1.0000
Filtrer	-0.0000	1.0000	-0.0000	-0.0000	-0.0000
Garder	1.0000	0.0000	0.0000	1.0000	-0.0000

Tab.14 : La génération du réseau fusionnée dissocié.

La variation des valeurs initiales des poids et des biais change le comportement du système. Les décisions données sont rarement fausses et des cas d'indécision peuvent apparaître.

5.5. DISCUSSIONS ET CONCLUSIONS :

D'après les résultats que nous avons obtenus (Tab.15), nous pouvons faire les commentaires qui suivent :

	Nombre de bonnes décisions	Nombre d'indécisions	Nombre d'erreurs
Système flou	3	0	2
RNA	3	2	0
RNF (max-min)	3	0	2
RNF (min-max)	3	0	2
Pedrycz (neurones formels)	5	0	0
Pedrycz (min-max)	5	0	0
Pedrycz (max-min)	2	2	1
Structure fusionnée compacte	5	0	0
Structure fusionnée dissociée	5	0	0
RNA à initialisation floue des poids inter-couches	5	0	0
RNA à initialisation floue de tous les poids	5	0	0

Tab.15 : Récapitulatif des résultats obtenus.

- La logique floue est une technique adaptée pour la classification et le diagnostic. En revanche, elle requiert une base de règles complète dont la détermination n'est pas toujours évidente. Si la base des règles n'est pas suffisante le système devient incapable de traiter tous les cas possibles. Ses résultats risquent d'être incohérents.

- Les réseaux de neurones artificiels sont également un bon outil d'aide à la décision mais les cas d'indécision et les réponses erronées sont assez fréquentes.

La base de données est un élément décisif dans le comportement des structures à architecture neuronale. Lorsqu'elle est petite, la capacité de génération du réseau est réduite. La majorité des sorties générées est erronée. L'augmentation de la taille de la base de données engendre une stagnation de l'erreur d'apprentissage. Le réseau devient lent et peut même diverger. Il y a donc un compromis à faire lors du choix du nombre d'exemples d'entraînement. Il doit y en avoir assez pour que le système génère les bonnes décisions et pas trop pour que le calcul de l'erreur ne stagne pas et que les décisions ne soit pas faussées.

- L'introduction des neurones flous ne donne de bons résultats que si l'architecture du réseau a un sens et qu'elle a été choisie en s'inspirant du raisonnement flou. Dans le cas contraire, aucune amélioration n'est apportée au système. Ainsi, un réseau de neurones flous dont la structure est arbitrairement choisie, génère des résultats obsolètes. Par contre, la structure de Pedrycz donne de meilleurs résultats lorsque les neurones sont flous.

L'utilisation des autres structures de neurones flous (KWAIN CAI, neurones flous compétitifs...) pourrait améliorer davantage la génération du réseau et régler le problème de saturation du calcul de l'erreur.

- L'initialisation des poids et des biais joue un rôle primordial dans le bon apprentissage du système de diagnostic et sa convergence.

L'initialisation floue des poids a conduit à une amélioration de la génération en éliminant les cas d'indécisions. Ce résultat n'est pas systématique à cause des valeurs aléatoires des biais. Une détermination floue des biais devrait régler ce problème de manière définitive car toutes les valeurs seraient significatives.

Il serait intéressant d'envisager une technique d'apprentissage qui permettrait de réadapter les poids à l'aide d'un système flou.

- La fusion des structures neuronales et du raisonnement flou améliore de manière impressionnante les capacités de génération du système. Néanmoins, lorsque toutes les opérations sont rassemblées en un seul réseau, la structure est encombrée et l'apprentissage devient extrêmement lent. Le partage du processus en un ensemble de tâches élémentaires réalisées par des sous-réseaux de taille réduite, permet de palier à ce problème et rend l'apprentissage beaucoup plus rapide tout en maintenant les capacités de génération.

Conclusion générale

CONCLUSION GENERALE

Un bon diagnostic des huiles de transformateurs est primordial, pour assurer la sécurité des biens et des personnes et la continuité de service. Il ne peut être fait que par des personnes compétentes. Le recours aux systèmes intelligents permet de capitaliser les connaissances et l'expérience humaine.

La logique floue et les réseaux de neurones artificiels sont de bons outils d'aide à la décision mais chacun présente des lacunes qui peuvent fausser les résultats. Leur fiabilité n'est pas absolue.

La combinaison de ces deux approches permet de surmonter les difficultés présentées par chacune des deux techniques et de rassembler leurs avantages. Le résultat est un système expert qui utilise le raisonnement humain tout en étant capable d'apprendre.

Il y a plusieurs manières de réaliser ces combinaisons. La première consiste à introduire la logique floue dans le neurone lui même. La génération de ces réseaux n'est correcte que si l'architecture du réseau reproduit d'une manière ou d'une autre le raisonnement flou.

La seconde approche consiste à fusionner totalement les deux techniques, soit en un seul réseau compact, soit en un ensemble de réseaux dissociés élémentaires. Dans les deux cas la génération est très bonne, cependant le temps de calcul est très grand pour la structure compacte.

La dernière possibilité consiste à introduire localement, l'une des deux techniques pour aider l'autre. L'initialisation floue des poids donne des résultats satisfaisant mais insuffisants à cause des valeurs aléatoires des biais.

Parmi tous les systèmes étudiés, le système fusionné dissocié est l'hybridation neuro-flou la plus performante pour le diagnostic des huiles de transformateurs. Sa robustesse, sa rapidité et la précision de ses sorties lui permettent de donner des décisions correctes et d'éviter les cas d'indécisions.

Les résultats obtenus lors de notre travail encouragent la poursuite des recherches dans ce sens. Les systèmes adoptés doivent être éprouvés et leurs performances exactes doivent être déterminées. Nous pouvons également envisager de les améliorer.

Plusieurs structures hybrides restent à tester. Nous citons : l'adaptation floue des poids, la détermination floue des biais, la détermination des règles d'inférence par un réseau neuronal (Tagaki et Hayashi) et la détermination des sous-ensembles flous par un réseau neuronal (Adeli et Hung).

Il serait également intéressant de voir l'effet de l'utilisation des autres algorithmes d'apprentissage qu'ils soient supervisés ou non.

Les algorithmes génétiques peuvent être utilisées pour régler n'importe quel paramètre (les paramètres d'apprentissage ou les biais par exemple).

Bibliographie

BIBLIOGRAPHIE

- [1] L.S. : 'Fraude et Factures Impayées à SONELGAZ : Cinq Milliard de Dinars de Pertes', El Wattan, dimanche 29 mai 2005, p.1et 5.
- [2] Nathalie Rolland, Philippe Magnier : 'Explosions et Incendies de Transformateurs : Méthode d'Evaluation des Coûts', 2002 :
www.sergi_France.com/html/documents/FTPoa03f.pdf
- [3] Michel Weinfeld : 'Réseaux de Neurones', Techniques de l'ingénieur, article H 1990, 2001.
- [4] Leïla Mokhnache : 'Application des Réseaux de Neurones dans le Diagnostique et la Prédiction des Isolations de Haute Tension'. Thèse de Doctorat d'Etat, ENP, Décembre 2003.
- [5] Eric Davalo et Patrick Naïm : 'Des Réseaux de Neurones', Edition Eyrolles, 1990.
- [6] Kary Främling : 'Les Réseaux de Neurones Comme Outil d'Aide à la Décision Floue', Rapport de D.E.A (Ingénierie Informatique), INSA Lyon , Juillet 1992.
- [7] Ahmad M. Ibrahim : ' Fuzzy Logic for Embedded Systems Applications', Newnes, 2004.
- [8] Claude Touzet : 'Les Réseaux de Neurones Artificiels, Introduction au Connexionnisme : cours, exercices et travaux pratiques', EC2 Edition, Paris, 1992.
- [9] M'hamed Yaddou : 'Etude de Synthèse sur les Réseaux de Neurones et leurs Applications', Thèse de Magistère, ENP, 1998.
- [10] L. Mokhnache, A. Boubakeur : 'Network Prediction of the Breakdown Voltage in a Point-Barrier-Plane Air Gap Using Neurol', CEIDP 2001, IEEE-annual report conference on Electric Insulation and Dielectric Phenomena, Ontario-Canada, Octobre 2001, pp.369-372.
- [11] L. Mokhnache, A. Boubakeur : 'Self-Organisation Neural Network In Prediction of the Breakdown Voltage in a Point Barrier-Plan Air Gaps Using Different Barrier Widths', CGC, IEEE CD-ROOM, paper ID : 8, Bahrein, Mai 2003.
- [12] L. Mokhnache, A. Boubakeur, N. Nait Said : 'Influencing Training Parameters in Some Supervised Neural Networks In Prediction of the Breakdown Voltage in a Point-Barrier-Plane Air Gap', Trans of 4th Int. Conf. on Neural Networks and application of World Scientific And Engineering Academy And society (WSEAS), ID : 454-187, Athena-Greece, Juin 2003.
- [13] L. Mokhnache, A. Boubakeur : ' The Use of Some Back-Propagation Paradigms in the Classification of the Insulating Transformer Oil', 4th QRM. Professional Publishers, IMecE, Oxford, UK, Avril 2002, pp.191-194.

- [14] L. Mokhnache, A. Boubakeur : ' Classification Of Transformer Oil Using Self-Organizing Network And Bayesian Neural Networks', Int.Conf. on Energy efficiency ICEE, 2003.
- [15] L. Mokhnache, A. Boubakeur, B. Oould Noureddine, M. A. R. Bedja : 'Prediction Of Long Thermal Ageing Of Transformer Oil Using Neural Network', Proceeding of international Conference on advance in Processing Testing and Application of Dielectric Materials APTADM'2001, N° Spécial Przegląd Elektrotechniczny, Wroclaw, Poland, Septembre 2001, pp.277-280
- [16] L. Mokhnache, A. Boubakeur, N. Ait Said : 'Application of Neural Networks Paradigms in the Diagnosis and Thermal ageing Prediction of Transformer Oil', Int.Conf. on Dielectric Liquid –ICDL'2002 proceeding, IEEE-DEIS, Paper N°048, Graz, Austria, Juillet 2002, pp.258-261
- [17] L. Mokhnache, A. Boubakeur : 'Comparison of Different Neural Network Algorithms Used in the diagnosis of Transformer Oil', IEEE-DEIS proceeding CEIDP'2002, Cancun, Quintana Roo, Mexique, Oct.2002, pp.244-247.
- [18] L. Mokhnache, A. Boubakeur, N. Nait Said : 'Influencing Training Parameters in Some Supervised Neural Networks for Transformer Oil Diagnosis', Int. Conf. on Condition Monitoring and Diagnostic Engineering Management (COMADEM), Växjö, Swiss, Août 2003.
- [19] L. Mokhnache, A. Boubakeur, N. Nait Said, A. Felliachi, R. Kattan : 'Supervised and Unsupervised Neural Networks Used in the Slassification and Diagnosis of Transformer Oil', soumis à IEEE-Energy Conversion Transaction, 2003.
- [20] L. Mokhnache, A. Boubakeur, A. L. Nemmour, A. Khelifi : 'Neural Network Application in the Field of HV Cable Insulation Thermal Ageing ', 3rd Int. Conf. on Quality, Reliability and Maintenance (QRM), Professional Publishers, ImecE, Oxford, UK, Avril 2000, pp.122-126.
- [21] L. Mokhnache, A. Boubakeur : 'Self-Organisation Neural Networks in Prediction of Insulation Thermal Ageing', ISH2003, IEEE-DEIS proceeding, Delft-Netherlands, Aout 2003.
- [22] L. Mokhnache, A. Boubakeur, A. Felliachi : 'Thermal Ageing Prediction of Transformer Oil and PVC of High Voltage Cables Using Neural Networks', IEE Proceedings– Science, Measurement and Technology, Vol. 150, Issue 03, Mai 2003, pp.107-112.
- [23] L. Mokhnache, A. Boubakeur : 'Use of Neural Networks in the Monitoring of High Voltage Insulation Thermal Ageing', acceptation finale pour publication dans Journal of COMADEM, Octobre 2003.
- [24] L. Mokhnache, A. Boubakeur : 'Application of Many Neural Networks Paradigms in the High Voltage Insulation', Archive of Electrical Engineering (AEE), Polish Academy of Science, Vol. 3, Octobre 2003.

- [25] Jelena Godjevac : ‘Idées Nettes sur la Logique Floue’, Presses Polytechniques et Universitaires Romandes, Lausanne, 1999.
- [26] Pierre Gabriel : ‘Introduction à la Logique Floue et à la Commande Floue’, 2001 : http://elap.montefiore.ulg.ac.be/fuzzy/Figure1/la_logique_floue.htm
- [27] L.A. Zadeh, ‘Fuzzy Sets’, Information and Control, Vol 8, 1965, pp.29-44.
- [28] E. Mamdani : ‘An Expériment in Linguistic Synthesis with a Fuzzy Logic Controller’, International Journal on Man-Machine Studies, Vol.7, 1975, pp.1-13.
- [29] François Chevré, François Guély : ‘La Logique Floue’, Cahier Technique n°191, Collection Technique du Groupe Schneider, mars 1998.
- [30] Yacine Chakhchoukh : ‘Modélisation et Commande par Logique Floue d’une Suspension Active d’un Véhicule’, Thèse de PFE, ENP, Juin 2004.
- [31] Fuzzy Logic Toolbox For use with MATLAB user’s guide version 2, The MathWorks, 1999.
- [32] Hansruedi Bühler : ‘Réglage par logique floue’, Presses Polytechniques et Universitaires Romandes, Lausanne, 1994.
- [33] L.A. Zadeh : ‘A Rational for Fuzzy Control’, Journal of Dynamic Systems, Measurement and Control, Mars 1972, pp. 3-4.
- [34] L.A. Zadeh : ‘Outline of a New Approach to the Analysis of Complex Systems and Decision Processes’, IEEE Trans, on systems, Man and Cybernetics, 3(1), Janvier 1973, pp.28-44.
- [35] T. Tagaki et M. Sugeno : ‘Derivation of Fuzzy Control Rules From Human Operator’s Control Actions’, Proc of the IFAC Symp. On Fuzzy Information, Knowledge Representation and Decision analysis, Juillet 1983, pp.55-60.
- [36] Ngee Ann : the polytechnic of industry: School of engineers : http://www.np.edu.sg/~yck/CCA_L6.pdf
- [37] Honglin Zhang : ‘Fuzzy sets and Fuzzy Neurons’, 23 Mars 2004 : www.engr.usask.ca/classes/ME/885/notes/ME885Slides.pdf
- [38] Eiji Uchinot and Takeshi Yamakawat,” Neo-Fuzzy-Neuron Based New Approach to System Modeling,with Application to Actual System”, IEEE, 1994, pp.564-565.
- [39] H. Nomura, I. Hayashi, N. Wakami, ‘A Self Method of Fuzzy Inference Rules by Descent Method’, Fuzzy logic state of art, R.Lowen and M.Roubens. Klumer Academic Puplishers, 1993, pp.465-470.
- [40] Abraham. Ajith : ‘ Neuro Fuzzy Systems: State-of-the-art Modeling Techniques’, School of Computing & Information Technology, Monash University, Churchill 3842, Australia : <http://ajith.softcomputing.net>

- [41] F. Hoffmann, B. Baesens, J. Martens, F. Put, J. Vanthienen : 'Comparing a Genetic Fuzzy and a Neuro-Fuzzy Classifier for Credit Scoring', Royal Institute of Technology, Center for Autonomous Systems, Stockholm, Sweden, International Journal of Intelligent Systems (15), 2000, pp.785–800.
- [42] Chr. W. Frey, M. Sajidman, H.B. Kuntze : 'A Neuro-Fuzzy Supervisory Control System For Industrial Batch Processes', VDI/VDE-GMA-Tagung, Computational Intelligence, Berlin, 3./4, Mars 1998.
- [43] Ludmila I. Kuncheva, Sushmita Mitra : 'A Two-Level Classification Scheme Trained by a Fuzzy Neural Network', Proc. 12 International Conference on Pattern Recognition, Jerusalem, Israel, 1994, pp.467-469.
- [44] H.O. Nyingesa, P.L. Rosin : 'Neural-Fuzzy Applications in Computer Vision', Journal of Intelligent and Robotic Systems (29) : Kluwer Academic Publishers. Netherlands, 2000, pp.309-315.
- [45] Faisel J. Uppal, Ron J. Patton, Vasile Palade : 'Neuro-Fuzzy Based Fault Diagnosis Applied to an Electro-Pneumatic Valve', IFAC, 2002.
- [46] P.J. Vuarchex : 'Caractéristique des Liquides Diélectriques', Technique de l'ingénieur, K714 (K4), Janvier 1988.
- [47] Djida Hamitouche : 'GAMM3 Analyse des Huiles Isolantes', ABB, 1996.
- [48] M. Lombard : 'Surveillance des transfos : L'isolation d'un Transformateur' : www.lgmtransfo.com
- [49] S. Bouazabia : 'Les Huiles Isolantes de Transformateur', Séminaire sur les huiles isolantes de transformateur, ENP le 1-8 Juin 1999.
- [50] S. Abdi : 'Influence du Vieillissement Thermique sur les Propriétés de l'Huile de Transformateur', Thèse de Magistère, ENP, 2000.

Annexes

Annexe 1 : La base de données utilisée pour l'entraînement des réseaux neuronaux.

Exemple n°	Indice de couleur	Viscosité cinématique (Cts)	Indice d'acidité (mgKOH/g)	Rigidité diélectrique (kV/cm)	Facteur de dissipation	Teneur en eau (ppm)	Décision
1	2	11,07	0.02	55	0.038	9	Garder
2	4	11,79	0.04	69	0.032	11	Garder
3	2.5	11,09	0.44	69	0.03	7	Garder
4	2.5	11,8	0.04	58	0.03	17	Garder
5	3.5	10,98	0.09	68	0.042	16	Régénérer
6	4	10,87	0.05	61	0.031	13	Garder
7	3.5	10,49	0.13	60	0.061	17	Régénérer
8	2	10,59	0.01	66	0.016	9	Garder
9	3.5	10,94	0.1	61	0.053	14	Régénérer
10	2.5	10,67	0.06	51	0.064	14	Garder
11	2	11,69	0.01	54	0.07	12	Garder
12	4	10,67	0.05	69	0.049	19	Garder
13	3.5	9,67	0.09	52	0.054	10	Régénérer
14	1	11,69	0.02	58	0.025	8	Garder
15	4	10,98	0.08	48	0.056	25	Filtrer
16	4.5	11,23	0.14	54	0.028	29	Filtrer
17	4	11,22	0.09	59	0.065	23	Régénérer
18	4.5	10,68	0.17	51	0.097	18	Jeter
19	2.5	10,86	0.09	24	0.019	42	Filtrer
20	2.5	10,45	0.04	58	0.019	11	Garder
21	4	9,98	0.05	59	0.036	14	Garder
22	3	9,87	0.03	57	0.024	9	Garder
23	1.5	10,09	0.01	60	0.025	12	Garder
24	2	10,21	0.03	48	0.035	8	Filtrer
25	2.5	10,13	0.02	60	0.032	12.5	Garder
26	5.5	9,8	0.16	50	0.086	18	Régénérer
27	3	11,59	0.08	55	0.049	28	Garder
28	3	10,21	0.02	52	0.047	12	Garder
29	3.5	11,89	0.04	53	0.053	23	Garder
30	2.5	11,13	0.13	46	0.061	29	Filtrer
31	2.5	11,65	0.01	60	0.02	13	Garder
32	3.5	9,98	0.03	59	0.038	14	Garder
33	3	10,76	0.04	60	0.042	29	Garder
34	3.5	11,87	0.09	57	0.045	19	Régénérer
35	2	11,01	0.02	68	0.016	8	Garder
36	1.5	10,12	0.02	46	0.052	19	Filtrer
37	1	10,87	0.02	59	0.02	11	Garder
38	3	10,15	0.04	30	0.042	29	Filtrer
39	3.5	10,02	0.09	57	0.045	19	Régénérer
40	2	10,58	0.02	68	0.016	8	Garder

Exemple n°	Indice de couleur	Viscosité cinématique (Cts)	Indice d'acidité (mgKOH/g)	Rigidité diélectrique (kV/cm)	Facteur de dissipation	Teneur en eau (ppm)	Décision
41	1.5	10,98	0.02	46	0.052	19	Filtrer
42	1	8,137	0.02	59	0.02	11	Garder
43	4	9,33	0.03	68	0.042	19	Garder
44	1.5	9,77	0.01	67	0.063	7	Garder
45	5	10,12	0.14	11	0.005	74	Régénérer
46	1.5	8,65	0.01	67	0.006	6	Garder
47	3.5	10,14	0.1	56	0.073	15	Régénérer
48	4	10,23	0.06	32	0.063	26	Filtrer
49	3	11,01	0.02	63	0.018	9	Garder
50	2.5	11,69	0.05	55	0.036	25	Garder
51	5	11,79	0.05	70	0.016	11	Garder
52	4.5	11,55	0.13	28	0.078	46	Jeter
53	1	11,62	0.01	43	0.034	18	Filtrer
54	4.5	11,43	0.08	10	0.066	43	Jeter
55	3.5	11,52	0.06	59	0.105	22	Régénérer
56	0.5	11,28	0.01	58	0.007	17	Garder
57	5	11,19	0.22	54	0.067	17	Régénérer
58	5	10,23	0.04	64	0.04	10	Garder
59	4	11,51	0.03	66	0.049	20	Filtrer
60	3	9,84	0.04	63	0.042	9	Garder
61	0.5	9,84	0.01	62	0.06	6	Garder
62	4	9,33	0.08	63	0.083	12	Garder
63	4	10,53	0.21	53	0.106	27	Jeter
64	3	10,67	0.03	70	0.047	10	Garder
65	6	9,52	0.08	55	0.043	21	Garder
66	2	10,35	0.02	48	0.033	21	Filtrer
67	2	11,41	0.02	59	0.019	10	Garder
68	2	10,98	0.04	51	0.017	8	Garder
69	3.5	10,81	0.06	45	0.032	24	Filtrer
70	3	10,98	0.06	54	0.047	12	Garder
71	4	11,36	0.13	65	0.078	19	Régénérer
72	4	11	0.4	30	0.5	40	Réformer
73	1	12	0.04	60	0.04	33	Garder
74	3	10	0.08	47	0.02	35	Filtrer
75	2	12.5	0.02	56	0.02	32	Garder
76	3	11	0.1	42	0.1	37	Filtrer
77	1	10	0.046	56	0.015	30	Garder
78	2	11.45	0.11	50	0.025	35	Filtrer
79	2	12.3	0.072	58	0.02	33	Garder
80	2	12	0.092	48	0.022	37	Régénérer

Exemple n°	Indice de couleur	Viscosité cinématique (Cts)	Indice d'acidité (mgKOH/g)	Rigidité diélectrique (kV/cm)	Facteur de dissipation	Teneur en eau (ppm)	Décision
81	4	12	0.3	30	0.6	41	Réformer
82	0.7	8.960	0.02	22	0.01515	59.5	Filtrer
83	1.9	11.75	0.03	27	0.024	41.9	Régénérer
84	<0.5	9.06	0.02	<15	0.0026	36.6	Filtrer
85	5	11.32	0.078	-	0.1274	-	Régénérer
86	1.6	11.12	0.042	-	0.1144	-	Régénérer
87	6.5	12.024	0.04	22	0.01245	31.5	Réformer
88	1	12.54	0.046	56	0.015	31	Garder
89	3	11	0.106	50	0.025	35	Régénérer
90	2	12.3	0.072	52	0.02	32	Garder
91	4	11	0.09	42	0.082	37	Filtrer
92	4	10	0.3	30	0.6	42	Réformer
93	5,4	10,64	0,4	38	0,1207	35,1	Jeter

Annexe 2 :

Les règles utilisées pour le système flou.

Règle n°	Indice de couleur	Viscosité cinématique	Indice d'acidité	Rigidité diélectrique	Facteur de dissipation	Teneur en eau	Décision
1	Transparente	–	Bonne	Bonne	Bonne	Bonne	Garder
2	Transparente	–	Bonne	Bonne	Moyenne	Bonne	Garder
3	Transparente	–	Bonne	Moyenne	Bonne	Bonne	Filtrer
4	Transparente	–	Bonne	Moyenne	Moyenne	Bonne	Filtrer
5	Transparente	Moyenne	Bonne	Mauvaise	Bonne	Mauvaise	Filtrer
6	Transparente	Moyenne	Moyenne	Bonne	Bonne	Bonne	Garder
7	Transparente	Moyenne	Moyenne	Bonne	Bonne	Moyenne	Garder
8	Transparente	Elevée	Moyenne	Bonne	Bonne	Moyenne	Garder
9	Jaune	–	Bonne	Bonne	Moyenne	Bonne	Garder
10	Jaune	–	Bonne	Moyenne	Bonne	Bonne	Filtrer
11	Jaune	–	Moyenne	Bonne	Bonne	Bonne	Garder
12	Jaune	–	Moyenne	Bonne	Moyenne	Bonne	Garder
13	Jaune	–	Moyenne	Moyenne	Bonne	Bonne	Filtrer
14	Jaune	–	Mauvaise	Bonne	Bonne	Bonne	Garder
15	Jaune	–	Mauvaise	Moyenne	Moyenne	Bonne	Filtrer
16	Jaune	–	Mauvaise	Mauvaise	Bonne	Mauvaise	Filtrer
17	Jaune	Moyenne	Bonne	Bonne	Bonne	Moyenne	Garder
18	Jaune	Moyenne	Mauvaise	Bonne	Bonne	Moyenne	Filtrer
19	Jaune	Elevée	Moyenne	Bonne	Bonne	Moyenne	Garder
20	Orange	–	Bonne	Bonne	Moyenne	Bonne	Garder
21	Orange	–	Moyenne	Bonne	Bonne	Bonne	Garder
22	Orange	–	Moyenne	Bonne	Moyenne	Bonne	Garder
23	Orange	–	Moyenne	Bonne	Mauvaise	Bonne	Régénérer
24	Orange	–	Moyenne	Moyenne	Bonne	Bonne	Filtrer
25	Orange	–	Moyenne	Mauvaise	Moyenne	Bonne	Filtrer
26	Orange	–	Mauvaise	Bonne	Moyenne	Bonne	Régénérer
27	Orange	Moyenne	Moyenne	Moyenne	Bonne	Moyenne	Filtrer
28	Orange	Moyenne	Mauvaise	Bonne	Bonne	Moyenne	Filtrer
29	Orange	Moyenne	Mauvaise	Moyenne	Mauvaise	Mauvaise	Régénérer
30	Brune	–	Moyenne	Bonne	Bonne	Bonne	Garder
31	Brune	–	Moyenne	Bonne	Moyenne	Bonne	Garder
32	Brune	–	Moyenne	Moyenne	Moyenne	Bonne	Filtrer
33	Brune	–	Moyenne	Mauvaise	Moyenne	Bonne	Filtrer
34	Brune	–	Moyenne	Mauvaise	Moyenne	Mauvaise	Régénérer
35	Brune	–	Mauvaise	Bonne	Bonne	Bonne	Filtrer
36	Brune	–	Mauvaise	Bonne	Moyenne	Bonne	Régénérer
37	Brune	–	Mauvaise	Mauvaise	Bonne	Mauvaise	Régénérer
38	Brune	–	Mauvaise	Mauvaise	Moyenne	Mauvaise	Jeter
39	Brune	Faible	Mauvaise	–	Mauvaise	–	Jeter
40	Brune	Moyenne	Moyenne	Mauvaise	Bonne	Moyenne	Filtrer
41	Brune	Moyenne	Moyenne	Mauvaise	Moyenne	Moyenne	Régénérer
42	Brune	Moyenne	Moyenne	Mauvaise	Mauvaise	Moyenne	Jeter
43	Brune	Moyenne	Mauvaise	Moyenne	Moyenne	Mauvaise	Jeter

Résumé :

La maintenance des éléments de production permet d'assurer la continuité de leur fonctionnement. Différentes techniques ont été appliquées dans des systèmes d'aides à la décision et ont donné des résultats satisfaisants. L'utilisation des méthodes intelligentes a montré leurs capacités de diagnostic et de prise de décision.

Les réseaux de neurones artificiels (RNA) sont caractérisés par leur capacité d'apprentissage, alors que les systèmes flous sont simples et proches du raisonnement humain. Ce travail s'intéresse à l'hybridation de ces deux approches pour le diagnostic des huiles de transformateurs utilisées par la SONELGAZ (Borak 22) à partir de tests physico-chimiques.

Plusieurs possibilités sont envisageables pour la combinaison des réseaux neuronaux et des systèmes flous. Deux grandes catégories existent : les réseaux de neurones flous et les systèmes neuro-flous.

Les méthodes de diagnostic proposées conduisent à de meilleurs résultats que ceux offerts par chacune techniques.

Mots clés : Diagnostic des huiles de transformateurs, Réseaux de neurones artificiels, logique floue, Systèmes hybrides neuro-flous.

Abstract :

The monitoring of production elements permit to ensure their permanence's utility. Many techniques were applied in expert systems and have given good results. The use of intelligent methods proved their diagnosis efficiency and their ability to give the appropriate decision.

The artificial neural networks (ANN) are characterized by their ability to learn from the database whereas the fuzzy systems are simple and approach the human reasoning. This work focus the hybridization of artificial neural networks (ANN) and the fuzzy inference systems (FIS) in the diagnosis of transformer oil used by the SONELGAZ (Borak 22) using physico-chemical tests.

Many possibilities are imaginable for the combination of neural networks and fuzzy systems. Two large categories exist : fuzzy-neuron networks and the neuro-fuzzy systems.

Proposed diagnosis methods lead to better results than those offered by each technique.

Keywords : Transformer oil diagnosis, Artificial neural networks, Fuzzy logic, Neuro-Fuzzy hybrid systems.

المخلص :

إن صيانة وسائل الإنتاج تسمح بضمان استمرارية تشغيلها. استعملت عدة طرق لإنشاء أنظمة مساعدة في التقرير أعطت نتائج مرضية. إستعمال الطرق الذكية أظهر قدراتها على المعاينة و إتخاذ القرارات. تتميز الشبكات العصبونية الإصطناعية بقدرتها على الأستعاب انطلاقا من بنك للمعلومات بينما تنفرد الأنظمة المعتمدة على المنطق الغامض ببساطتها و قربها من طريقة التفكير الأنساني. يهدف هذا العمل الى هجن الشبكات العصبونية الأصطناعية والمنطق الغامض في معاينة زيوت المحوّلات المستعملة من طرف سونالغاز (بوراك 22) اعتمادا على التحاليل الفيزيو-كيميائية. لادماج الأليتين هناك عدة امكانيات. تنقسم الأنظمة الهجينة الى صنفان: شبكات ذات عصبونات غامضة و أنظمة عصبو-غامضة.

الأنظمة المقترحة أعطت نتائج أكثر صّحة من القرارات المستخلصة من الآ استعمال الاحادي لكل تقنيّة. **الكلمات المفتاحية :** معاينة زيوت المحوّلات، الشبكات العصبونية الأصطناعية، المنطق الغامض، الأنظمة الهجينة العصبو-غامضة.