



Développement d'une méthode connexionniste pour la détection et le diagnostic de défauts de systèmes de chauffage

Xiaoming Li

► To cite this version:

Xiaoming Li. Développement d'une méthode connexionniste pour la détection et le diagnostic de défauts de systèmes de chauffage. Interface homme-machine [cs.HC]. Ecole Nationale des Ponts et Chaussées, 1996. Français. ⟨NNT : ⟩. ⟨tel-00529470⟩

HAL Id: tel-00529470

<https://pastel.hal.science/tel-00529470v1>

Submitted on 25 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Thèse de Doctorat de l'Ecole Nationale des Ponts et Chaussées

Spécialité

Sciences et techniques du bâtiment

Présentée par

LI Xiaoming

pour obtenir le grade de Docteur de l'Ecole Nationale des Ponts et Chaussées

Sujet de la thèse

**Développement d'une méthode connexionniste pour la détection
et le diagnostic de défauts de systèmes de chauffage**

Soutenue le 25 Janvier 1996 devant le jury composé de

Mme. BENARD C.	(Président)
MM. CHEREL D.	
DUBUISSON B.	(Rapporteur)
HUETZ J.	(Rapporteur)
LEFEBVRE G.	
VISIER J.C.	(Directeur de thèse)

NS 19 856 (4)

X

à toute ma famille

32



*« La pensée digne de ce nom ne se
nourrit pas d'abstractions creuses mais
de réflexions sur la réalité »*

REMERCIEMENTS

*Ces travaux ont été effectués au Service ENergie, Environnement intérieur et Automatismes (ENEA) du Centre Scientifique et Technique du Bâtiment (**CSTB**) à Marne-la-Vallée.*

Je voudrais adresser un grand merci à Monsieur Jean-Christophe VISIER, Chef de Division du CSTB et Directeur de thèse, qui a suivi et encadré de près cette thèse et m'a fait bénéficier de son expérience. Que Monsieur Philippe DUCHENE-MARULLAZ, Chef du Service ENEA, trouve aussi l'expression de ma sincère gratitude pour m'avoir accueilli dans son Service et pour sa sympathie.

*J'adresse ma profonde gratitude à Monsieur Bernard DUBUISSON, Professeur à l'Université de Technologie de Compiègne (**UTC**), Directeur du Laboratoire Heuristique et Diagnostic des Systèmes Complexes (URA 817 CNRS et UTC), pour avoir rapporté, avec soin et en toute rigueur, ce mémoire et avoir accepté de participer au jury de cette thèse.*

*J'ai été très sensible à l'honneur que m'a fait Monsieur Jacques. HUETZ, Professeur à l'Ecole Centrale de Paris (**ECP**), ancien Responsable de la formation doctorale en Energétique (UPR 288 CNRS et ECP), pour avoir accepté d'être rapporteur et de participer au jury.*

*J'exprime toute ma reconnaissance à Madame Christine BENARD, Directeur de la Recherche et des Etudes à l'Ecole Normale Supérieure (**ENS**), pour m'avoir fait l'honneur de présider ce jury, et pour sa rigueur et son intérêt vis-à-vis de cette thèse.*

*Je remercie Monsieur Didier CHEREL, Ingénieur à l'Agence de l'Environnement et de la Maîtrise de l'Energie (**ADEME**), ainsi que Monsieur Gilles LEFEBVRE, Directeur de laboratoire de recherche à l'Ecole Nationale des Ponts et Chaussées (**ENPC**), pour avoir accepté de participer au jury.*

Je tiens à remercier Monsieur Hossein VAEZI-NEJAD, Ingénieur au CSTB, pour l'aide qu'il m'a prodiguée lors de l'élaboration d'un simulateur numérique de systèmes de chauffage collectif à eau chaude.

J'adresse également mes remerciements à tous mes collègues du Service ENEA - ingénieurs, techniciens, thésards, stagiaires, etc., pour les conseils et l'assistance que ceux-ci m'ont prodigués, et pour les amitiés et la sympathie exprimées tout au long des années passées ensemble.

Pour terminer, je remercie vivement le secrétariat pour le soin apporté au travail de dactylographie, en particulier Madame Chantal PIZALIS.

RESUME

L'accroissement des performances des systèmes de génie climatique de ces dernières années s'est accompagné d'une complexité accrue de ceux-ci. La compréhension de leur fonctionnement ainsi que la détection et le diagnostic de leurs défauts deviennent de plus en plus difficile pour les équipes de maintenance. Ces derniers souhaitent donc disposer d'outil performant d'aide à la détection des défauts ou dérives de fonctionnement et, éventuellement, d'aide au diagnostic des causes de ceux-ci.

Cette thèse s'intéresse au développement d'un tel outil adapté aux systèmes de chauffage collectif à eau chaude.

Les défauts pour lesquels il paraissait le plus utile de développer un outil de détection et de diagnostic ont d'abord été recensés. La modélisation-simulation du fonctionnement des cinq systèmes de chauffage avec et sans défauts a permis ensuite d'obtenir une base de données destinée au développement d'un outil de détection et de diagnostic.

Un prototype de détection et de diagnostic basé sur la *reconnaissance des formes* a été développé en utilisant, comme outil de classification, un modèle connexionniste (*réseaux de neurones multi-couches*).

Ce prototype a été testé sur les cinq systèmes simulés. Le résultat est satisfaisant avec un taux de réussite supérieur à 90% et un risque de fausse alarme inférieur à 2% pour l'ensemble des défauts des cinq systèmes simulés. Ceci bien que seules les données provenant d'un des systèmes aient été utilisées pour la phase d'apprentissage du prototype.

Cette étude permet de penser que la généralisation du prototype vers des systèmes réels devrait donner des résultats intéressants. A l'issue de ce travail les principales perspectives consistent à :

- valider l'outil de détection et de diagnostic ainsi développé sur des systèmes réels,
- implanter celui-ci dans des systèmes de *gestion technique des bâtiments* en collaborant avec des industriels,
- appliquer l'approche développée dans cette thèse à d'autres systèmes de génie climatique.

MOTS CLES :

Bâtiment - Chaudière à gaz - Contrôle-commande - Détection et diagnostic de défauts - Gestion de l'énergie - Maintenance - Reconnaissance des formes - Réseau de neurones - Système de chauffage à eau chaude

ABSTRACT

Nowadays, the Heating, Ventilating and Air Conditioning (HVAC) systems are becoming more and more complicated due to the increase of their performances. Consequently it is more difficult for the maintenance teams to understand the running states of these systems and to detect and diagnose their faults. They wish to have performant tools which can help to detect and diagnose (if necessary) the operating faults of HVAC-systems.

This thesis concerns the development of one such tool for Fault Detection and Diagnosis (FDD) suitable for hydraulic heating systems.

The most important faults of hydraulic heating systems, for which a FDD tool must be developed, have been selected at the first stage of the thesis. And then, the modelling-simulation of 5 heating systems with and without faults allowed us to obtain a data base which will be used to develop a FDD tool.

Based on *pattern recognition* theory, a FDD prototype has been developed by using a connectionist model (Multi-Layers *neural networks*) as classifier.

The prototype has been tested for 7 cases in each of the 5 simulated systems. The overall results are satisfactory (a successful detection rate greater than 90% and a false alarm probability less than 2%) although only the simulation data of one heating system were used to train the FDD prototype.

This study shows that the generalization of the FDD prototype to real heating systems could give interesting results. So, the main prospect, resulting from this thesis, consists of:

- validating this prototype with real heating systems,
- implementing it in *Building Energy Management systems* in cooperation with industrialists,
- applying the methodology developed in this thesis to other HVAC-systems.

KEY WORDS

Building - Gas fired boiler - Optimal control - Fault detection and diagnosis - Energy management - Maintenance - Pattern recognition - Neural network - Hydraulic heating system

SOMMAIRE

INTRODUCTION	3
CHAPITRE 1 - Recensement et classification des défauts	10
1.1 - Introduction	11
1.2 - Description d'un système de référence	12
1.2.1 - Description générale du système	12
1.2.2 - Description des raccordements de chaudière	12
1.2.3 - Description du système de contrôle-commande	14
1.3 - Recensement et classification des défauts de systèmes de chauffage	16
1.3.1 - Analyse des défauts avec deux méthodes qualitatives	16
1.3.2 - Entretien avec un nombre réduit d'exploitants	16
1.3.3 - Enquête auprès d'un nombre important d'intervenants extérieurs	17
1.3.4 - Conclusion.....	30
CHAPITRE 2 - Modélisation et simulation de défauts	31
2.1 - Introduction	32
2.2 - Modélisation du système de chauffage	33
2.2.1 - Description du système en vue de la simulation des défauts.....	33
2.2.2 - Présentation des modèles existants.....	33
2.2.3 - Modélisation d'une chaudière à gaz	34
2.3 - Construction d'un simulateur du système de chauffage	45
2.3.1 - Présentation des logiciels MATLAB et SIMULINK	45
2.3.2 - Elaboration d'un simulateur du système de chauffage.....	47
2.4 - Simulation du système de chauffage sans et avec défauts.....	49
2.4.1 - Modélisation des défauts à simuler	49
2.4.2 - Simulation du fonctionnement du système comportant des défauts.....	50

CHAPITRE 3 : Méthode connexionniste pour le diagnostic	55
3.1 - Reconnaissance des formes statistique et diagnostic.....	56
3.1.1 - Généralité sur la reconnaissance des formes statistique.....	56
3.1.2 - Diagnostic de défauts automatisé	58
3.1.3 - Concept de rejet.....	60
3.1.4 - Prétraitement des données	62
3.2 - Méthode connexionniste en tant qu'outil pour la reconnaissance des formes	65
3.2.1 - Introduction.....	65
3.2.2 - Définition d'un système connexionniste	65
3.2.3 - Un modèle connexionniste - réseaux perceptron multi-couches	70
CHAPITRE 4 - Développement d'un réseau connexionniste pour le diagnostic de systèmes de chauffage collectif	81
4.1 - Introduction	82
4.2 - Mise au point des structures de réseaux de neurones.....	84
4.2.1 - Prétraitement de données	84
4.2.2 - Mise au point des structures de réseaux de neurones	93
4.3 - Test de la généricité des réseaux développés	123
4.3.1 - Influence d'une modification de la consigne de température intérieure de la période d'occupation	125
4.3.2 - Influence d'un changement du dimensionnement des radiateurs.....	128
4.3.3 - Influence d'un changement de l'inertie du bâtiment	131
4.3.4 - Influence d'une modification de la consigne de température intérieure de la période d'inoccupation.....	134
4.3.5 - Conclusion.....	137
CONCLUSION	139
ANNEXES	143
Annexe A - Description du système de référence de chauffage collectif à eau chaude	144
Annexe B - Recensement et classification de défauts	153
BIBLIOGRAPHIE	188

INTRODUCTION

POSITIONNEMENT DU PROBLEME ET ENJEUX DE LA RECHERCHE

Un processus d'optimisation du bâtiment consiste à maintenir un bâtiment dans un état de fonctionnement optimal, c'est-à-dire à optimiser des critères (performances) globaux, qui ne sont en général pas des fonctions mathématiquement explicites, [Hyvärinen, 1993].

Ce processus doit en principe tenir compte de tous les éléments liés au bâtiment (contraintes environnementales - confort et protection de l'environnement, maintenance énergétique, éclairage, coût personnel, ...) ; il comprend deux phases principales : la conception et l'exploitation (ou fonctionnement du bâtiment).

Ces deux phases sont a priori interdépendantes car :

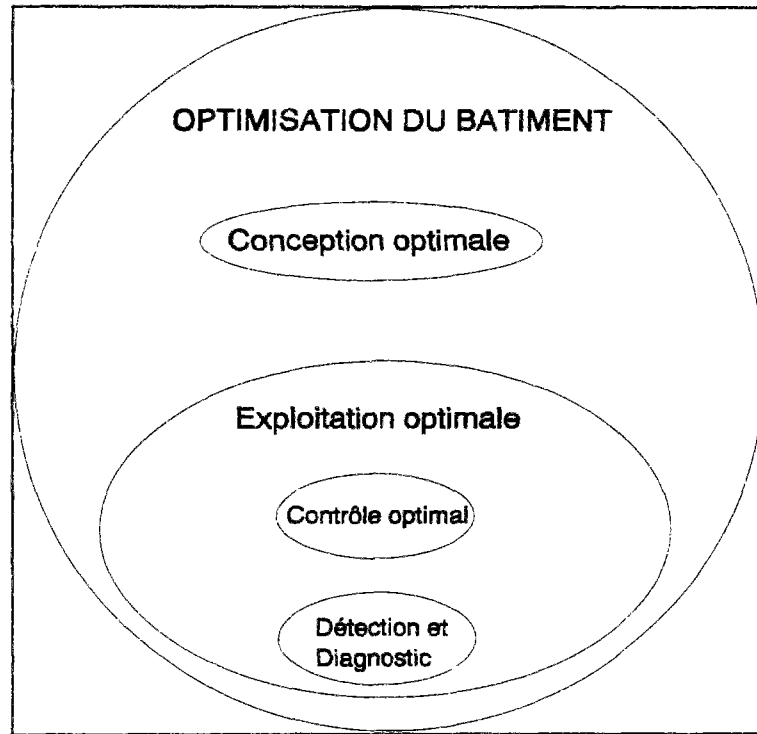
- d'une part la conception optimale (ou optimisation du bâtiment durant la phase conception), réalisée souvent en fonction de simulations du fonctionnement du bâtiment, doit être complétée sur le terrain par un processus d'exploitation optimale afin d'assurer le fonctionnement optimal du bâtiment. Ce dernier processus consiste notamment à optimiser les consignes, les horaires de fonctionnement, etc. et à appliquer le contrôle optimal pour optimiser le fonctionnement du système de génie climatique, etc.,
- d'autre part le retour d'expériences de la phase fonctionnement du bâtiment, grâce à un processus d'optimisation des modes d'exploitation, aide à améliorer la conception de nouveaux bâtiments.

Pour ce qui concerne le contrôle optimal au niveau du bâtiment, deux conditions principales sont nécessaires pour sa réalisation :

- chaque composant (équipement) doit être dans un état de fonctionnement normal,
- chaque sous-système doit se trouver dans un état de fonctionnement optimal.

Il faut donc intégrer, dans un processus d'exploitation optimale, un système de détection et de diagnostic des défauts dont l'objectif est de permettre de s'assurer que chaque composant et sous-système se trouve en permanence dans un état de fonctionnement désiré. Ce système doit en principe permettre :

- de surveiller en temps réel l'ensemble des sous-systèmes et des composants,
- de détecter les défauts et diagnostiquer leurs causes,
- d'évaluer au niveau du bâtiment la gravité de dégradation de performances en présence de défauts,
- de planifier des actions (corrections, reconfigurations, etc.) si nécessaire.



Hiérarchie d'un processus d'OPTIMISATION DU BATIMENT

La figure ci-dessus montre les différentes étapes d'un processus d'optimisation du bâtiment. L'objectif de cette schématisation hiérarchique est de sensibiliser le lecteur au fait que les outils fiables de détection et diagnostic sont des éléments de base indispensables pour résoudre le problème d'optimisation du bâtiment.

Or l'accroissement des performances des systèmes de génie climatique de ces dernières années s'est accompagnée d'une complexité accrue de ces systèmes. La compréhension de leur fonctionnement et la détection et le diagnostic de leurs défauts deviennent de plus en plus difficiles pour un opérateur moyen. Il est donc nécessaire de mettre à la disposition des équipes d'exploitation des outils performants d'aide à la détection des défauts de fonctionnement et des dérives de fonctionnement et, éventuellement, d'aide au diagnostic des causes de ces défauts.

Le développement de systèmes de télégestion et de Gestion Technique du Bâtiment (GTB) fait que l'on dispose de plus en plus fréquemment, dans des installations de taille moyenne ou grande, de systèmes capables de mesurer un grand nombre de variables sur une installation. On constate aujourd'hui que ces systèmes sont massivement utilisés pour automatiser le fonctionnement des installations, pour commander celles-ci à distance ou pour détecter des défauts de fonctionnement francs ou simples. Par contre les constructeurs ne proposent pas à l'heure actuelle sur ces systèmes de fonctions permettant de détecter les dérives de fonctionnement ou de mettre en évidence des défauts plus complexes. Quelques utilisateurs avancés de ces systèmes de télégestion ou de GTB ont développé des fonctions de ce type ; le manque de temps et de moyens n'ont toutefois permis aux plus entreprenants d'entre eux de développer qu'un nombre réduit de fonctions de ce type et ont empêché toute diffusion des fonctions ainsi développées.

L'Agence Internationale de l'Energie (AIE) a mis en place dans le cadre de l'Annexe 25 du programme Conservation d'Energie dans les Bâtiments (CEB), un programme de recherche dont l'objectif est « de développer des méthodes et des procédures permettant l'optimisation en temps réel, la détection et le diagnostic des défauts des installations de génie climatique ». L'Annexe 25 s'est fixé comme objectif de développer des prototypes de méthodes d'optimisation de fonctionnement et de détection diagnostic de défaut qui pourraient être implantés sur des systèmes de GTB. La France, représentée par le CSTB et l'Ecole des Mines de Paris, s'est fortement impliquée dans ces travaux avec le soutien de l'Agence De l'Environnement et de la Maîtrise de l'Energie (ADEME).

La cible de ces différents travaux de recherche est double.

D'une part, les constructeurs de systèmes devraient à terme proposer, dans les logiciels d'application fournis avec leurs produits, des fonctions automatiques d'optimisation de bâtiment, dont de détection et de diagnostic de défaut. L'implication forte des centres de recherche des leaders mondiaux du marché de la régulation dans l'Annexe 25 est le reflet de l'intérêt des constructeurs de Télégestion et de GTB pour ce thème.

D'autre part, les utilisateurs de ces systèmes souhaitent, de plus en plus fréquemment, disposer de systèmes ouverts dans lesquels ils puissent implanter des fonctions non forcément fournies par le constructeur du matériel qu'ils utilisent. Ils sont en ce sens fortement intéressés par une participation à la définition et à la réalisation de ces fonctions.

Le travail de cette thèse s'intéresse en particulier aux systèmes de chauffage collectif à eau chaude. Son objectif est de tenter de développer un outil de détection diagnostic automatique, permettant d'aider les équipes d'exploitation du bâtiment à détecter plus facilement et plus fiablement les défauts d'une part, et à diagnostiquer et analyser leurs causes d'autre part. Cet outil devra être implantable dans un système de télégestion et de GTB.

GENERALITE SUR LA DETECTION ET LE DIAGNOSTIC DE DEFAUTS

Avant de proposer une méthode de détection diagnostic pour les systèmes de génie climatique, on passe en revue les différentes étapes d'une procédure de détection diagnostic afin de donner au lecteur une idée globale sur une stratégie de détection diagnostic.

La surveillance en temps réel du fonctionnement d'un système technologique, quel que soit le domaine concerné, est un enjeu économiquement important. La détermination, sans nécessiter son arrêt, de l'état de fonctionnement d'un système permet une gestion à moindre coût de sa maintenance.

Pour montrer la démarche d'une stratégie de détection diagnostic, on décrit d'abord les concepts de différents états de fonctionnement, ensuite les différentes étapes d'une procédure de détection diagnostic.

Quels sont les états de fonctionnement d'un système ?

Un processus industriel est caractérisé par un certain nombre de variables d'état, de variables externes (sorties), et de sollicitations ou commandes (entrées) ; ces variables sont significatives du fonctionnement du processus et généralement présentées en tout ou en partie, à l'opérateur. Un processus sera dit en état de fonctionnement normal si ces variables sont au voisinage du point de fonctionnement affiché.

Les variables observées sont liées par des relations de cause à effet, relations statiques (température d'équilibre, courbe de consommation, ...), et relations dynamiques (temps de réponse à une perturbation - amortissement, ...).

Nous dirons qu'un processus est en état de défaut, de détérioration ou panne si les relations de cause à effet se sont modifiées. Différents niveaux d'état de défaut peuvent être constatés, et l'on notera par ordre croissant d'importance les défauts en reconnaissant leur caractère, soit purement aléatoire (défauts cataleptiques), soit progressif dans le cas de l'usure ou du vieillissement (défauts évolutifs) : défaut naissant, défaut passager, panne permanente et panne catastrophique.

Les défauts seront également classés selon leur origine. En effet, la modification des relations de cause à effet peut parvenir de l'observation que l'on a du phénomène (capteur) ou encore de la cause que l'on croit appliquée au processus (organe de commande) ou du processus lui-même (fuite, encrassement, rupture d'organe, ...).

Nous classerons donc : défaut de capteur, défaut d'organes de commande, défaut de l'unité de commande et défaut de processus et de composant.

Quelles sont les étapes d'une procédure de détection diagnostic ?

La conduite et la supervision d'un processus impliquent un certain nombre d'opérations nécessaires que nous présenterons ci-dessous.

(1) Acquisition de données

C'est la base même de la procédure et elle doit être totalement fiable.

(2) Perception

L'information sur l'état de fonctionnement du processus est contenue dans les variables mesurées. Cependant, elle est très diluée et le module de perception va jouer le rôle de traitement (compression, etc.) de l'information, recevant les données acquises et fournissant des indices ou paramètres, indicateurs de défauts. Les méthodes réalisant cette fonction peuvent être classées en deux catégories : à base de traitement du signal et à base de modèles, [Brunet, 1990], [Carrier, 1992], [Frank, 1987], [Isermann, 1983], [Wagner, 1992].

(3) Détection

C'est l'opération qui permet de décider si le système est ou n'est pas en état de fonctionnement normal. Les techniques employées sont nombreuses, [Brunet, 1990], [Ragot, 1990], [Vaézi-Néjad, 1990], on peut citer, comme exemple, la logique à seuil, le test d'hypothèse et le test de Page-Hinkley.

(4) Diagnostic

Le diagnostic est une opération de classification qui caractérise le défaut par type et degré de sévérité. Les outils principaux pour effectuer cette opération, [Anderson, 1989], [Brunet, 1990], [Culp, 1989], [Dubuisson, 1990a], [Pau, 1975], [Pau, 1981], sont les suivants :

- la classification, reconnaissance des formes,
- l'analyse discriminante floue
- la logique floue,
- les arbres logiques,
- les systèmes experts et interpréteurs de règles.

(5) Pronostic

Le pronostic qui, sur la base de l'évolution passée d'un défaut, permet de prédire son évolution future dans un contexte similaire. Les outils utilisés sont d'une part, ceux provenant des techniques du traitement de signal (méthodes de lissage, extrapolation, etc.), d'autre part ceux reposant sur une modélisation du processus d'évolution du défaut quand celui-ci peut être identifié.

(6) Analyse de conséquences

Elle indique le temps restant avant la panne ; et de façon plus spécifique les impacts de la panne sur la disponibilité et la sécurité, la qualité, les aspects économiques et environnementaux, etc..

(7) Planification des actions

Prenant en compte le résultat des étapes de diagnostic, pronostic et l'analyse de conséquences, ce module va planifier dans le temps les actions d'urgence, les reconfigurations ou passage en mode dégradé et les évitements de pannes.

Ces sept fonctions ou modules n'apparaissent pas toutes dans chaque application. C'est ainsi que les défauts cataleptiques doivent être détectés aussi rapidement que possible et la détection précoce consiste en la prévention des catastrophes éventuelles et permet la reconfiguration du système. Par exemple, la panne de capteur doit être détectée aussi rapidement que possible pour permettre des actions telles que la reconfiguration de la boucle de commande ou le transfert en mode manuel ou une action de maintenance. En revanche, les défauts évolutifs dont le degré d'appartenance à telle ou telle classe mesure le niveau de sévérité, doivent être surveillés de façon périodique.

METHODE PROPOSEE POUR UNE PROCEDURE DE DETECTION DIAGNOSTIC DE DEFAUTS DE SYSTEMES DE GENIE CLIMATIQUE

En ce qui concerne les systèmes de génie climatique (chauffage, climatisation, ventilation, conditionnement d'air, pompes à chaleur, etc.), une procédure de détection diagnostic de défauts peut s'effectuer de deux manières différentes, [Hyvärinen, 1993] :

- méthode ascendante (bottom-up en anglais), qui consiste à surveiller en temps continu, le système i.e. sous-systèmes et composants, détecter leurs défauts, et évaluer si possible la gravité de la dégradation de performance au niveau du bâtiment,
- méthode descendante (top-down en anglais), qui commence par la détection ou/et le diagnostic d'une ou des performances au niveau du bâtiment ; par exemple, dès qu'une consommation excessive d'énergie a été remarquée, on commence à chercher ses causes (sous-systèmes en fonctionnement non-optimal, composant en état de défaut, ...).

Il est souhaitable que ces deux approches soient conçues et intégrées dans un même système de détection et de diagnostic. En ce qui concerne l'objectif de cette thèse, il s'agit principalement de développer une méthode ascendante. Bien évidemment, la méthode descendante devrait, dans les travaux ultérieurs, être couplée avec une méthode ascendante.

En ce qui concerne la technique de détection diagnostic proprement dite, nous nous attacherons à l'étude d'une méthode connexionniste (réseaux de neurones artificiels) en tant qu'outil de classification (reconnaissance des formes). En effet, la reconnaissance des formes appliquée au problème du diagnostic a fait ses preuves depuis de longues années, [Dubuisson, 1990] et [Masson, 1992]. La dualité entre la notion de classe en reconnaissance des formes, et les états ou modes de fonctionnement en diagnostic, a suscité un nombre de travaux importants dans le domaine de détection diagnostic.

PLAN DE LA THESE

Les travaux présentés dans cette thèse consistent à développer une méthode de détection diagnostic de défauts pour les systèmes de chauffage collectif à eau chaude. Notons que ce type de système est très utilisé et représentatif dans de nombreux pays de la Communauté Européenne.

Le chapitre 1 est consacré d'une part au recensement et à la classification des défauts, d'autre part à la sélection des défauts les plus importants pour lesquels on souhaite disposer de méthodes de détection diagnostic.

Le chapitre 2 présente la modélisation-simulation des défauts d'un système de chauffage à eau chaude afin d'établir une base de données comportant des fonctionnements normaux et défectueux du système. La modélisation-simulation devra permettre de simuler facilement d'autres types de systèmes de chauffage comportant des défauts en changeant simplement un nombre très réduit de paramètres des modèles.

Ces deux chapitres constituent la phase de construction d'une base de données ; celle-ci doit être fiable afin de développer des méthodes de détection diagnostic pouvant répondre aux besoins réels des exploitants des systèmes de chauffage.

Le chapitre 3 rappelle d'abord les différentes étapes de la construction d'un système de détection diagnostic par la méthode de reconnaissance des formes. Il décrit ensuite un modèle connexionniste (réseaux Perceptron Multi-Couches) en tant qu'outil de classification.

Enfin le chapitre 4 présente le développement (apprentissage et test) des réseaux de neurones pour la classification (i.e. diagnostic) des 7 modes de fonctionnement d'un système de chauffage simulé, ainsi que la généralisation des réseaux développés en les appliquant aux 4 autres systèmes de chauffage simulés et non utilisés durant la phase d'apprentissage. Cette généralisation est très importante en terme d'application sur le terrain, car on ne souhaite pas réapprendre sur chaque système (bâtiment) mais trouver une solution pour l'ensemble des systèmes (bâtiments) similaires à diagnostiquer.

CHAPITRE 1

RECENSEMENT ET CLASSIFICATION DES DEFAUTS

Ce chapitre est consacré au recensement et classification des défauts pour les systèmes de chauffage à eau chaude et de production d'eau chaude sanitaire utilisés dans les bâtiments collectifs ou tertiaires. A l'issue de ce travail, les défauts les plus importants vis-à-vis des caractéristiques prédéfinies seront obtenus et serviront de base au développement de méthodes de détection et de diagnostic.

1.1 - Introduction	11
1.2 - Description d'un système de référence	12
1.2.1 - Description générale du système	12
1.2.2 - Description des raccordements de chaudière	12
1.2.3 - Description du système de contrôle-commande	14
1.3 - Recensement et classification des défauts de systèmes de chauffage	16
1.3.1 - Analyse des défauts avec deux méthodes qualitatives	16
1.3.1.1 - Méthode d'Analyse des Dérives de Variable et de leurs Causes	16
1.3.1.2 - Méthode d'Analyse des Modes de Défaillance et de leurs Effets	16
1.3.2 - Entretien avec un nombre réduit d'exploitants	16
1.3.3 - Enquête auprès d'un nombre important d'intervenants extérieurs	17
1.3.3.1 - Introduction	17
1.3.3.2 - Analyse des résultats d'enquête	17
1.3.3.2.1 - Classifications en fonction d'une caractéristique	18
1.3.3.2.2 - Classifications prenant en compte simultanément plusieurs caractéristiques	21
1.3.4 - Conclusion	30

1.1 - INTRODUCTION

Le travail de recensement et de classification des défauts sera entrepris pour les systèmes de chauffage à eau chaude et de production d'ECS (Eau Chaude Sanitaire) utilisés dans les bâtiments collectifs ou tertiaires.

Pour faciliter le recensement et la classification on décrira d'abord un système de référence représentatif. Ce système doit comporter les éléments clés du système de chauffage et de production d'ECS ; mais les options prises dans la description de celui-ci ne doivent pas conduire à trop restreindre le recensement et la classification des défauts..

Après la définition du système de référence, nous effectuerons le recensement et la classification des défauts en suivant trois approches complémentaires :

- application de deux méthodes d'analyse des défauts permettant pour la première (la méthode d'Analyse des Dérives de Variable et de leurs Causes) de rechercher les causes possibles de dérives des variables mesurées et pour la seconde (la méthode d'Analyse des Modes de Défaillance et de leurs Effets) [Villemeur, 1988] d'évaluer les effets des défauts d'un composant sur les variables mesurées,
- entretien avec un nombre réduit d'exploitants de chaufferies collectives pour recenser les défauts des installations et évaluer l'importance des composants de système de chauffage et de production d'eau chaude sanitaire,
- enquête, via un questionnaire, auprès d'un nombre important d'intervenants extérieurs ayant une connaissance du terrain (Membres du Groupement des Utilisateurs des Systèmes de Télégestion, ingénieurs des villes de France, exploitants de chaufferies, bureaux d'études, etc.) afin d'évaluer plus précisément l'importance des défauts et de les classer.

On présente successivement le système de référence et les trois approches suivies. L'accent est mis, dans la présentation, sur la troisième approche. Les principes des deux autres sont décrits dans le texte, les résultats étant présentés en annexe (Annexes B1, B2 et B3).

1.2 - DESCRIPTION D'UN SYSTEME DE REFERENCE

Le système, dit de référence, de chauffage et de production d'ECS pour lequel on cherche à détecter et diagnostiquer les défauts est conçu en faisant référence à un système qui a été décrit initialement par un Bureau d'étude pour Gaz de France. La description du système comprend une description générale, une description des chaudières utilisées et une description du système de contrôle-commande.

1.2.1 - DESCRIPTION GENERALE DU SYSTEME

Les caractéristiques principales du système sont les suivantes (Figure 1.1) :

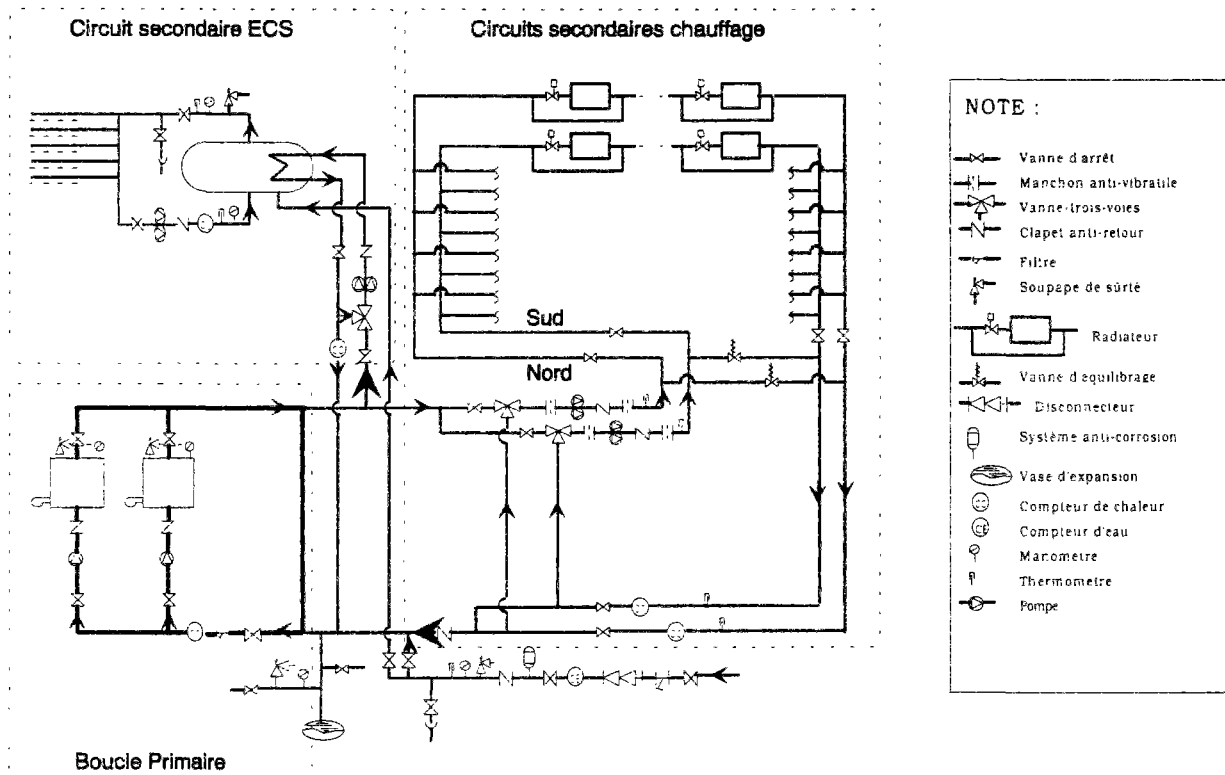


Figure 1.1 Schéma de principe du système de référence

- deux chaudières classiques à gaz sont utilisées pour la génération de chaleur,
- le système de chauffage et de production d'ECS est installé dans un bâtiment de bureau de moyenne dimension,
- le réseau de distribution (la boucle secondaire) comprend trois circuits : un pour le chauffage de la façade nord, un pour celui de la façade sud, et un pour la production d'eau chaude sanitaire,
- les circuits de tous les étages sont connectés parallèlement,
- au niveau de chaque étage on trouve une série de radiateurs par circuit (nord ou sud). Chaque radiateur est raccordé en dérivation sur le circuit monotube de l'étage par l'intermédiaire d'un robinet thermostatique à 4 voies,
- la production d'eau chaude sanitaire est instantanée,
- la ventilation mécanique contrôlée ne fonctionne qu'en période d'occupation,
- une liste de composants du système de référence (voir Annexe A1),
- une liste de variables mesurables du système de référence (voir Annexe A2).

1.2.2 - DESCRIPTION DES RACCORDEMENTS DE CHAUDIERE

Chaque chaudière est connectée avec les autres parties du système par les entrées-sorties suivantes (Figure 1.2) :

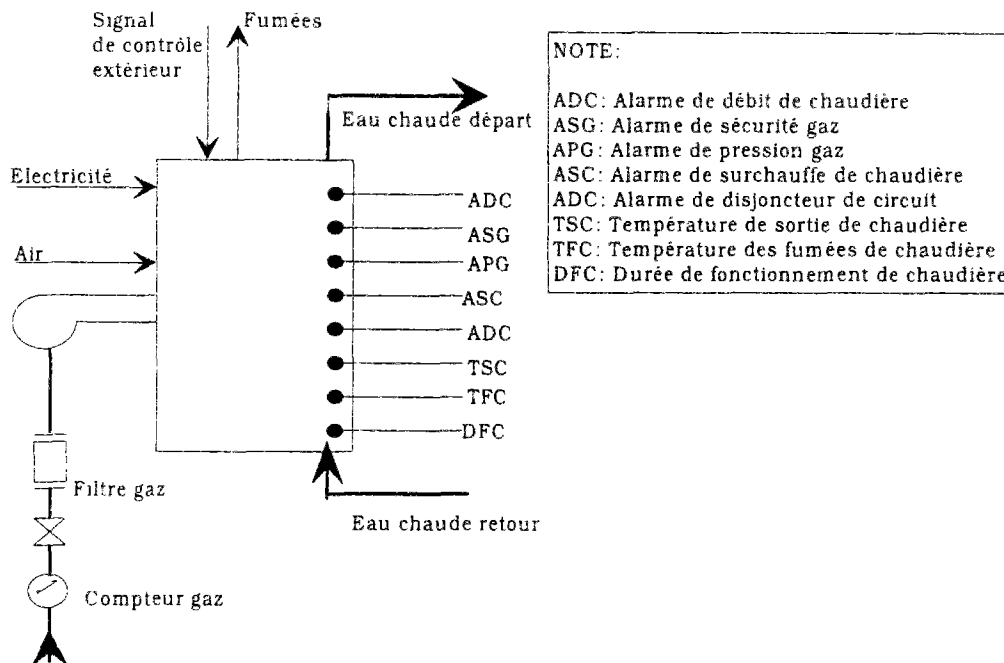


Figure 1.2 Raccordement de la chaudière avec les autres parties du système

Entrées physiques :

- électricité
- air
- gaz naturel
- eau chaude de retour

Sorties physiques :

- fumées
- eau chaude de départ

Entrées logiques et numériques :

- signal extérieur de contrôle
- réarmement du disjoncteur de circuit
- réarmement de l'interrupteur de sécurité
- réarmement du thermostat de surchauffe (thermostat de sécurité)

Sorties logiques et numériques :

- alarme : débit d'eau trop bas dans la chaudière
- alarme : arrêt automatique du brûleur pour des raisons de sécurité
- alarme: pression du gaz naturel trop haute
- alarme: pression du gaz naturel trop basse
- alarme: surchauffe d'eau
- alarme: arrêt du disjoncteur de circuit
- mesure de la température d'eau chaude
- mesure de la température des fumées
- comptage des heures de fonctionnement

- ### 1.2.3 - DESCRIPTION DU SYSTEME DE CONTROLE-COMMANDE

La conduite en séquence (en cascade) des chaudières consiste à déterminer le nombre des chaudières à utiliser et la température de consigne de chaque chaudière.

Ce système de conduite reçoit les informations des programmeurs d'intermittence des deux circuits pour adapter les températures de consigne des deux chaudières aux besoins énergétiques de la boucle secondaire (chauffage et production d'eau chaude sanitaire).

Ces chaudières sont pilotées avec permutation de priorité sur la durée de fonctionnement. Cette permutation a lieu tous les deux jours (i.e. 48 heures) et est assurée par un dispositif automatique sans intervention manuelle.

La régulation en fonction de la température extérieure (régulation en boucle ouverte) consiste à déterminer la température de départ secondaire. Le chauffage comprend deux circuits (nord et sud) afin de tenir compte des ensoleillements différents des deux façades (nord et sud). Le raccordement hydraulique de ces circuits permet de réaliser une indépendance hydraulique de réglage.

Le régulateur calcule une température de consigne en sortie de la vanne-trois-voies en utilisant une courbe de chauffe, et agit en boucle fermée sur le servomoteur de la vanne-trois-voies de chacun des circuits nord et sud pour contrôler (généralement en mode proportionnel et intégral) la température de départ secondaire. En général, cet asservissement fonctionne en tenant compte des trois aspects suivants :

- la consigne de départ (i.e. courbe de chauffe),
- la température de départ mesurée à la sortie de la vanne-trois-voies,
- la température d'eau de retour à chaudière (limite basse exigée par certaines chaudières).

La régulation locale est réalisée par un robinet thermostatique connecté au radiateur. Il régule la température intérieure en boucle fermée, en agissant sur le débit du radiateur. Le débit maximal est obtenu lorsque la température dans le bâtiment est inférieure ou égale à 19°C. Le débit s'annule lorsque cette température atteint 21°C (soit une bande proportionnelle de 2°C). La loi de correspondance entre la température et le débit a été choisie comme non-linéaire.

Cette régulation permet de satisfaire à la fois le confort des occupants et les impératifs de réduction de la consommation d'énergie par le maintien de la température intérieure à un niveau prédéfini et par la récupération des apports gratuits.

Le système de contrôle de production d'ECS est indépendant de celui de chauffage, et agit sur le servomoteur de la vanne-trois-voies pour réaliser le réglage de la température de départ du circuit secondaire. Un aquastat de sécurité est installé à la sortie du circuit secondaire permettant d'interrompre impérativement l'alimentation du circuit primaire par la fermeture de la vanne trois voies en cas de dépassement du seuil haut 60°C par exemple (la température d'ECS est limitée à 60°C pour minimiser les pertes thermiques par distribution).

De plus la température de départ primaire est contrôlée par un seuil bas (60°C par exemple) permettant la fourniture d'ECS même pendant la période d'arrêt ou de maintien du chauffage.

1.3 - RECENSEMENT ET CLASSIFICATION DES DEFAUTS DE SYSTEMES DE CHAUFFAGE

Après avoir défini le système de référence, on présente ci-après les trois approches utilisées pour le recensement et la classification des défauts :

- analyse des défauts avec les méthodes ADVC et AMDE,
- entretien avec un nombre réduit d'exploitants de chaufferies collectives,
- enquête via un questionnaire auprès d'un nombre important d'intervenants extérieurs.

1.3.1 - ANALYSE DES DEFAUTS AVEC DEUX METHODES QUALITATIVES

Les deux méthodes de recensement de défauts utilisées permettent pour la première de rechercher les causes possibles de dérive d'une variable mesurée et pour la seconde d'évaluer les effets des défauts d'un composant sur les variables mesurées.

Ces deux méthodes ont été appliquées de façon systématique pour établir des bases de données décrivant d'une part les symptômes et leurs causes possibles et d'autre part les défauts et les effets qui leurs sont associés.

1.3.1.1 - Méthode ADVC (Analyse des Dérives de Variable et de leurs Causes)

La méthode ADVC est une méthode qualitative qui a été développée dans l'industrie chimique pour identifier des risques, [Ozog, 1985].

Cette méthode permet d'analyser les différentes causes (défauts) des dérives d'une variable mesurée. Pour l'appliquer, nous avons utilisé la liste de variables présentée dans l'Annexe A2.

Nous avons recensé (voir Annexe B1) tous les défauts possibles sans tenir compte de leur importance, c'est-à-dire que les défauts les plus importants et les défauts peu importants sont simultanément présentés. De plus, pour les causes de dérives d'une variable, nous avons séparé les causes provenant directement de défauts d'un composant et celles de dérives d'une autre variable.

1.3.1.2 - Méthode AMDE (Analyse des Modes de Défaillance et de leurs Effets)

Cette méthode est utilisée pour analyser les effets des défauts d'un composant via les variables mesurées, et fournir un système d'audit systématique. Pour appliquer la méthode AMDE, nous avons sélectionné une partie des composants présentés dans l'Annexe A1, cette partie étant considérée comme relativement plus importante par rapport aux autres composants. Le résultat d'analyse en utilisant la méthode AMDE est présenté dans l'Annexe B2.

1.3.2 - ENTRETIEN AVEC UN NOMBRE REDUIT D'EXPLOITANTS

L'application des méthodes ADVC et AMDE nous a permis de recenser en détail les défauts et les dérives de systèmes de chauffage. Or tous les défauts recensés avec ces méthodes ne présentent pas les mêmes importances pour ce qui concerne les quatre aspects suivants : le confort des occupants, les consommations d'énergie, la fréquence de défaut, l'existence ou non de méthodes de détection.

Afin d'évaluer l'importance de ces défauts, nous avons interviewé plusieurs exploitants des chaufferies collectives à partir d'un questionnaire ouvert en utilisant comme support le système de référence décrit précédemment.

D'abord nous avons demandé à ces exploitants quels sont les défauts les plus importants dans les installations dont ils ont la charge vis-à-vis des quatre aspects cités ci-dessus.

Ensuite nous leur avons demandé d'évaluer l'importance de tous les composants listés dans l'Annexe A1 en donnant une note (1 très important, 2 important, 3 moyennement important, 4 peu important) à chaque composant en respectant toujours les quatre aspects.

A l'issue de cette phase nous avons obtenu une liste restreinte des défauts typiques pour ce type d'installation. Cette liste est présentée dans l'Annexe B3. A chaque défaut important, une information particulièrement utile est donnée pour la suite de l'étude, la nécessité qu'il y a à développer des méthodes de détection de ce défaut.

1.3.3 - ENQUETE AUPRES D'UN NOMBRE IMPORTANT D'INTERVENANTS EXTERIEURS

1.3.3.1 - Introduction

Pour confirmer la liste des défauts typiques (voir Annexe B3), et pour avoir une classification des défauts plus représentative et plus solide afin de mieux guider le développement de méthodes de détection et de diagnostic, nous avons lancé une enquête sous forme de questionnaire. Dans ce questionnaire, nous avons listé, grâce aux études précédentes, le plus exhaustivement possible les défauts des composants de systèmes de chauffage et de production d'eau chaude sanitaire. On arrive au total à une liste de 106 défauts (voir Annexe B4).

L'enquête doit permettre pour chacun de ces défauts de déterminer :

- sa fréquence d'apparition,
- s'il est facile aux équipes d'exploitation de détecter ce défaut,
- son incidence sur le confort des usagers,
- son incidence sur les consommations d'énergie.

Pour ce faire on a demandé aux experts de donner pour chaque défaut et pour chaque caractéristique une note allant de 0 à 3 ; les caractéristiques et la signification de ces notes sont présentées dans l'Annexe B4.

1.3.3.2 - Analyse des résultats d'enquête

Le dépouillement des résultats du questionnaire doit nous permettre d'évaluer l'importance des composants d'installations d'une part, et de classer les défauts suivant leur importance d'autre part.

Le questionnaire a été envoyé à 137 personnes, nous avons reçu au total 46 réponses. Les résultats de cette enquête ont été analysés, [Li, 1995 c], et seront présentés comme suit :

- classifications en fonction d'une seule caractéristique,
- classifications prenant en compte simultanément plusieurs caractéristiques.

On a choisi de distinguer les défauts en fonction de la valeur moyenne de la note donnée par les experts. On distingue dans la plupart des cas les défauts ayant pour une caractéristique donnée une note moyenne supérieure ou égale à 1,5 de ceux pour lesquels la note moyenne est inférieure à 1,5. Les premiers sont considérés comme importants au regard de cette caractéristique, les autres étant jugés comme peu importants.

Ces classifications nous permettent de mettre en évidence pour différentes caractéristiques ou groupements de caractéristiques les défauts qui paraissent les plus importants.

1.3.3.2.1 - Classifications en fonction d'une caractéristique

- fréquence d'apparition⁽¹⁾,
- difficulté de détection,
- influence sur le confort,
- influence sur les consommations d'énergie.

(1) Classification en fonction de la fréquence de défaut

Les défauts les plus fréquents ($C1 \geq 1.5$) peuvent être regroupés de la façon suivante :

- défauts du système de production de chaleur,
- défauts de la régulation des circuits secondaires,
- défauts spécifiques à une zone,
- défaut sur l'échangeur ECS,
- défauts divers.

En ce qui concerne la production de chaleur :

- les défauts conduisant à l'arrêt de cette production ont un impact fort sur le confort. Des systèmes de détection particuliers et permettant de les déceler existent déjà dans les installations de chauffage à eau chaude,
- ceux entraînant une moindre efficacité de cette production ne sont pas traités actuellement. Ce sont par exemple : encrassement de l'échangeur principal, mauvaise combustion, mauvaise gestion de la cascade des chaudières.

En ce qui concerne la régulation des circuits secondaires :

- les pannes électriques pour lesquelles il existe déjà des systèmes de détection,
- les défauts liés à un mauvais réglage sont très fréquents et il n'existe pas de procédure automatique pour les détecter. On remarquera que ces défauts sont d'autant plus difficiles à détecter que la régulation des circuits primaires devient complexe. Dans le cas d'une régulation simple en fonction de la température extérieure, la détection est simple ; dès que l'on essaie de faire de l'intermittence, le problème devient plus complexe.

En ce qui concerne les zones particulières (air dans radiateur, usage incorrect de vanne thermostatique, ...) :

- les défauts une fois détectés sont faciles à résoudre. Les défauts conduisant à un inconfort seront probablement rapidement détectés par les plaintes des occupants. En revanche, ceux conduisant à une dérive de consommation sans augmentation de l'inconfort risquent de ne pas être détectés si l'on ne met pas en place de système de détection spécifique (système de détection automatique par exemple). Ces systèmes de détection spécifique n'existent pas aujourd'hui.

En ce qui concerne la production d'ECS :

- seul l'entartrage de l'échangeur est fréquent. Il peut, comme tout défaut qui évolue lentement avec le temps être difficile à détecter. Il pourrait avoir peu d'impact sur le confort et est généralement difficile à détecter, il risque donc de ne pas l'être en l'absence de méthode spécifique.

En ce qui concerne les défauts divers (pannes de pompe, fuite d'eau dans la chaufferie, ...) :

- ils sont généralement faciles à détecter et relèvent d'inspections de routine.

⁽¹⁾ Il apparaît nécessaire en ce qui concerne la fréquence d'apparition des défauts de distinguer les défauts de conception pour lesquels on a cherché à déterminer le taux de présence (pourcentage des installations présentant ce type de défaut) des autres défauts pour lesquels on a cherché à déterminer la fréquence des défauts (durée moyenne d'apparition entre deux défauts).

(2) Classification en fonction du taux de présence pour les défauts de conception

Les défauts de conception ne font a priori pas partie de l'étude qui nous intéresse. Certains d'entre eux peuvent toutefois avoir un effet majeur sur le fonctionnement de l'installation et il paraissait indispensable de les recenser.

Les défauts de conception les plus fréquents ($C2 \geq 1.5$) sont liés :

- au dimensionnement et à l'équilibrage des radiateurs et des circuits. Ce type de défaut est très connu et de nombreuses méthodes permettant de l'identifier et d'y remédier ont été développées. Il reste à les appliquer et à les associer aux méthodes automatiques de détection de défauts ce qui est lourd et pas toujours facile.
- au positionnement des sondes de mesure. Les défauts liés au positionnement des sondes sont bien connus. L'application des règles de l'art pourrait permettre de les limiter. Deux problèmes restent toutefois à résoudre. Comment identifier la présence de ces défauts? Comment développer des méthodes de détection de défauts si on n'est pas sûr de la validité des données fournies par les capteurs?
- au dimensionnement de la vanne trois voies et donc à son autorité. Les défauts liés à l'autorité hydraulique de la vanne trois voies mettent une fois de plus en évidence l'aspect crucial de ces vannes de régulation. On peut se demander si l'on ne met pas sur le compte du dimensionnement de cette vanne des défauts liés également à la commande qui lui est appliquée et donc au système de régulation programmation.
- au dimensionnement des pompes. Les pompes sont toujours surdimensionnées. Il faut tout de même être conscient dès la conception d'une installation de chauffage des problèmes du gaspillage d'énergie afin d'éviter le surdimensionnement exagéré.

(3) Classification en fonction de la difficulté de détection

Les défauts les plus difficiles à détecter ($C3 \geq 1.5$) peuvent être regroupés en trois groupes :

- les défauts de conception. Les personnes en charge de la maintenance de l'installation ont des difficultés importantes à détecter les défauts de conception. Il faut noter que cette détection ne relève pas directement de leur compétence et que de tels défauts auraient normalement dû être détectés lors de la mise en service,
- les défauts à évolution lente. La détection des défauts à évolution lente tels que l'entartrage, l'embouage, la corrosion impose de prendre du recul par rapport au fonctionnement au jour le jour de l'installation. Il est nécessaire pour détecter ce type de défaut de pouvoir archiver à intervalle régulier des "photographies" du fonctionnement de l'installation et de pouvoir les comparer. Ceci impose une organisation qui ait une visée à long terme. La difficulté de détection de ce type de défaut semble montrer que cette démarche visant à prendre du recul est rarement mise en oeuvre.
- les autres défauts.

En ce qui concerne ces autres défauts on trouve principalement :

- un grand nombre de défauts liés à la régulation des circuits secondaires,
- quelques défauts liés à la production de chaleur,
- des défauts liés à des fuites invisibles de l'extérieur de l'installation conduisant à des circulations parasites.

D'abord, la régulation des circuits secondaires apparaît comme un des éléments clés. Les défauts cités comme difficiles à détecter ont trait à la fois à la définition de la consigne de la température de départ par les matériels de régulation programmation et à la vanne de commande (mauvaise autorité ou fuite). Il apparaît très difficile de différencier ces diverses causes qui ont toutes des effets sur la même partie de l'installation et dont les effets peuvent facilement être confondus.

Ensuite les défauts liés à la production de chaleur portent à la fois sur la mauvaise combustion et sur la gestion des chaudières. Il apparaît difficile aux exploitants de déterminer si le nombre de chaudières utilisées est bon et si les réglages des régulations du primaire sont bien effectués.

Enfin les fuites visibles de l'extérieur de l'installation sont rapidement détectées. Toutefois celles qui peuvent se produire entre le primaire et le secondaire des échangeurs de production d'ECS risquent de ne pas l'être.

(4) Classification en fonction de l'influence sur le confort

Les défauts ayant une influence importante sur le confort sont au nombre de 26. Il s'agit presque uniquement de défauts de fonctionnement. Ceci semble signifier que les défauts de conception pouvant avoir une influence importante sur le confort sont détectés au début de la vie de l'installation.

Une grande partie de ces défauts est des défauts francs se traduisant par un arrêt de fonctionnement de tout ou partie de l'installation. On peut citer parmi ceux-ci :

- les pannes électriques ou mécaniques des pompes, des vannes, du brûleur, du système de régulation programmation,
- l'arrêt de l'alimentation du brûleur, la mise en sécurité du brûleur,
- et quelques défauts rares tels que l'éclatement du manchon anti-vibratile, le feu dans le conduit de cheminée ou le gel de canalisations.

Un deuxième groupe est constitué par les défauts du système de régulation programmation : mauvais calculs des heures de relance ou d'arrêt par le programmeur d'intermittence, horloge mal réglée.

Viennent ensuite deux défauts isolés. Le premier est la présence d'air dans les radiateurs, le second est le sous dimensionnement du préparateur d'ECS. Il faut noter que ce deuxième défaut est le seul défaut de conception ayant un impact important sur le confort.

(5) Classification en fonction de l'influence sur les consommations d'énergie

Trente et un défauts recensés ont une incidence forte sur les consommations d'énergie.

9 défauts de conception ou d'installation apparaissent dans cette liste : isolation du bâtiment ou de l'installation non conforme aux calculs, mauvais équilibre thermique ou hydraulique de l'installation, surdimensionnement du brûleur, régulateur pas adapté à la vanne trois voies, sondes de mesures de température intérieure ou extérieure mal positionnées. Ceci semble montrer que les défauts de conception ou d'installation ayant une incidence non négligeable sur les consommations sont beaucoup plus difficilement détectés lors du début de la vie de l'installation que les défauts ayant une incidence sur le confort.

Parmi les défauts de fonctionnement une part importante est constituée par des défauts apparaissant lentement : dégradation de l'isolation, embouage et entartrage ce type de défaut pose par nature des difficultés particulières de détection.

Les autres défauts ayant une incidence forte sur le confort se répartissent comme suit :

- fuites d'eau : vers les locaux ou la chaufferie, au niveau des vannes, au niveau des échangeurs,
- défauts liés aux occupants (ouverture de fenêtre, usage incorrect des vannes thermostatiques),
- mauvaise combustion,
- mauvais fonctionnement ou mauvais réglage du système de régulation programmation.

1.3.3.2.2 - Classifications prenant en compte simultanément plusieurs caractéristiques

- fréquence d'apparition et difficulté de détection,
- fréquence d'apparition et influence sur le confort,
- fréquence d'apparition et influence sur les consommations,
- confort des occupants et difficulté de détection,
- consommations d'énergie et difficulté de détection,
- consommations d'énergie et influence sur le confort,
- fréquence d'apparition, difficulté de détection et influence à la fois sur le confort et les consommations d'énergie.

(1) Classification en fonction de la fréquence d'apparition (ou du taux de présence pour les défauts de conception) et de la difficulté de détection

Après une analyse sur chacune des caractéristiques il nous a paru nécessaire d'analyser les corrélations entre les notes obtenues pour les différentes caractéristiques.

L'objectif de cette première classification est de recenser les défauts fréquents que l'on a du mal à détecter.

Pour les défauts autres que les défauts de conception, La figure 1.4 montre que parmi les 37 défauts difficiles à détecter ($C3 \geq 1.5$) il y en a 12 qui sont fréquents ($C1 \geq 1.5$).

Ces défauts sont les suivants :

- défauts concernant la régulation des circuits secondaires,
- mauvais calcul de l'heure de relance, de l'heure d'arrêt, et mauvais réglage du programmeur d'intermittence,
- horaires d'occupation affichés non cohérents avec l'occupation,
- courbe de chauffe mal réglée,
- consigne de température de retour mal réglée et mauvaise loi de régulation de la température de retour.
- défauts concernant la gestion de la cascade des chaudières : mauvais choix du nombre des chaudières utilisées, mauvais réglage des chaudières en cascade, mauvais réglage des aquastats des chaudières, mauvaise séquence d'arrêt entre pompe et brûleur,
- défaut concernant la production d'ECS : entartrage de l'échangeur de la production d'ECS.

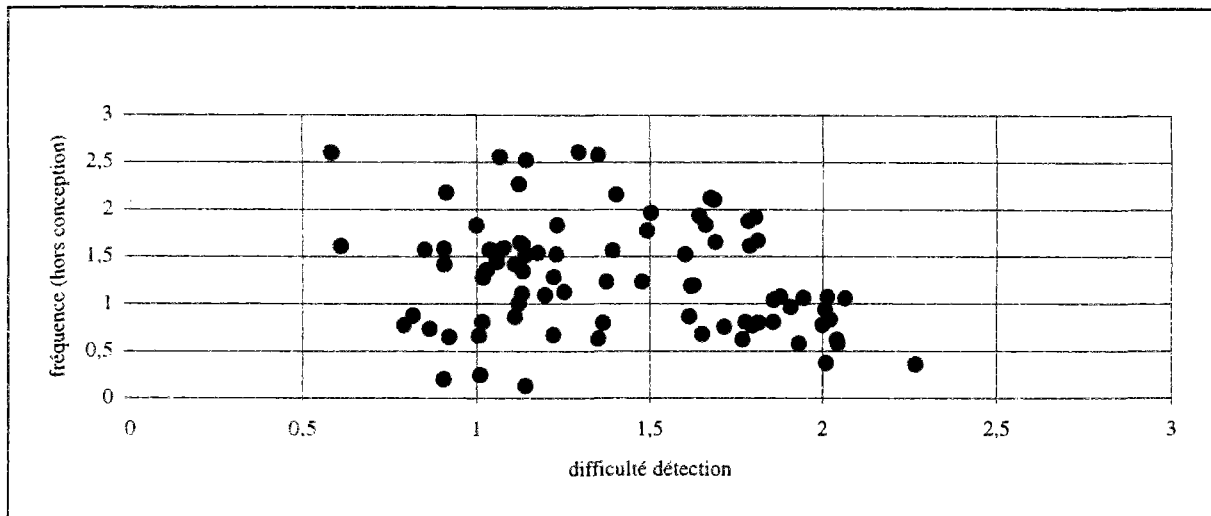


Figure 1.4 Fréquence / difficulté de détection

On peut donc dire que les deux tiers des défauts fréquents (hors défauts de conception) ne posent pas de difficulté de détection et que le dernier tiers est composé très majoritairement de défauts liés à la régulation et à la programmation.

En ce qui concerne les défauts de conception, la figure 1.5 montre clairement que presque tous les défauts de conception (19/20) sont difficiles à détecter et qu'une grande partie de ces défauts difficiles à détecter sont fréquents ($C2 \geq 1.5$).

Ces défauts sont les suivants :

- défauts concernant l'ensemble de l'installation : déséquilibre du chauffage entre différentes zones et mauvais équilibre hydraulique entre différents circuits ; mauvais dimensionnement des radiateurs dans certaines zones,
- défauts concernant le positionnement des sondes : sonde de mesure de la température intérieure placée dans une pièce non-représentative ; sondes de température extérieure mal positionnées ; sonde de la vanne thermostatique influencée par le radiateur ou d'autres sources de chaleur (rayonnement solaire, ...),
- défaut de vanne-trois-voies : mauvaise autorité hydraulique,
- défaut de pompe : surdimensionnement.

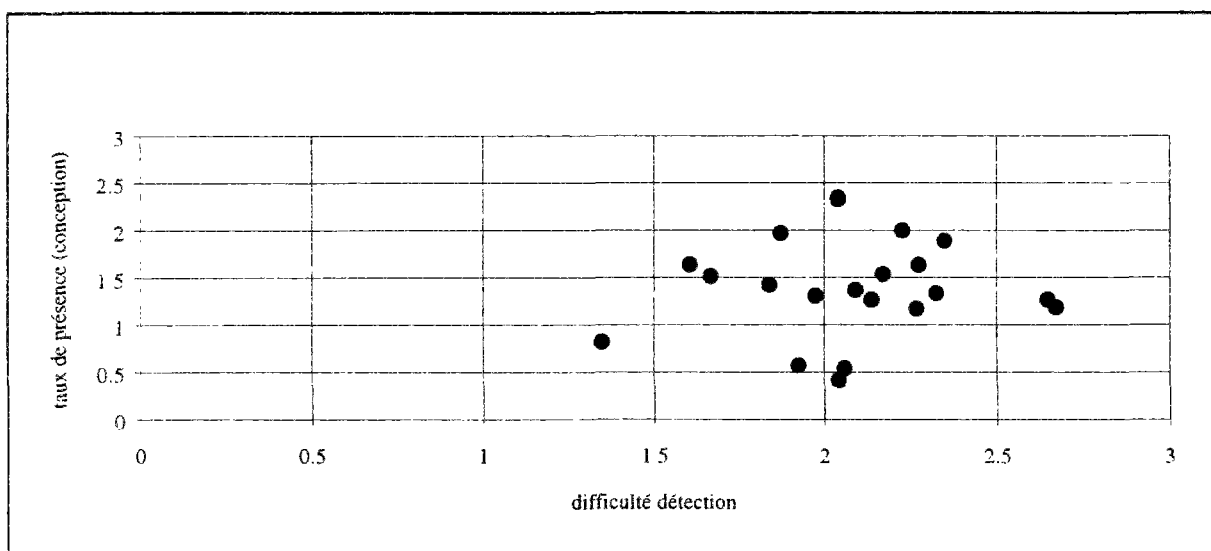


Figure 1.5 Taux de présence / difficulté de détection

(2) Classification en fonction de la fréquence d'apparition (ou du taux de présence) et de l'influence sur le confort des occupants

Parmi les défauts autres que les défauts de conception qui ont un impact important sur le confort ($C4 \geq 1.5$), plus de la moitié (15/26) sont fréquents (Figure 1.6). Il s'agit des défauts suivants :

- mauvais calcul de l'heure d'arrêt ou de l'heure de relance par le programmeur d'intermittence,
- panne électrique de la régulation des circuits secondaires (programmeur d'intermittence, régulateur en fonction de la température extérieure, horloge, ...),
- horloge non à l'heure ou horaires d'occupation affichés non cohérents avec l'occupation,
- pannes de vanne thermostatique,
- pannes ou arrêt de sécurité de brûleur,
- air dans radiateur ou dans des tuyaux,
- pannes électriques ou mécaniques de pompe.

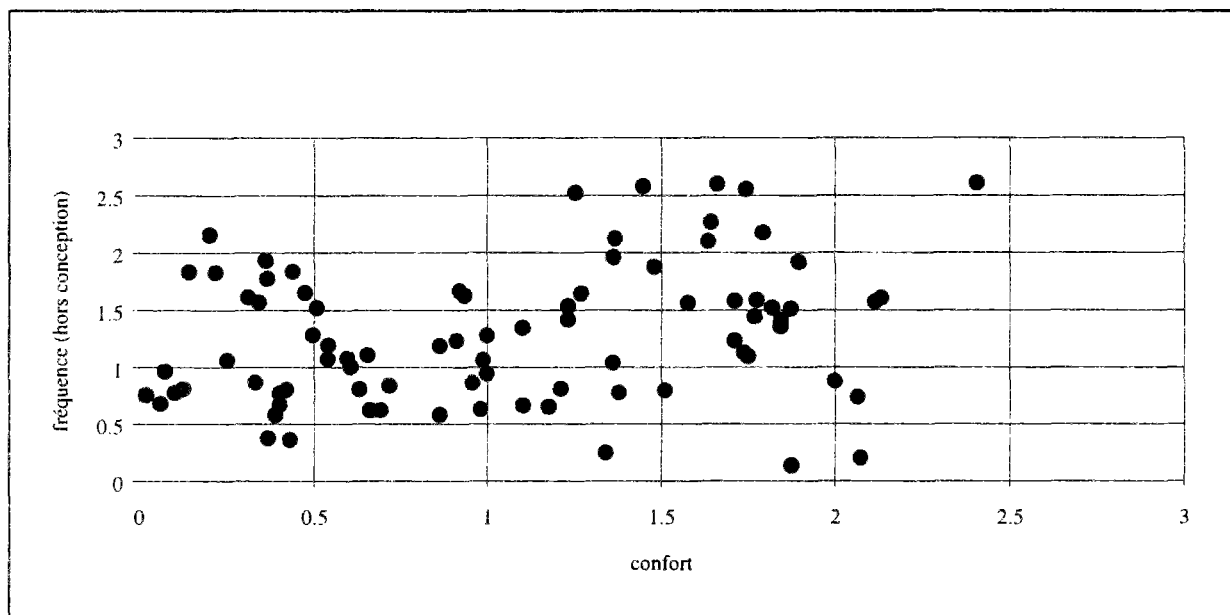


Figure 1.6 Fréquence (hors conception) / confort

En revanche, pour les défauts de conception un seul (parmi 21 défauts de conception recensés) a un impact important sur le confort (Figure 1.7). Il s'agit du sous dimensionnement du préparateur d'ECS. Il faut toutefois noter que ce n'est pas un défaut très fréquent (note 1,3 sur le critère fréquence d'apparition).

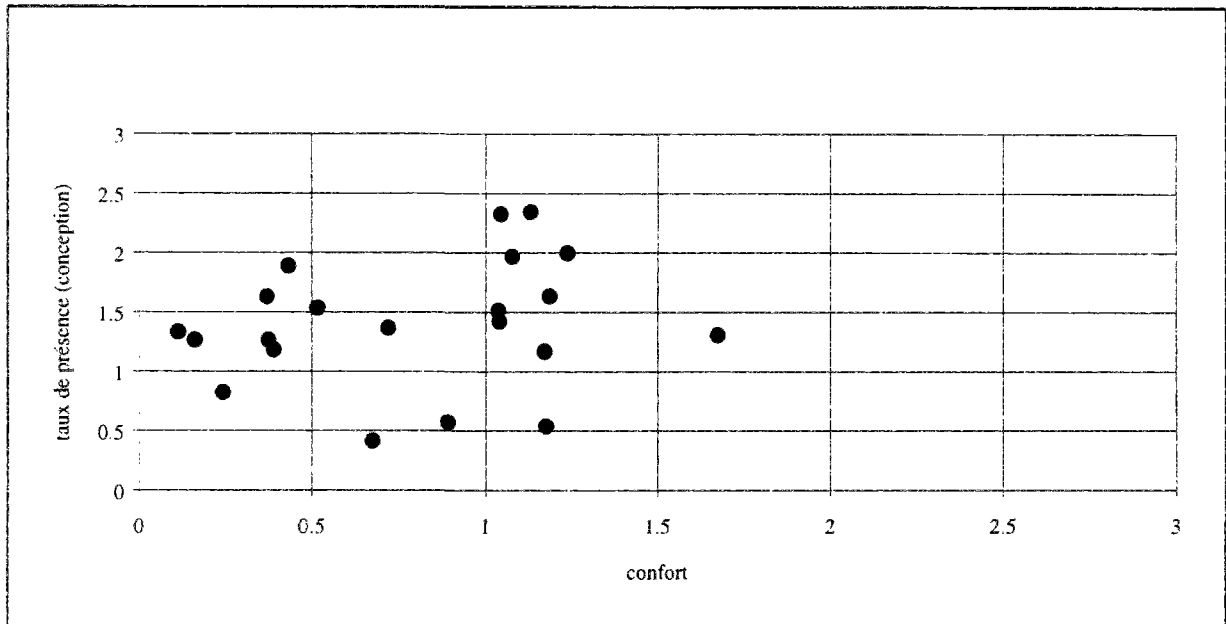


Figure 1.7 Taux de présence (conception) / confort

On peut donc dire que presque tous les défauts de conception ayant un impact important sur le confort sont détectés lors de la mise en service de l'installation et que, généralement, les défauts ayant un impact important sur le confort des occupants sont des défauts de fonctionnement.

(3) Classification en fonction de la fréquence d'apparition (ou du taux de présence) et de l'influence sur la consommation d'énergie

Parmi les défauts autres que les défauts de conception qui ont un impact important sur les consommations ($C5 \geq 1.5$), la moitié (13/25) est fréquente (Figure 1.8). Ces défauts sont les suivants :

- défauts liés aux occupants : ouverture inadéquate de fenêtre, usage incorrect de thermostat,
- défauts liés à la régulation des circuits secondaires : courbe de chauffe mal réglée, mauvais calcul de l'heure de relance ou de l'heure d'arrêt par le programmateur d'intermittence et mauvais réglage de celui-ci, horaires d'occupation affichés non cohérents avec l'occupation,
- défauts liés à la production de chaleur : encrassement de l'échangeur principal de chaudière (côté fumées), mauvaise combustion du brûleur, mauvais choix du nombre des chaudières utilisées, mauvais réglage du gestionnaire de la cascade chaudière ou panne électrique de celui-ci,
- défaut lié au préparateur d'ECS : entartrage de l'échangeur.

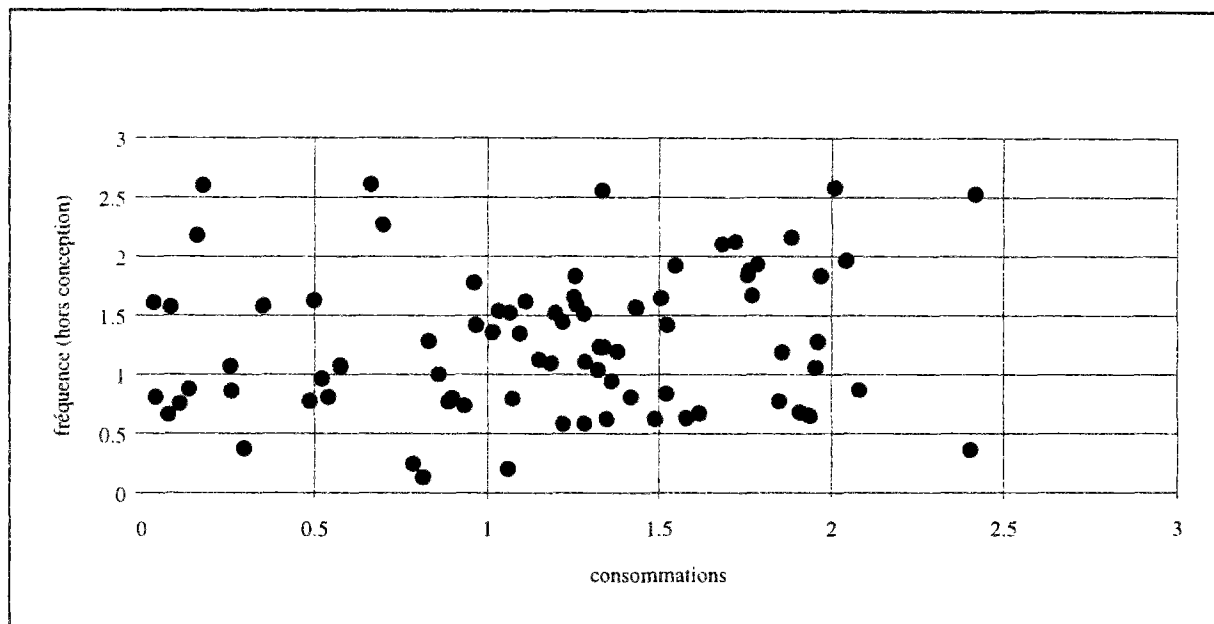


Figure 1.8 Fréquence (hors conception) / consommations

Pour les défauts de conception, parmi les 9 défauts ayant un impact important sur les consommations, 5 sont fréquents (Figure 1.9). Il s'agit :

- du déséquilibre entre différentes zones ou entre les circuits,
- du sousdimensionnement et du surdimensionnement des radiateurs dans certaines pièces ou zones particulières,
- du positionnement de sonde : sonde de mesure de la température intérieure placée dans une pièce non représentative, sondes de température extérieure mal positionnées.

En comparant les figures 1.6 et 1.8 on constate que les défauts de fonctionnement, c'est-à-dire les défauts autres que les défauts de conception, ayant un impact important sur le confort des occupants sont aussi nombreux et fréquents que ceux qui ont un impact important sur les consommations de l'énergie. En revanche si on compare les figures 1.7 et 1.9 on constate que les défauts de conception ont un impact beaucoup plus important sur les consommations que sur le confort.

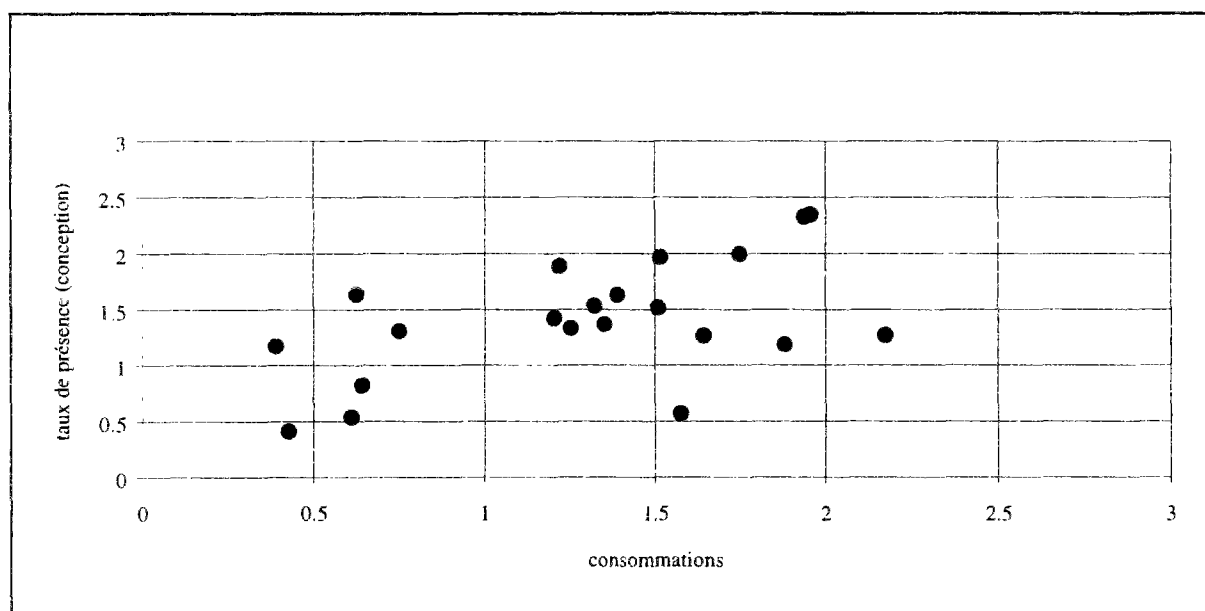


Figure 1.9 Taux de présence (conception) / consommations

(4) Classification en fonction de l'influence sur le confort et de la difficulté de leur détection

La figure 1.10 présente pour les défauts étudiés la corrélation entre l'influence sur le confort et la difficulté de détection. En analysant cette figure, on constate qu'une grande partie des défauts (qu'il s'agisse de défauts de conception ou de fonctionnement) ayant un impact important sur le confort est facile à détecter ou déjà détectée. Dans de nombreux cas des alarmes ou des indicateurs de défauts sont déjà prévus sur les installations.

Parmi les 27 défauts ayant un impact important sur le confort ($C4 \geq 1.5$) on ne trouve que 4 défauts difficiles à détecter ($C3 \geq 1.5$). Ces défauts peuvent être regroupés de la façon suivante :

- dysfonctionnements du programmeur d'intermittence, dysfonctionnements dus à une panne ou à un mauvais réglage (3 défauts),
- sous dimensionnement du préparateur d'ECS (1 défaut). C'est un défaut de conception.

Les défauts liés à une mauvaise programmation de l'intermittence nous semblent devoir être étudiés dans le détail dans la mesure où il paraît possible à peu de frais de définir des méthodes permettant de les détecter.

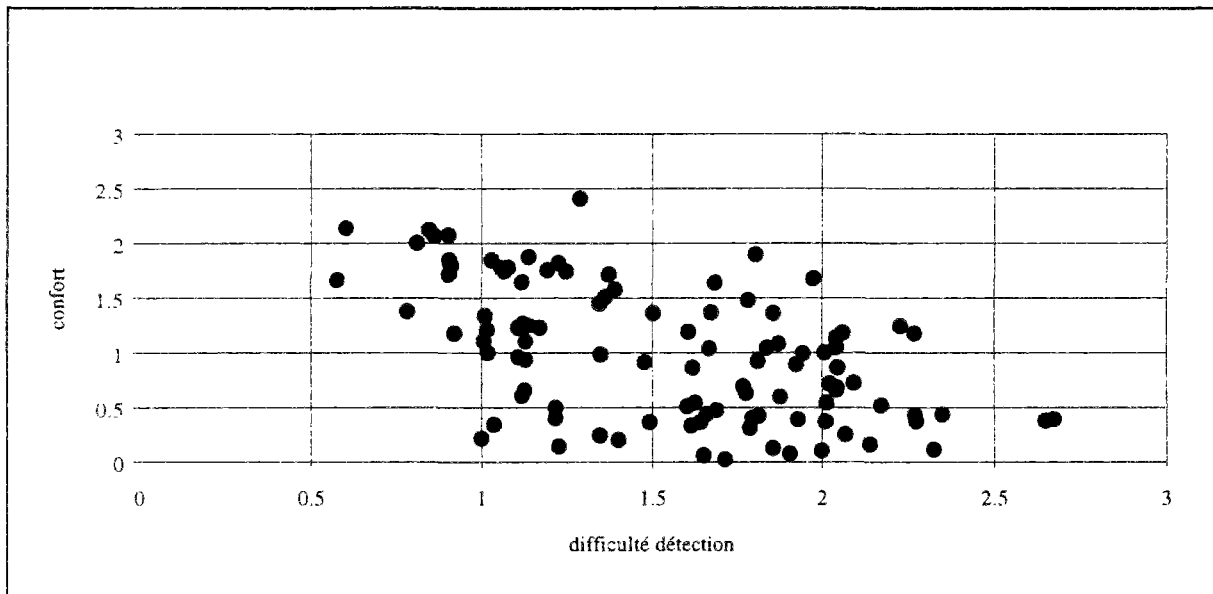


Figure 1.10 Confort / difficulté de détection

On peut donc dire que les outils aujourd'hui disponibles sur les installations permettent de détecter sans trop de difficulté la plus grande partie des défauts ayant une incidence sensible sur le confort.

(5) Classification en fonction des consommations d'énergie et de la difficulté de détection

La figure 1.11 montre la corrélation entre l'influence sur les consommations d'énergie et la difficulté de détection. On constate que les résultats obtenus sont très différents de ceux de la figure 1.10 portant sur la corrélation confort/difficulté de détection. En effet l'analyse de la figure 1.11 montre qu'une grande partie des défauts ayant un impact important sur les consommations d'énergie est difficile à détecter. Parmi les 35 défauts ayant un impact fort sur les consommations ($C5 \geq 1.5$), 24 sont difficiles à détecter ($C3 \geq 1.5$).

Ceci montre clairement que les méthodes de détection de défaut ont, dans un premier temps été développées principalement en vue de limiter l'impact des défauts sur le confort. Un travail important reste donc à faire si on veut limiter l'impact des défauts sur les consommations.

Les défauts ayant un impact important sur les consommations et qui sont difficiles à détecter se retrouvent sur différentes parties de l'installation :

- la production de chaleur (entartrage de l'échangeur principal de la chaudière, dégradation de l'isolation thermique de la chaudière, surdimensionnement du brûleur, mauvaise combustion, mauvaise gestion de la cascade chaudière),
- la régulation des circuits secondaires (programmation d'intermittence, régulation en fonction de la température extérieure, vannes de régulation non étanches en position fermée),
- le mauvais positionnement des sondes (température intérieure ou extérieure),
- l'ensemble de l'installation et du bâtiment (dégradation de l'isolation thermique, entartrage et embouage des circuits hydrauliques, déséquilibre thermique),
- des zones particulières (mauvais dimensionnement de radiateur, usage incorrect de vanne thermostatique, ouverture "inadéquate" de fenêtre),
- le préparateur d'ECS (fuite, entartrage...).

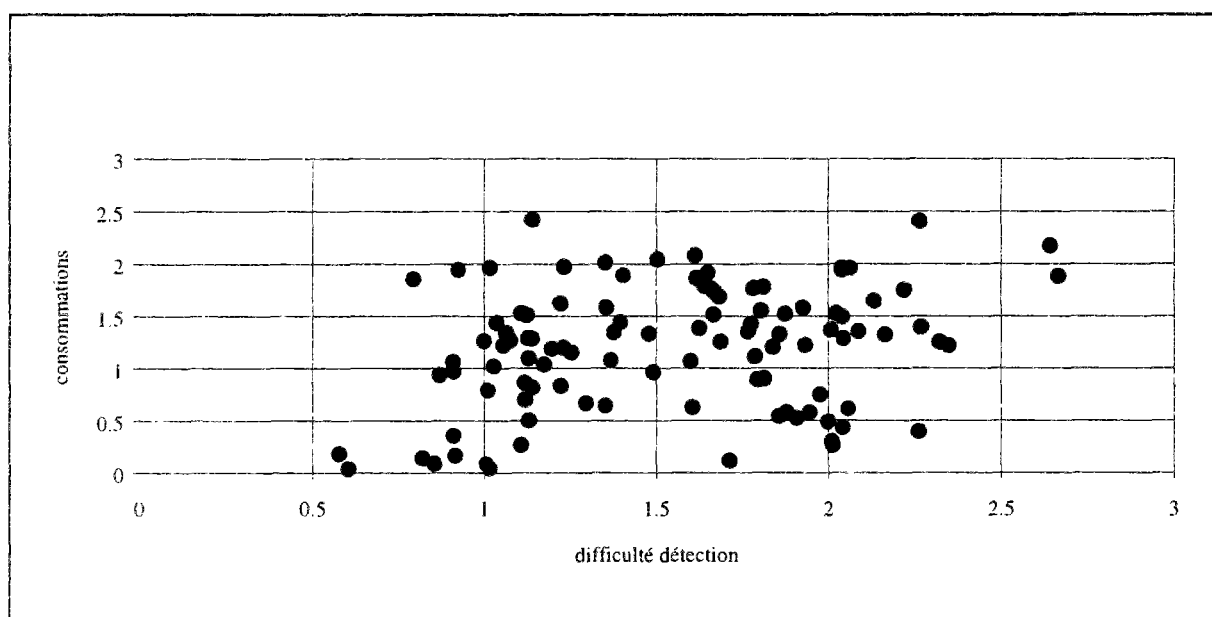


Figure 1.11 Consommations / difficulté de détection

(6) Classification en fonction des consommations d'énergie et du confort

La figure 1.12 qui montre la corrélation entre l'influence sur le confort et l'influence sur les consommations montre que les défauts ayant un impact fort sur les consommations ne sont pas les mêmes que ceux ayant un impact fort sur le confort. Les seuls défauts ayant un impact fort à la fois sur les consommations et le confort sont ceux liés à la programmation de l'intermittence.

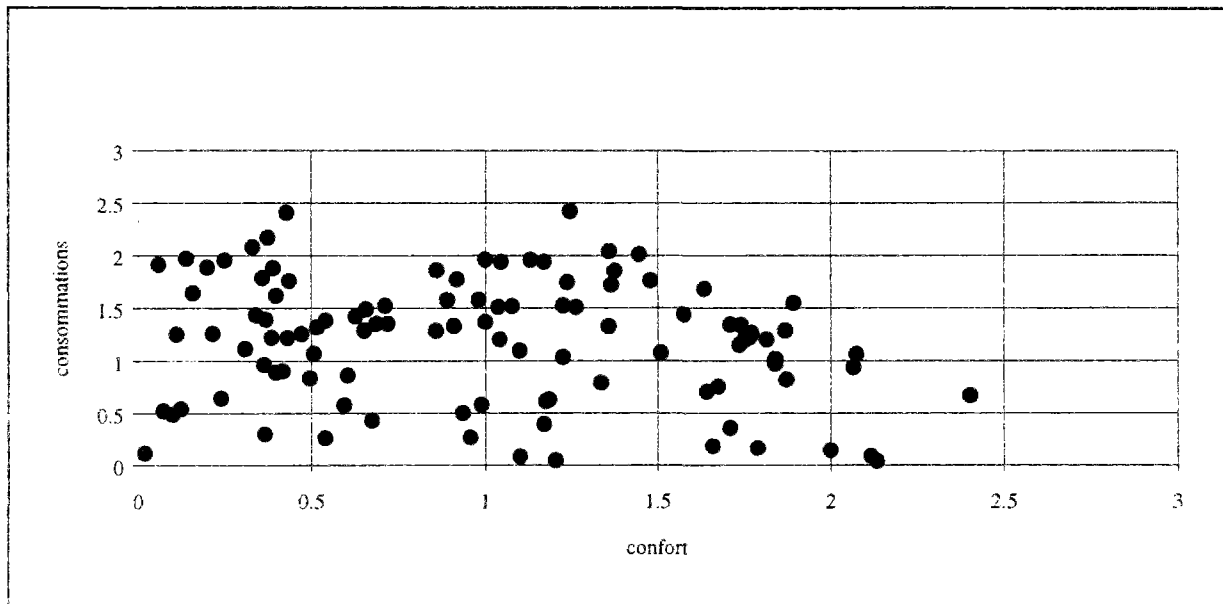


Figure 1.12 Consommations / confort

(7) Classification en fonction de l'ensemble des quatre caractéristiques

Pour faire apparaître les défauts les plus importants en tenant compte de 4 caractéristiques : C1 ou C2, C3, C4, C5 (voir Annexe B4), nous avons calculé la moyenne arithmétique des notes obtenues pour ces quatre caractéristiques. Les défauts les plus importants (moyenne arithmétique ≥ 1.5) peuvent être regroupés de la façon suivante :

- défauts liés à la régulation des circuits secondaires (programmeur d'intermittence, horloge, régulateur en fonction de la température extérieure, vanne de régulation),
- défauts liés à des zones particulières (radiateur, mauvaise manipulation de régulation locale et ouverture de fenêtre par des occupants),
- défauts liés au positionnement incorrect des sondes de température et pannes de sondes,
- défauts liés à la production de chaleur (mauvaise combustion, conduite des chaudières en cascade),
- défauts liés à la production d'ECS (entartrage, sousdimensionnement),
- défauts liés à l'ensemble de l'installation et au bâtiment (déséquilibre du chauffage entre différentes zones, l'isolation thermique ne correspond pas au calcul),
- défaut de surdimensionnement de pompe.

Pour faciliter l'élaboration de méthodes de détection de défaut, nous avons séparé en deux groupes les défauts les plus importants que nous avons choisis dans un premier temps afin de développer des méthodes automatiques de détection. Nous décrivons ci-après succinctement ces défauts et leurs conséquences :

Le premier groupe concerne les défauts de conception ou d'installation :

- défaut 1 : (composant : l'ensemble de l'installation) : déséquilibre de chauffage entre différentes zones du bâtiment,
- défaut 2 : (composant : l'ensemble de l'installation) : mauvais équilibre hydraulique entre les circuits,
- défaut 3 : (composant : radiateur) : sur-/sous-dimensionnement de radiateur dans certaines pièces ou des zones particulières,
- défaut 4 : (composant : sondes) : Sonde de mesure de la température intérieure placée dans une pièce non représentative.

Les conséquences des défauts 1, 2, 3 sont pratiquement similaires : surchauffe ou souschauffe dans certaines zones du bâtiment, gaspillage de l'énergie.

Et le défaut 4 peut entraîner : une durée de relance matinale trop courte, une courbe de chauffe pas adaptée au bâtiment, une température intérieure dans certaines pièces trop basse surtout en début de matinée. La présence du défaut 4 conduit à utiliser avec précaution les mesures de températures intérieures pour la détection d'autres défauts.

Le deuxième groupe concerne les défauts qui apparaissent durant le fonctionnement de l'installation et varient au cours du temps. Ces défauts doivent être surveillés et diagnostiqués en temps réel :

- défaut 5 : (composant : régulateur en fonction de température extérieure) : courbe de chauffe mal réglée. Les conséquences de ce défaut sont : température intérieure corrélée avec la température extérieure, inconfort des occupants, vanne thermostatique presque fermée. Les exploitants ne sont pas toujours satisfaits des courbes de chauffe auto-adaptatives qui sont proposées par les constructeurs pour résoudre ce problème.
- défaut 6 : (composant : programmeur d'intermittence) : mauvais calcul de l'heure de relance ou d'arrêt par le programmeur d'intermittence. Ce type de défaut peut entraîner : une température intérieure trop basse (surtout le lundi matin et à la fin de l'après-midi) ou trop haute, un inconfort des occupants, une augmentation de la consommation d'énergie.
- défaut 7 : (composant : brûleur) : mauvaise combustion. La mauvaise combustion comprend plusieurs aspects : excès d'air trop important, défaut d'air. Les conséquences de ce défaut sont : augmentation de la température de fumées, baisse de l'efficacité de l'échangeur principal de chaudière, augmentation des consommations d'énergie.
- défaut 8 : (composant : chaudière) : entartrage et encrassement de l'échangeur principal de chaudière. Les conséquences sont : baisse de l'efficacité de l'échangeur principal, augmentation de la consommation de l'énergie, température d'eau à la sortie de chaudière trop basse. On doit signaler qu'en pratique, l'entartrage est beaucoup plus difficile à détecter que l'encrassement.
- défaut 9 : (composant : vanne-trois-voies / vanne-deux-voies de régulation des circuits secondaires) : vanne-trois-voies non étanche en position fermée (fuite de la voie directe vers bypass ou course de limiteur mal positionnée). Les conséquences sont : température de départ de circuit secondaire trop haute quand la consigne de température intérieure trop basse (par exemple pendant la période d'arrêt), température intérieure trop haute pendant la période d'inoccupation, gaspillage de l'énergie.
- défaut 10 : (composant : préparateur d'ECS) : entartrage. Les conséquences de ce défaut sont : débit d'ECS trop bas ou température d'ECS trop basse, plainte d'usager contre l'insuffisance de la température ou du débit d'ECS, augmentation de la consommation d'énergie.

1.3.4 - CONCLUSION

Le travail concernant le recensement et la classification des défauts nous a permis d'obtenir une base de données importante sur les défauts de fonctionnement des installations de chauffage à eau chaude. Cette base de données comporte pour une centaine de défauts des indications sur leur fréquence, sur la difficulté de leur détection, sur leur influence sur les consommations d'énergie et sur le confort.

L'analyse des données de cette base qui est présentée dans cette thèse permet de tirer une première série de conclusions. Ces conclusions sont les suivantes :

- les défauts de conception qui ont un impact important sur le confort sont généralement détectés au début de la vie de l'installation,
- les défauts de conception qui ont un impact important uniquement sur les consommations risquent fortement de ne pas être détectés au début de la vie de l'installation,
- les défauts de conception sont difficilement détectés en cours de fonctionnement de l'installation par les équipes d'exploitation,
- les défauts de fonctionnement, c'est-à-dire les défauts autres que les défauts de conception, qui ont un impact sensible sur le confort sont à quelques exceptions près bien détectés par les équipes d'exploitation,
- de nombreux défauts de fonctionnement ayant un impact important sur les consommations sont aujourd'hui difficiles à détecter,
- de nombreux défauts ayant un impact tant sur le confort que sur les consommations apparaissent au niveau de la régulation des circuits secondaires sans qu'il existe de méthode simple pour les détecter,
- il n'existe pas de systèmes de détection automatique des défauts apparaissant au niveau d'une zone, et ceux-ci ne sont donc généralement détectés que par l'intermédiaire des plaintes des occupants.

Ce travail nous a donc permis de sélectionner les défauts pour lesquels il paraissait le plus utile de développer des méthodes de détection automatique.

Les défauts sélectionnés ont été simulés, dans un premier temps, par la méthode numérique ; ceci nous permet d'obtenir une base de données représentant le fonctionnement du système comportant différents types de défaut. Cette simulation sera présentée dans le Chapitre 2

CHAPITRE 2

MODELISATION ET SIMULATION DE DEFAUTS

L'objectif de ce chapitre est de générer, par la voie de simulation, une base de données numériques qui représente le fonctionnement de systèmes de chauffage sans et avec défauts. Cette base de données devra nous servir dans la phase suivante au développement d'une méthode automatique de détection et de diagnostic de défauts pour les systèmes de chauffage.

2.1 - Introduction	32
2.2 - Modélisation du système de chauffage	33
2.2.1 - Description du système en vue de la simulation des défauts.....	33
2.2.2 - Présentation des modèles existants.....	33
2.2.3 - Modélisation d'une chaudière à gaz	34
2.2.3.1 - Introduction	36
2.2.3.2 - Calcul de la température de départ d'eau de chaudière	36
2.2.3.2.1 - Bilan énergétique	36
2.2.3.2.2 - Paramétrage	39
2.2.3.3 - Calcul de la température des gaz de combustion à la sortie de l'échangeur.....	39
2.2.3.3.1 - Température des gaz de combustion au niveau du brûleur.....	40
2.2.3.3.2 - Température des gaz de combustion à la sortie de l'échangeur en régime permanent.....	43
2.2.3.4 - Conclusion	44
2.3 - Construction d'un simulateur du système de chauffage.....	45
2.3.1 - Présentation des logiciels MATLAB et SIMULINK	45
2.3.2 - Elaboration d'un simulateur du système de chauffage.....	47
2.4 - Simulation du système de chauffage sans et avec défauts.....	49
2.4.1 - Modélisation des défauts à simuler	49
2.4.2 - Simulation du fonctionnement du système comportant des défauts.....	50
2.4.2.1 - Présentation des classes ou modes de fonctionnement	50
2.4.2.2 - Présentation graphique de données de simulation.....	50

2.1 - INTRODUCTION

Le développement de méthodes de détection impose que l'on dispose d'une base de données représentant le fonctionnement d'installations comportant différents défauts. Ces données peuvent être obtenues par mesure sur des installations réelles, par mesure sur un banc d'essai ou par simulation numérique [Spriet, 1982]. Disposant d'un ensemble des modèles permettant de simuler le fonctionnement d'une installation en l'absence de défaut, on a choisi la voie de la simulation.

Le travail présenté dans ce chapitre consiste à :

- simplifier le système de référence (voir paragraphe 1.2) pour faciliter la simulation,
- assembler les modèles existants qui décrivent un système de chauffage à eau chaude,
- élaborer un modèle de chaudière puisque le modèle existant ne s'adapte pas à la simulation de certains défauts de chaudière recensés,
- modifier certains modèles (en particulier le modèle du bâtiment et le modèle de programmeur-régulateur, afin de s'adapter au logiciel de simulation (SIMULINK) d'une part, et d'autre part de pouvoir simuler plus objectivement les défauts, surtout les défauts liés au programmeur-régulateur,
- déterminer les paramètres de l'ensemble des modèles,
- choisir, dans un premier temps, un nombre réduit des défauts les plus importants en s'appuyant sur les résultats des recensement et classification des défauts (voir paragraphe 1.3) et, ensuite, les modéliser par changement de certains paramètres de modèles,
- simuler, durant une saison de chauffage entière, le fonctionnement de l'installation en présence ou non des défauts.

On présente successivement la modélisation du système de chauffage, l'élaboration d'un simulateur du système de chauffage et la simulation des défauts effectuée.

2.2 - MODELISATION DU SYSTEME DE CHAUFFAGE

2.2 1 - DESCRIPTION DU SYSTEME EN VUE DE LA SIMULATION DES DEFAUTS

Le système de chauffage simulé est une version simplifiée du système de référence de chauffage décrit dans le paragraphe 1.2. Les modifications principales dans cette version vis-à-vis du système de référence sont les suivantes :

- compte tenu de la complexité du système, la partie production d'eau chaude sanitaire et la partie alimentation d'eau ne sont pas simulées dans un premier temps,
- le modèle du bâtiment a été simplifié pour ne créer que deux zones (nord et sud) au lieu des six zones initialement prévues. Les paramètres du modèle tels que résistances, capacités, surfaces des baies vitrées, surfaces d'échange des radiateurs, etc. sont donc recalculés en fonction du bâtiment structuré en deux zones,
- le modèle de programmeur-régulateur du système simulé a été modifié par rapport à celui du système de référence pour pouvoir simuler plus réellement et plus facilement les différents défauts issus des résultats d'enquête.

Notons que les modifications du modèle de bâtiment et du modèle de programmeur ne sont pas présentées ici pour ne pas alourdir la lecture.

2.2.2 - PRESENTATION DES MODELES EXISTANTS

Les modèles, décrivant les composants du système simplifié sans défaut, existent déjà en grande partie au CSTB, [Caccavelli, 1991], [François, 1990] et [Laret, 1980]. Ce sont principalement les suivants :

- chaudière classique à gaz,
- bâtiment six zones,
- programmeur-régulateur :
 - . horloge hebdomadaire, programmeur d'intermittence
 - . conduite en séquence des chaudières, régulateur sur vanne-trois-voies
 - . éléments de régulation (aquastat de chaudière, sonde de température, robinet thermostatique),
- composants hydrauliques :
 - . pompe, bouteille casse-pression, tuyau, vanne-trois-voies, convergent, divergent, radiateur

L'assemblage de ces modèles permet de faire une simulation du fonctionnement d'un système complet de chauffage à eau chaude qui ne comporte pas de défaut.

Or, pour construire une base de données qui décrit le comportement du système avec défauts, il a fallu modifier certains modèles afin de pouvoir changer facilement les valeurs numériques de certains paramètres de modèle, ce changement exprimant la présence de tel ou tel défaut. Citons par exemple qu'un défaut d'encrassement de chaudière peut s'exprimer par un changement (diminution) de la conductance thermique entre l'eau d'un côté et les fumées de l'autre.

Toutes les modifications de modèles ne sont pas ici présentées. En revanche pour mieux montrer l'approche modélisation-simulation des défauts, on décrit en détail un modèle de chaudière permettant d'introduire certains défauts du brûleur et de l'échangeur d'une chaudière à gaz.

2.2.3 - MODELISATION D'UNE CHAUDIERE A GAZ

Notations

Les notations suivantes seront employées par le modèle de chaudière présenté dans ce paragraphe.

C_m : fraction massique de carbone

C_{eq_eau} : capacité thermique équivalente globale de chaudière (eau + métal) (J/°C)

$[CO_2]$: fraction molaire (volumique) en CO_2 des fumées (gaz de combustion) sèches

C_{p_eau} : chaleur massique d'eau (J/kg.°C)

$\overline{C_{p_air}}$: chaleur massique moyenne de l'air comburant entre T_{prech} et T_o (J / kg.°C)

$\overline{C_{p_g}}$: chaleur massique moyenne des gaz de combustion entre T_{g1} et T_o (J /kg.°C)

$\overline{DCM_{max}}$: débit maximal de capacité thermique moyenne parmi les deux fluides d'échange (W / °C)

$\overline{DCM_{min}}$: débit minimal de capacité thermique moyenne parmi les deux fluides d'échange (W / °C)

$\overline{DCM_{T_{g1} - T_{g2}}}$: débit capacité thermique moyenne des gaz de combustion entre T_{g1} et T_{g2} (W / °C)

$\overline{DCM_{T_{g1} - T_o}}$: débit de capacité thermique moyenne des gaz de combustion entre T_{g1} et T_o (W / °C)

$\overline{DCM_{T_1 - T_2}}$: débit de capacité thermique moyenne des gaz de combustion entre T_1 et T_2 (W / °C)

E_{air} : fraction molaire d'excès d'air de combustion comparée avec la quantité d'air nécessaire pour la combustion stoechiométrique

H_m : fraction massique d'hydrogène

K : conductance thermique entre les gaz de combustion et l'eau circulant dans l'échangeur (W / °C)

K_{bal} : coefficient de déperdition thermique par balayage en phase d'entretien (W / °C)

K_{eau} : conductance thermique d'eau (W / °C)

K_p : conductance thermique (parois) entre l'eau circulant dans l'échangeur et l'air ambiant (W / °C)

K_{pg} : coefficient de perte par les gaz de combustion dépendant de combustible et de $[CO_2]$ (1 / °C)

$\dot{M}_{air}$: débit massique d'air comburant (kg / s)

$\dot{M}_{br}$: débit massique de combustible du brûleur (kg / s)

$\dot{M}_{eau}$: débit massique d'eau circulant dans l'échangeur (kg / s)

M_{eq_eau} : masse équivalente en eau de chaudière (l'eau et le métal) (kg)

$\dot{M}_g$: débit massique des gaz de combustion (kg / s)

$\dot{M}_v$: débit massique de combustible de la veilleuse (kg / s)

N_m :	fraction massique d'azote	
NUT :	nombre d'unités de transfert	
O_m :	fraction massique d'oxygène	
P_{atm} :	pression atmosphérique	(atm)
P_{bal} :	perte par le balayage	(W)
PCI :	pouvoir calorifique inférieur de combustible	(J / kg)
P_g :	perte par les gaz de combustion	(W)
P_p :	perte par les parois	(W)
R_{ap} :	rapport des débits de capacité thermique moyenne des deux fluides : ($\overline{DCM}_{min} / \overline{DCM}_{max}$)	
S_m :	fraction massique de soufre	
t :	temps	(s)
T_{air} :	température de l'air ambiant	(°C)
T_{d_eau} :	température de départ d'eau de chaudière en régime dynamique	(°C)
T_{g1} :	température des gaz de combustion dans le foyer en régime permanent	(°C)
T_{g2} :	température des gaz de combustion à la sortie de l'échangeur en régime permanent	(°C)
T_{g2_dym} :	température des gaz de combustion à la sortie de l'échangeur en régime dynamique	(°C)
T_{g2_min} :	température minimale des gaz de combustion à la sortie d l'échangeur (en phase d'entretien)	(°C)
T_o :	température de référence dans les conditions normales	(°C)
T_{prech} :	température de préchauffage de l'air comburant	(°C)
T_{r_eau} :	température de retour d'eau de chaudière en régime dynamique	(°C)
T_{ros} :	température de rosée de l'air comburant	(°C)
ε_1 :	efficacité de l'échangeur en phase de marche de chaudière	
ε_2 :	efficacité de l'échangeur en phase d'entretien de chaudière	
τ_1 :	constante de temps de l'évolution de la température des gaz de combustion en phase de marche de chaudière	(s)
τ_2 :	constante de temps de l'évolution de la température des gaz de combustion en phase d'entretien de chaudière	(s)
Φ_{max} :	puissance d'échange maximum théoriquement possible	(W)
$\Phi_{réel}$:	puissance thermique réellement échangée entre deux fluides	(W)
Φ_{br} :	puissance fournie par le brûleur	(W)
Φ_v :	puissance fournie par la veilleuse	(W)

2.2.3.1 - Introduction

Le modèle présenté ici a pour objectif de simuler le comportement dynamique de systèmes de chauffage à eau chaude en présence de certains défauts importants de chaudière à gaz. Comme référence nous avons utilisé une chaudière acier de marque GUILLOT - OPTIMAGAZ (E87). L'OPTIMAGAZ (E87) est une chaudière à haut rendement destinée à la production d'eau chaude pour les réseaux de chauffage équipés de convecteurs, radiateurs, aérothermes, plancher chauffant et les réseaux de préparations d'eau chaude sanitaire dans les bâtiments d'habitation, de bureaux ou d'industries.

Ce modèle calcule d'une part la température de départ d'eau de chaudière en régime dynamique et d'autre part la température des gaz de combustion à la sortie de l'échangeur, ce qui devra permettre de simuler certains défauts de brûleur et d'échangeur.

2.2.3.2 - Calcul de la température de départ d'eau de chaudière

2.2.3.2.1 - Bilan énergétique

Le bilan énergétique au niveau de chaudière sera établi avec les 6 hypothèses suivantes (de H1 à H6), qui ont été utilisées et validées pour une autre étude, [Caccavelli, 1991] :

- H1 : une chaudière traditionnelle à gaz sans condensation.
- H2 : deux phases de fonctionnement :
 - 1^{ère} : marche c'est-à-dire brûleur et ventilateur en fonctionnement (débit brûleur constant),
 - 2^{ème} : entretien c'est-à-dire veilleuse en fonctionnement et ventilateur à l'arrêt.
- H3 : échangeur principal caractérisé par un seul noeud (chaudière peu inerte), c'est-à-dire toute la masse d'eau et de métal supposée concentrée en un point.
- H4 : température du noeud de l'échangeur principal assimilée à la température de départ d'eau de chaudière.
- H5 : pertes par balayage pendant la phase d'entretien (P_{bal}) et par les parois de chaudière (P_p) supposées linéaires par rapport à la différence entre la température de départ d'eau et la température de l'air ambiant dans la chaufferie ($T_{d_eau} - T_{air}$).
- H6 : perte par les gaz de combustion négligée en phase d'entretien.

Ce bilan s'écrit alors :

en phase de marche de chaudière :

$$\Phi_{br} - P_g - P_p = M_{eq_eau} C_{p_eau} \frac{dT_{d_eau}}{dt} + \dot{M}_{eau} C_{p_eau} (T_{d_eau} - T_{r_eau}) \quad (2.1)$$

en phase d'entretien de chaudière :

$$\Phi_v - P_{bal} - P_p = M_{eq_eau} C_{p_eau} \frac{dT_{d_eau}}{dt} + \dot{M}_{eau} C_{p_eau} (T_{d_eau} - T_{r_eau}) \quad (2.2)$$

avec :

$$\Phi_{br} = \dot{M}_{br} PCI \quad (2.3)$$

$$\Phi_v = \dot{M}_v PCI \quad (2.4)$$

$$P_{bal} = K_{bal} (T_{d_eau} - T_{air}) \quad (2.5)$$

$$P_g = \frac{K_{pg} (T_{g2_dym} - T_{air})}{10000[CO_2]} \Phi_{br} , [Morant, 1982] \quad (2.6)$$

$$P_p = K_p (T_{d_eau} - T_{air}) \quad (2.7)$$

Les équations (2.1) et (2.2) peuvent s'écrire respectivement :

$$\Phi_{br} - P_g = C_{eq_eau} \frac{dT_{d_eau}}{dt} + \frac{T_{d_eau} - T_{r_eau}}{\frac{1}{K_{eau}}} + \frac{T_{d_eau} - T_{air}}{\frac{1}{K_p}} \quad (2.8)$$

$$\Phi_v - P_{bal} = C_{eq_eau} \frac{dT_{d_eau}}{dt} + \frac{T_{d_eau} - T_{r_eau}}{\frac{1}{K_{eau}}} + \frac{T_{d_eau} - T_{air}}{\frac{1}{K_p}} \quad (2.9)$$

avec :

$$C_{eq_eau} = M_{eq_eau} C_{p_eau} \quad (2.10)$$

$$K_{eau} = \dot{M}_{eau} C_{p_eau} \quad (2.11)$$

A l'aide des équations (2.8) et (2.9), l'analogie électrique du modèle de premier ordre de chaudière peut être schématiquement présentée sur la Figure 2.1.

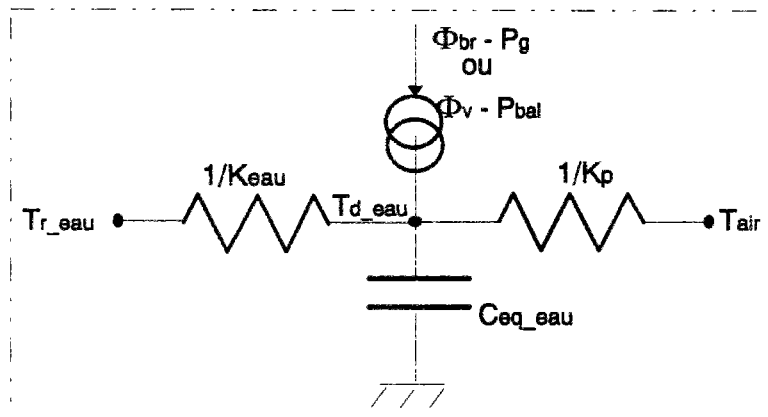


Figure 2.1 Analogie électrique du modèle de chaudière de 1^{er} ordre

Les équations (2.8) et (2.9) peuvent s'écrire encore sous forme d'équation d'état :

$$\frac{dx}{dt} = Ax + Bu \quad (2.12)$$

$$y = Cx + Du \quad (2.13)$$

avec :

$$x = T_{d_eau}$$

$$u = \begin{Bmatrix} \Phi_{br} - P_g \\ T_{r_eau} \\ T_{air} \end{Bmatrix} \text{ en phase de marche de brûleur}$$

ou

$$u = \begin{Bmatrix} \Phi_v - P_{bal} \\ T_{r_eau} \\ T_{air} \end{Bmatrix} \text{ en phase d'entretien de brûleur}$$

$$A = \begin{bmatrix} -\frac{K_{eau} + K_p}{C_{eq_eau}} \end{bmatrix}$$

$$B = \begin{bmatrix} \frac{1}{C_{eq_eau}} & \frac{K_{eau}}{C_{eq_eau}} & \frac{K_p}{C_{eq_eau}} \end{bmatrix}$$

$$C = [1]$$

$$D = [0 \ 0 \ 0]$$

L'évolution de la température de départ d'eau de chaudière au cours du temps se fait donc avec les équations (2.12) et (2.13).

2.2.3.2.2 - Paramétrage

Etant donnés la puissance de chaudière, le type de combustible, et le débit nominal de l'eau circulant dans la chaudière, les quantités Φ_{br} , Φ_v , K_{eau} , C_{eq_eau} , sont facilement calculées. Par contre les déterminations de P_g (liée à K_{pg} , T_{g2_dym} et $[CO_2]$, voir la formule (2.6)), P_{bai} (liée à K_{bai}), et K_p sont beaucoup plus délicates. Pour les déterminations de K_{pg} , K_{bai} et K_p nous avons recouru à d'autres études, [Caccavelli, 1991] & [Morant, 1982].

On présentera dans §2.2.3.3 la modélisation de l'évolution de la température des gaz de combustion à la sortie de l'échangeur en fonction de la fraction molaire en CO_2 dans les gaz de combustion secs. Ceci aura pour objectif de **modéliser un défaut de combustion - excès d'air de combustion trop important**.

2.2.3.3 - Calcul de la température des gaz de combustion à la sortie de l'échangeur

Le calcul de la température des gaz de combustion à la sortie de l'échangeur en régime dynamique T_{g2_dym} se fait avec l'hypothèse que l'évolution de la température des gaz de combustion à la sortie de l'échangeur se réalise selon un modèle d'ordre 1, [Caccavelli, 1991], c'est-à-dire :

en phase de marche de chaudière :

$$\tau_1 \frac{dT_{g2_dym}}{dt} = -T_{g2_dym} + T_{g2} \quad (2.14)$$

en phase d'entretien de chaudière :

$$\tau_2 \frac{dT_{g2_dym}}{dt} = -T_{g2_dym} + T_{g2_min} \quad (2.15)$$

avec :

$$T_{g2_min} = T_{air} + \varepsilon_2 (T_{d_eau} - T_{air}) \quad (2.16)$$

Après avoir déterminé les constantes de temps τ_1 et τ_2 , l'efficacité de l'échangeur en phase d'entretien ε_2 , ainsi que la température des gaz de combustion à la sortie de l'échangeur en régime permanent T_{g2} , nous utilisons les équations (2.14) et (2.15) pour évaluer la température des gaz de combustion à la sortie de l'échangeur en régime dynamique T_{g2_dym} .

Les déterminations des τ_1 , τ_2 , et ε_2 font référence à une autres étude, [Caccavelli, 1991] ; mais en revanche le calcul de T_{g2} sera présenté dans les paragraphes suivants.

Pour calculer T_{g2} , nous commençons par le calcul de la température des gaz de combustion dans le foyer T_{g1} , c'est-à-dire au niveau du brûleur.

2.2.3.3.1 - Température des gaz de combustion au niveau du brûleur

(1) Bilan énergétique

A l'aide du pouvoir calorifique inférieur⁽¹⁾ et en admettant une valeur moyenne pour la chaleur massique des gaz de combustion entre la température de référence (T_0) et la température des gaz de combustion (T_{g1}) (hypothèse de combustion adiabatique⁽²⁾), le bilan énergétique de combustion peut s'écrire en première approximation [Sacadura, 1978] :

$$\dot{M}_{br} PCI + \dot{M}_{air} \overline{C_{p_air}} (T_{prech} - T_0) = \dot{M}_g \overline{C_{p_g}} (T_{g1} - T_0) = \overline{DCM_{T_{g1}-T_0}} (T_{g1} - T_0) \quad (2.17)$$

En ce qui concerne notre étude, nous supposons qu'il n'y a pas de préchauffage de l'air comburant, l'équation (2.17) est simplifiée en l'équation (2.18) :

$$\dot{M}_{br} PCI = \overline{DCM_{T_{g1}-T_0}} (T_{g1} - T_0) \quad (2.18)$$

où, nous avons encore :

$$T_{g1} = T_0 + \frac{\dot{M}_{br} PCI}{\overline{DCM_{T_{g1}-T_0}}} \quad (2.19)$$

Puisque le débit de capacité thermique moyenne $\overline{DCM_{T_{g1}-T_0}}$ des gaz de combustion entre T_{g1} et T_0 dépend de la température T_{g1} , on effectue le calcul de T_{g1} par itération, en évaluant, au départ, la valeur de $\overline{DCM_{T_{g1}-T_0}}$ à des températures estimées. Le nombre d'itérations dépendra de la qualité de cette estimation initiale.

(2) Débit de capacité thermique moyenne des gaz de combustion entre deux températures

Afin de calculer le débit de capacité thermique moyenne des gaz de combustion entre deux températures différentes T_1 et T_2 , nous commençons par le calcul de la masse des constituants des gaz de combustion pour 1 kg de combustible.

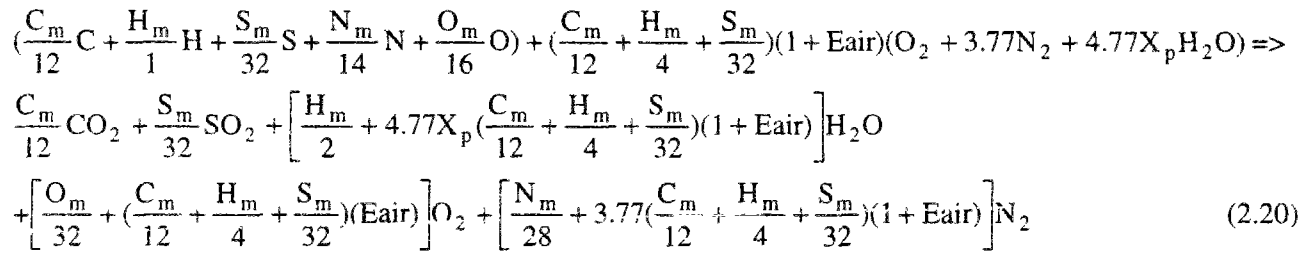
Masse des constituants des gaz de combustion pour 1 kg de combustible

Le calcul de la masse des constituants des gaz de combustion est fait d'une part à partir de l'équation de la réaction chimique qui dégage de l'énergie thermique, et d'autre part en fonction de la composition pondérale du combustible, de l'excès d'air (ou de la teneur en molaire du CO_2 des fumées sèches) et de la pression ainsi que de la température de rosée du comburant.

⁽¹⁾ Le pouvoir calorifique inférieur (PCI) de combustible correspond à 1 kg de combustible pris à la température de référence (T_0), les gaz de combustion étant ramenés à T_0 et l'eau restant sous forme de vapeur. Notons également qu'on peut indifféremment utiliser les notions de pouvoir calorifique "inférieur" ou "supérieur", à condition d'utiliser la même convention pour définir les débits thermiques et les rendements de chaudière.

⁽²⁾ La combustion adiabatique suppose qu'il n'y a pas d'échanges thermiques entre les gaz de combustion et l'environnement.

Dans le cas où l'air est utilisé comme comburant et où on suppose une combustion complète avec excès d'air, c'est-à-dire sans oxyde de carbone (CO) dans les gaz de combustion, nous pouvons écrire l'équation de la réaction chimique :



avec :

$$C_m + H_m + S_m + N_m + O_m = 1.$$

$$X_p = P_{H_2O} / (P_{atm} + P_{H_2O})$$

$$P_{H_2O} = \exp[15.87 - 5714 / (T_{ros} + 273)]$$

On a donc la fraction molaire de CO₂ dans les gaz de combustion secs

$$[CO_2] = \frac{\frac{C_m}{12}}{\frac{C_m}{12} + \frac{S_m}{32} + \left[\frac{O_m}{32} + \left(\frac{C_m}{12} + \frac{H_m}{4} + \frac{S_m}{32} \right) (E_{air}) \right] + \left[\frac{N_m}{28} + 3.77 \left(\frac{C_m}{12} + \frac{H_m}{4} + \frac{S_m}{32} \right) (1 + E_{air}) \right]} \quad (2.21)$$

d'où : l'excès d'air de combustion E_{air} :

$$E_{air} = \frac{\frac{C_m}{12} - [CO_2] \left[\frac{C_m}{12} + \frac{S_m}{32} + \frac{N_m}{28} + \frac{O_m}{32} + 3.77 \left(\frac{C_m}{12} + \frac{H_m}{4} + \frac{S_m}{32} \right) \right]}{4.77 [CO_2] \left(\frac{C_m}{12} + \frac{H_m}{4} + \frac{S_m}{32} \right)} \quad (2.22)$$

A l'aide de l'équation (2.20), on obtient la masse (en kg) des constituants des gaz de combustion pour 1 kg de combustible :

$$MCO_2 = (C_m / 12) 44 = 3.67 C_m \quad (2.23)$$

$$MH_2O = 18 [H_m / 2 + 4.77 X_p (1 + E_{air}) (C_m / 12 + H_m / 4 + S_m / 32)] \quad (2.24)$$

$$MN_2 = 28 [N_m / 28 + 3.77 (1 + E_{air}) (C_m / 12 + H_m / 4 + S_m / 32)] \quad (2.25)$$

$$MO_2 = 32 [E_{air} (C_m / 12 + H_m / 4 + S_m / 32) + O_m / 32] \quad (2.26)$$

$$MSO_2 = (S_m / 32) 64 = 2 S_m \quad (2.27)$$

Débit de capacité thermique moyenne des gaz de combustion entre deux températures

Une fois que nous avons calculé la masse des constituants des gaz de combustion, le débit de capacité thermique moyenne ($\overline{DCM}_{T_1-T_2}$) des gaz de combustion entre T₁ et T₂ peut être évalué selon la variation de l'enthalpie de chaque constituant avec la température [Reid, 1977].

D'abord, l'enthalpie (J/kg) des constituants des gaz de combustion en fonction de température s'écrit :

$$HCO_2(T) = \frac{4180}{0.044} (3.01 + 0.942 \cdot 10^{-2} T + 0.190 \cdot 10^{-5} T^2 - 0.210 \cdot 10^{-9} T^3) \quad (2.28)$$

$$HH_2O(T) = \frac{4180}{0.018} (2.16 + 0.667 \cdot 10^{-2} T + 0.188 \cdot 10^{-5} T^2 - 0.168 \cdot 10^{-9} T^3) \quad (2.29)$$

$$HN_2(T) = \frac{4180}{0.028} (2.01 + 0.648 \cdot 10^{-2} T + 0.771 \cdot 10^{-6} T^2 - 0.805 \cdot 10^{-10} T^3) \quad (2.30)$$

$$HO_2(T) = \frac{4180}{0.032} (2.15 + 0.697 \cdot 10^{-2} T + 0.695 \cdot 10^{-6} T^2 - 0.571 \cdot 10^{-10} T^3) \quad (2.31)$$

$$HSO_2(T) = 795 T \quad (2.32)$$

Ensuite, on obtient la variation d'enthalpie des constituants des gaz de combustion de T_1 à T_2 pour 1 kg de combustible :

$$VHCO_2 = [HCO_2(T_1) - HCO_2(T_2)] MCO_2 \quad (2.33)$$

$$VHH_2O = [HH_2O(T_1) - HH_2O(T_2)] MH_2O \quad (2.34)$$

$$VHN_2 = [HN_2(T_1) - HN_2(T_2)] MN_2 \quad (2.35)$$

$$VHO_2 = [HO_2(T_1) - HO_2(T_2)] MO_2 \quad (2.36)$$

$$VHSO_2 = [HSO_2(T_1) - HSO_2(T_2)] MSO_2 \quad (2.37)$$

Et enfin, on obtient le débit de capacité thermique moyenne des gaz de combustion entre T_1 et T_2 :

$$\overline{DCM}_{T_1-T_2} = \dot{M}_{br} \frac{(VHCO_2 + VHH_2O + VHN_2 + VHO_2 + VHSO_2)}{T_1 - T_2} \quad (2.38)$$

2.2.3.3.2 - Température des fumées (régime permanent) à la sortie de l'échangeur

Connaissant maintenant la température des gaz de combustion au niveau du brûleur T_{g1} , la température des gaz de combustion à la sortie de l'échangeur en régime permanent T_{g2} est calculée par la méthode de Nombre d'Unités de Transfert (NUT) à condition que la conductance thermique entre le gaz de combustion et l'eau circulant dans la chaudière et la température de retour d'eau de chaudière soient connues.

En effet, la méthode NUT suppose que T_{g1} et T_{g2} sont reliées à la température de retour d'eau de chaudière T_{r_eau} par l'efficacité de l'échangeur ε_1 , cette dernière étant définie comme le rapport de la puissance thermique réellement échangée à la puissance d'échange maximum théoriquement possible avec les mêmes conditions d'entrée des fluides dans l'échangeur (mêmes fluides, débits et températures d'entrée). Dans le cas d'un échangeur à contre-courant, l'efficacité ε_1 peut s'écrire :

$$\varepsilon_1 = \frac{\Phi_{réel}}{\Phi_{max}} = \frac{\overline{DCM_{T_{g1}-T_{g2}}} (T_{g1} - T_{g2})}{\overline{DCM_{min}} (T_{g1} - T_{r_eau})} \quad (2.39)$$

Comme nous sommes dans le cas où $\overline{DCM_{T_{g1}-T_{g2}}} = \overline{DCM_{min}}$, on a encore :

$$\varepsilon_1 = \frac{(T_{g1} - T_{g2})}{(T_{g1} - T_{r_eau})} \quad (2.40)$$

d'où :

$$T_{g2} = T_{g1} - \varepsilon_1 (T_{g1} - T_{r_eau}) \quad (2.41)$$

En utilisant l'hypothèse de convection thermique pure avec un échangeur à contre-courant, l'efficacité ε_1 peut s'écrire :

$$\varepsilon_1 = \frac{1 - \exp[-NUT(1 - R_{ap})]}{1 - R_{ap} \exp[-NUT(1 - R_{ap})]} \quad (2.42)$$

Dans cette expression, NUT est représentatif du pouvoir d'échange de l'échangeur. Il s'exprime en fonction d'une conductance interne (caractéristique de la chaudière considérée) entre les gaz de combustion et l'eau circulant dans l'échangeur, et du débit de capacité thermique moyenne $\overline{DCM_{T_{g1}-T_{g2}}}$ des gaz de combustion entre T_{g1} et T_{g2} :

$$NUT = K / \overline{DCM_{T_{g1}-T_{g2}}} \quad (2.43)$$

Tandis que R_{ap} représentant le rapport des débits de capacité thermique moyenne des deux fluides fait intervenir le débit d'eau circulant dans l'échangeur et sa chaleur massique :

$$R_{ap} = \frac{\overline{DCM_{T_{g1}-T_{g2}}}}{M_{eau} C_{p_eau}} \quad (2.44)$$

Notons qu'en pratique, la valeur de K n'est pas toujours facile à obtenir ; en revanche les rendements de chaudière sont généralement accessibles chez les constructeurs de chaudière. Pour ce qui nous concerne, la variation de la valeur K sera utilisée pour **modéliser et simuler les défauts d'entartrage, de dépôt des boues dans l'échangeur, et d'encrassement de l'échangeur par dépôt de produits de combustion**. Il faudrait signaler que la valeur référence de K, représentant un échangeur neuf sans entartrage ni encrassement, est déduite selon une efficacité de l'échangeur ε_1 d'environ 0.96.

Allant de l'équation (2.41) à l'équation (2.44), on voit clairement que le calcul de T_{g2} dépend de la valeur de ε_1 . Or pour calculer ε_1 , il faut connaître $\overline{DCM}_{T_{g1} - T_{g2}}$ qui est fonction de T_{g2} . Il s'agit donc un calcul itératif de T_{g2} à partir de la valeur de $\overline{DCM}_{T_{g1} - T_{g2}}$ calculée à des températures T_{g2} estimées.

2.2.3.4 - Conclusion

On a dans ce paragraphe présenté en détail un modèle de chaudière qui permet de simuler deux types de défaut :

- défauts de brûleur en faisant varier la fraction molaire en CO_2 des fumées sèches $[\text{CO}_2]$,
- défauts de l'échangeur en faisant varier la conductance thermique entre les gaz de combustion et l'eau circulant dans l'échangeur K.

Mais il faut signaler :

- que le choix du paramètre $[\text{CO}_2]$ est délicat car sa valeur pour le fonctionnement normal de brûleur est très variée d'une chaudière à une autre,
- que le choix du paramètre K est très difficile car sa valeur est difficilement accessible chez les constructeurs d'une part, et évolue au cours du fonctionnement de la chaudière d'autre part.

2.3 - CONSTRUCTION D'UN SIMULATEUR DU SYSTEME DE CHAUFFAGE

Une fois que les modèles ont été établis, il faut les traduire dans l'environnement SIMULINK afin de construire un simulateur du système de chauffage permettant de simuler certains défauts de fonctionnement.

Dans la plupart des modèles, il s'agit de résoudre des équations différentielles ; avec SIMULINK on emploiera la formulation des équations d'état (par exemple les formules (2.12) et (2.13)) aussi souvent que possible.

Avant de décrire le simulateur du système, on présente succinctement les logiciels MATLAB et SIMULINK, [Matlab, 1992] et [Simulink, 1992].

2.3.1 - PRESENTATION DES LOGICIELS MATLAB ET SIMULINK

MATLAB est un langage de programmation évolué destiné à la résolution des problèmes mathématiques et physiques. L'origine de MATLAB, MATrix LABoratory, réside dans le désir de rendre le calcul matriciel, généralement long et fastidieux à programmer, plus simple d'emploi. En effet, ceci se fait très aisément car le type de variable utilisé sous MATLAB est la matrice.

SIMULINK est quant à lui une extension de MATLAB. Ce logiciel garde toutes les propriétés de MATLAB auquel a été ajouté un environnement de simulation. Cet environnement utilise une interface graphique très évoluée pour la configuration, la définition et le suivi d'une simulation. Cette interface est basée sur une programmation sous forme de schéma bloc : des blocs élémentaires (objets) sont configurés, paramétrés et reliés entre eux par leurs entrées/sorties afin de créer un simulateur. Les principaux schémas-blocs de SIMULINK se présentent ci-dessous (Figures 2.2 - 2.5):

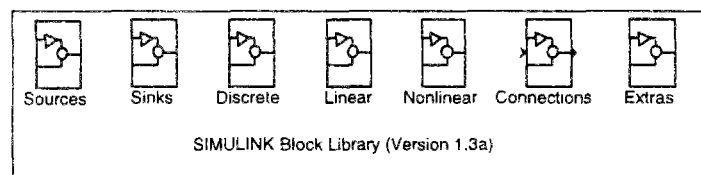


Figure 2.2 *Librairie des schémas-bloc SIMULINK*

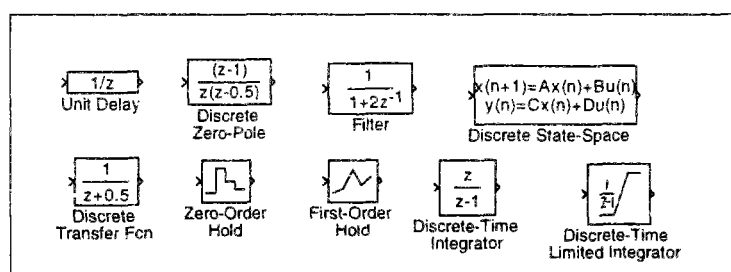


Figure 2.3 *Librairie des fonctions discrètes*

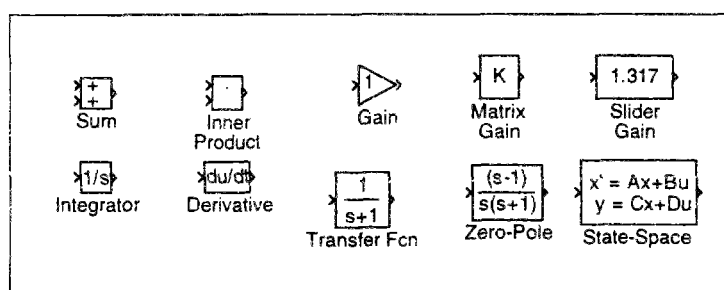


Figure 2.4 *Librairie des fonctions linéaires*

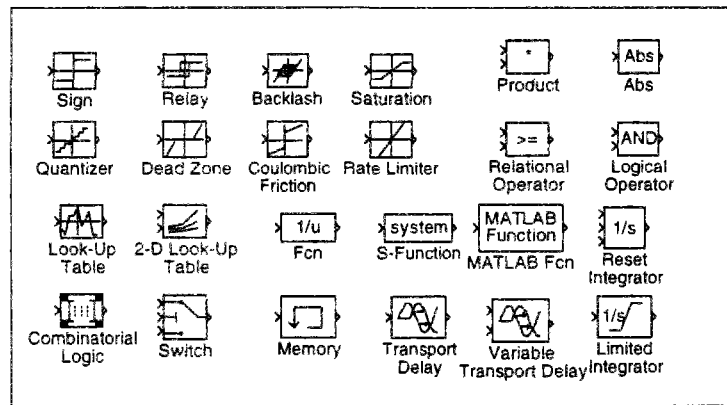


Figure 2.5 Librairie des fonctions non linéaires

SIMULINK permet donc de résoudre les systèmes décrits sous forme d'équations différentielles discrètes, linéaires et non linéaires. Il dispose également de nombreuses fonctions non linéaires et instructions discrètes permettant de simuler un grand nombre de processus.

L'évolution d'une simulation peut être visualisée au cours de simulation, ou alors le résultat final d'une séance de simulation se trouve dans l'espace de travail (Workspace en anglais) de MATLAB. On peut ainsi faire subir à ces vecteurs ou matrices toutes les opérations nécessaires au traitement des données.

On présentera dans le paragraphe suivant un simulateur du système de chauffage construit avec ces schémas bloc SIMULINK.

2.3.2 - ELABORATION D'UN SIMULATEUR DU SYSTEME DE CHAUFFAGE

L'ensemble des modèles du système présentés dans le paragraphe 2.2 ont été traduits en SIMULINK. Ils vont maintenant servir à monter un simulateur. Le montage se fait en reliant les différents modèles ; chaque modèle, qui représente un composant ou un sous-ensemble du système, est regroupé pour la clarté du graphisme. Nous avons tenté par souci de convivialité de rapprocher le plus possible la présentation du simulateur d'un système réel.

L'allure du simulateur est donnée par la Figure 2.6 :

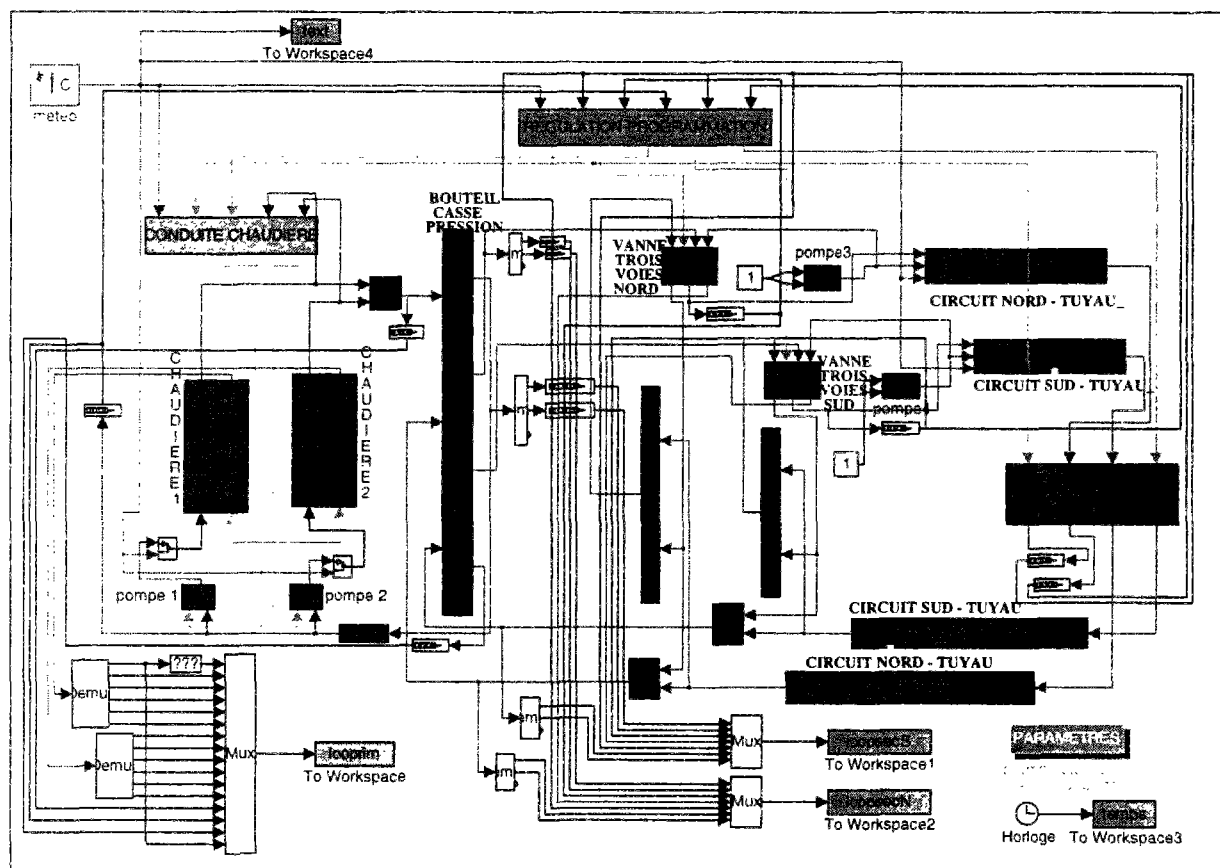


Figure 2.6 Simulateur du système de chauffage

Ce simulateur comprend principalement :

- le bloc METEO qui fournit, en cours de simulation, la température extérieure à différents modèles (équipements ou systèmes de contrôle) du système de chauffage,
- le bloc REGULATION-PROGRAMMATION qui représente le programmeur d'intermittence et la régulation des températures de départ des circuits nord et sud,
- le bloc CONDUITE CHAUDIERE qui représente la conduite en séquence (cascade) de deux chaudières à gaz,
- deux blocs CHAUDIERE 1 et CHAUDIERE 2 qui se situent dans la boucle primaire,
- le bloc BOUTEILLE-CASSE-PRESSION qui sépare la boucle primaire de la boucle secondaire,
- la boucle secondaire comprend un BATIMENT 2 ZONES, deux VANNES-TROIS-VOIES (nord et sud), des TUYAUX, deux POMPES de circulation,

- le bloc PARAMETRES qui initialise les paramètres des modèles, et définit la durée (i.e. début - fin) de simulation, le pas de simulation, etc.
- le bloc HORLOGE qui compte le temps s'écoulant au cours de simulation,
- des blocs TO WORKSPACE pour le stockage de données dans l'espace de travail de MATLAB.

La plupart des blocs se décomposent encore en sous-blocs.

Prenons, comme exemple, le bloc CHAUDIERE 1. Quand on clique avec une souris deux fois sur ce bloc, on voit apparaître sur l'écran un schéma bloc du modèle de chaudière à gaz (Figure 2.7). Là-encore, certains blocs du schéma se décomposent, ...

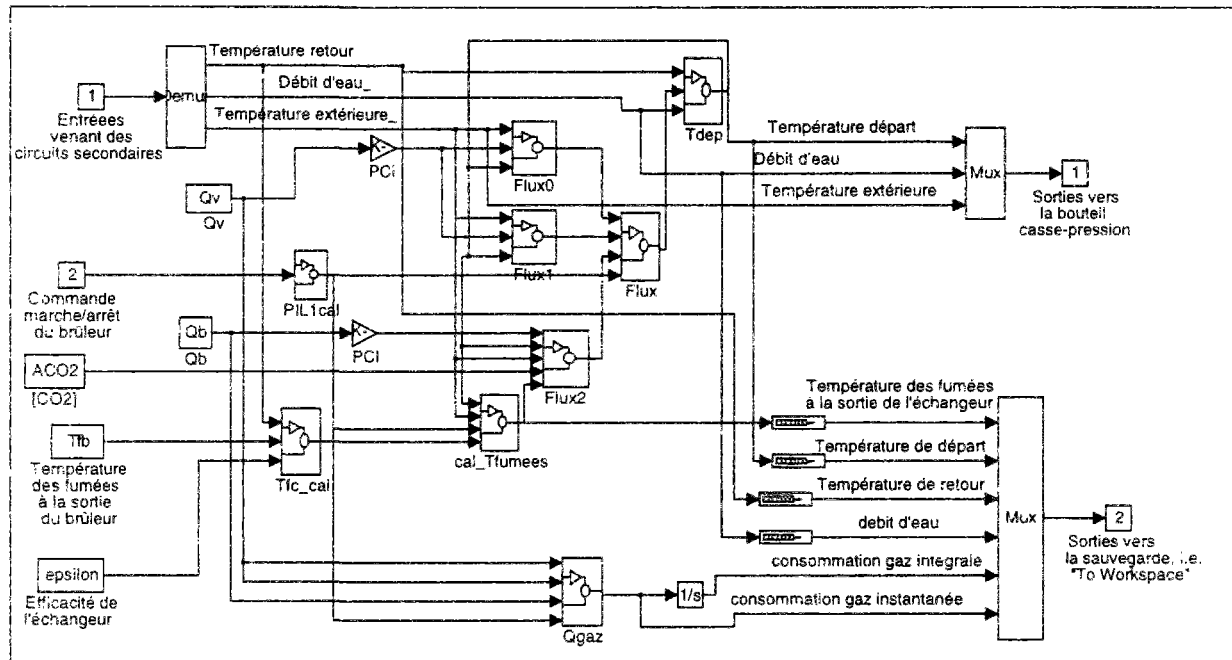


Figure 2.7 *Modèle de chaudière à gaz*

2.4 - SIMULATION DU SYSTEME DE CHAUFFAGE SANS ET AVEC DEFAUTS

2.4.1 - MODELISATION DES DEFAUTS A SIMULER

Les défauts choisis pour la génération d'une base de données destinée au développement de méthodes de détection de défauts sont issus des résultats de l'enquête décrits dans le chapitre 1. Les défauts de conception (défauts n'évoluant pas au cours du temps) ne sont pas simulés dans un premier temps. Notons toutefois que notre objectif est de définir des méthodes utilisables pour les différents défauts d'exploitation constatés, ces méthodes doivent donc rester applicables en présence de défauts de conception les plus courants. Par conséquent il faudra bien vérifier la robustesse des méthodes de détection développées en présence de certains défauts importants de conception.

Les défauts⁽³⁾ pris en compte dans la simulation sont au nombre de 6 ; leurs modélisations sont brièvement décrites ci-dessous :

- mauvaise combustion du **brûleur**. La mauvaise combustion comprend plusieurs aspects : excès d'air trop important, défaut d'air, variation importante de PCI (Pouvoir Calorifique Inférieur), etc.. Pour commencer, nous avons choisi le défaut de l'excès d'air trop important pour la simulation de ce type de défaut. La simulation de ce défaut se fait par un changement du taux d'excès d'air de 40 % (cas sans défaut : correspondant à un taux de CO₂ de 8 %) à 85 % (cas avec défaut : correspondant à un taux de CO₂ de 6.5 %).
- encrassement et entartrage de l'**échangeur** principal de chaudière. La simulation de ce défaut se fait par un changement du coefficient d'échange entre les fumées et l'eau circulant dans la chaudière de 185 W/°C (cas sans défaut : correspondant à une efficacité de l'échangeur de 96 %) à 130 W/°C (cas avec défaut : correspondant à une efficacité de l'échangeur de 89 %).
- **courbe de chauffe** trop basse. On n'a simulé dans un premier temps que le défaut de courbe de chauffe trop basse pour les deux zones (nord et sud) en faisant diminuer parallèlement de 10°C la température de consigne d'eau à la sortie de vanne-trois-voies (circuits nord et sud) par rapport à la valeur de référence calculée. Notons que le modèle de courbe de chauffe utilisé dans la simulation est basé sur le régime permanent ; or cette hypothèse n'est pas toujours valide car notre système de chauffage est en régime non-permanent, [Fabron, 1995]. Cela augmentera, on le verra plus loin, la difficulté de discrimination de défauts liés à la courbe de chauffe.
- relance précoce et relance tardive calculées par programmeur d'intermittence. La simulation se fait en augmentant (pour le défaut de relance précoce) ou diminuant (pour le défaut de relance tardive) dans le module de l'optimiseur la durée maximum de préchauffage. La durée maximum de préchauffage est calculée pour une température extérieure de base par exemple de -7°C et une température intérieure minimum par exemple de 14°C.

Que ce soit sur le terrain ou à l'aide d'un simulateur numérique, cette valeur est obtenue par essais successifs et est généralement différente suivant le bâtiment. En ce qui concerne notre système de chauffage simulé, nous avons obtenu, après plusieurs essais sur le simulateur numérique et avec un compromis vis-à-vis de différents jours de la semaine, une valeur de 6 heures qui représente la durée maximum de préchauffage correctement calculée. La relance précoce a été donc simulée par une durée maximum de préchauffage de 12 heures et la relance tardive par une durée maximum de préchauffage de 3 heures.

Notons qu'avec cette valeur de 6 heures, il est possible que le système relance, précocement ou tardivement, le chauffage matinal. Par conséquent la détection du fonctionnement normal est parfois difficile à cause de ces imperfections du programmeur utilisé dans la simulation.

⁽³⁾ On simule dans un premier temps des défauts "francs", c'est-à-dire une variation grande de paramètre physique qui représente un défaut quelconque, afin de bien mettre en évidence les différents types de défauts.

On doit signaler que ces imperfections du programmeur d'intermittence représentent bien la réalité, c'est-à-dire que les programmeurs existants dans le marché de Télégestion et de Gestion Technique du Bâtiment sont généralement imparfaits quel que soit le niveau de programmation appliqué, et qu'ils ont souvent du mal d'une part à bien contrôler le relance matinale de chauffage et d'autre part à réguler la courbe de chauffe [Visier, 1994]. C'est la raison pour laquelle on utilise dans notre simulation un programmeur non-parfait mais fidèle à la réalité, ce qui nous permettra de voir comment un programmeur réel influe sur la détection diagnostic des défauts et, par conséquent, de pouvoir plus facilement améliorer la performance du programmeur.

- **vanne-trois-voies** non étanche en position fermée (fuite de la voie directe vers bypass ou course de limiteur mal positionnée). Ce défaut a été simulé par une ouverture de la vanne-trois-voies du circuit nord de 10% de l'ouverture totale au moment où la vanne doit être complètement fermée (chauffage en arrêt).

2.4.2 - SIMULATION DU FONCTIONNEMENT DU SYSTEME COMPORTANT DES DEFAUTS

2.4.2.1 - Présentation des classes ou modes de fonctionnement

En fonction de la description des défauts à simuler, les résultats de simulation ont été structurés sous forme de classe (terme employé en reconnaissance des formes, voir le chapitre 3), chaque classe représentant un mode (état) de fonctionnement du système. Au total 7 classes ont été simulées ; la durée de simulation de chacune de ces classes est une saison de chauffage entière, qui représente dans notre simulation 172 jours.

Ces classes sont présentées ci-dessous :

- classe **normale** correspondant à un mode de fonctionnement normal de référence,
- classe **brûleur** correspondant au mauvais fonctionnement (mauvaise combustion) du brûleur,
- classe **échangeur** correspondant à l'entartrage ou l'encrassement de l'échangeur d'une chaudière,
- classe **courbe de chauffe** correspondant à la courbe de chauffe trop basse réglée par un régulateur en fonction de la température extérieure,
- classe **relance précoce** correspondant au mauvais calcul de l'heure de relance par un programmeur d'intermittence,
- classe **relance tardive** correspondant également au mauvais calcul de l'heure de relance par un programmeur d'intermittence,
- classe **vanne de régulation** correspondant à un défaut de vanne-trois-voies qui régule la température de départ du circuit secondaire, par exemple fuite de vanne en position fermée.

2.4.2.2 - Présentation graphique de données de simulation

Les simulations se sont effectuées sous l'environnement SIMULINK. Leurs résultats (variables) ont été stockés dans l'espace de travail (workspace en anglais) MATLAB sous forme *fichier.mat*. Pour chaque classe simulée, l'ensemble des fichiers comporte au total 54 variables (voir Annexe A3).

L'objectif de stocker un nombre si important de variables est double :

- d'une part, le fonctionnement du système (chauffage intermittent, régulation en fonction de température extérieure, couplage bâtiment-système, etc.) étant complexe, les comportements dynamiques de ces variables sont très utiles pour vérifier et justifier la crédibilité du simulateur développé,
- d'autre part, l'ensemble des variables constitue une base de données représentant le fonctionnement du système comportant différents défauts. Cette base de données devrait nous permettre de ressortir des variables portant les informations de défauts les plus significatives et les plus pertinentes, qui seront utilisées pour développer des méthodes de détection et de diagnostic.

Bien que les variables stockées soient nombreuses, le nombre des variables portant les signatures de différents défauts est malheureusement très réduit. On sait a priori que certaines variables portent des indications très significatives d'un de tel défaut, mais les autres n'en contiennent aucune indication. La recherche de ces variables significatives et pertinentes s'appuie essentiellement sur deux approches : analyse physique du système, méthodes mathématiques de traitement de données (analyse en composantes principales, etc.).

Pour mettre en évidence cet aspect, une comparaison de deux variables (température des fumées de chaudière et température intérieure) entre la classe normale et la classe brûleur est présentée par les Figures 2.8 - 2.9 et 2.10 - 2.11. On remarque facilement que la température des fumées indique bien un défaut (mauvaise combustion) de brûleur (Figures 2.8, 2.10) ; en revanche, la température intérieure n'en porte pratiquement aucune signature (Figures 2.9, 2.11).

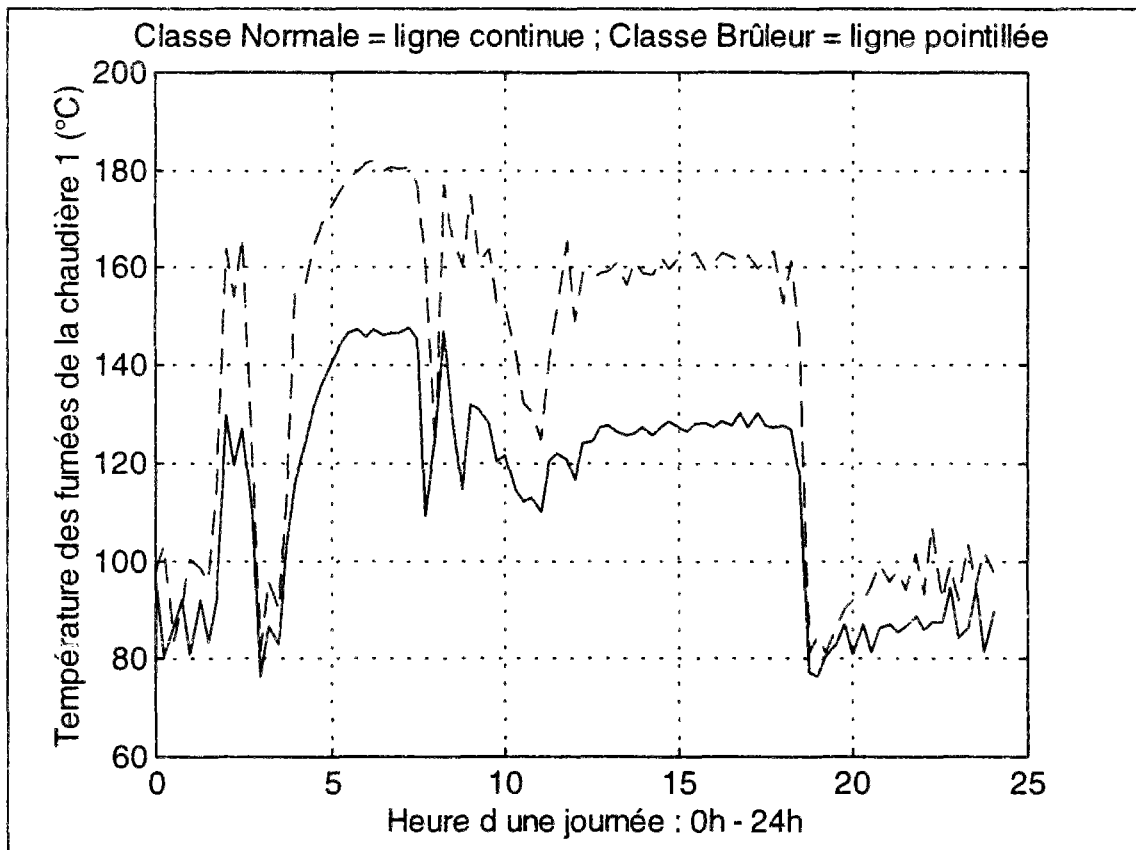


Figure 2.8 Température des fumées -
Comparaison entre classe normale et classe brûleur (1 journée)

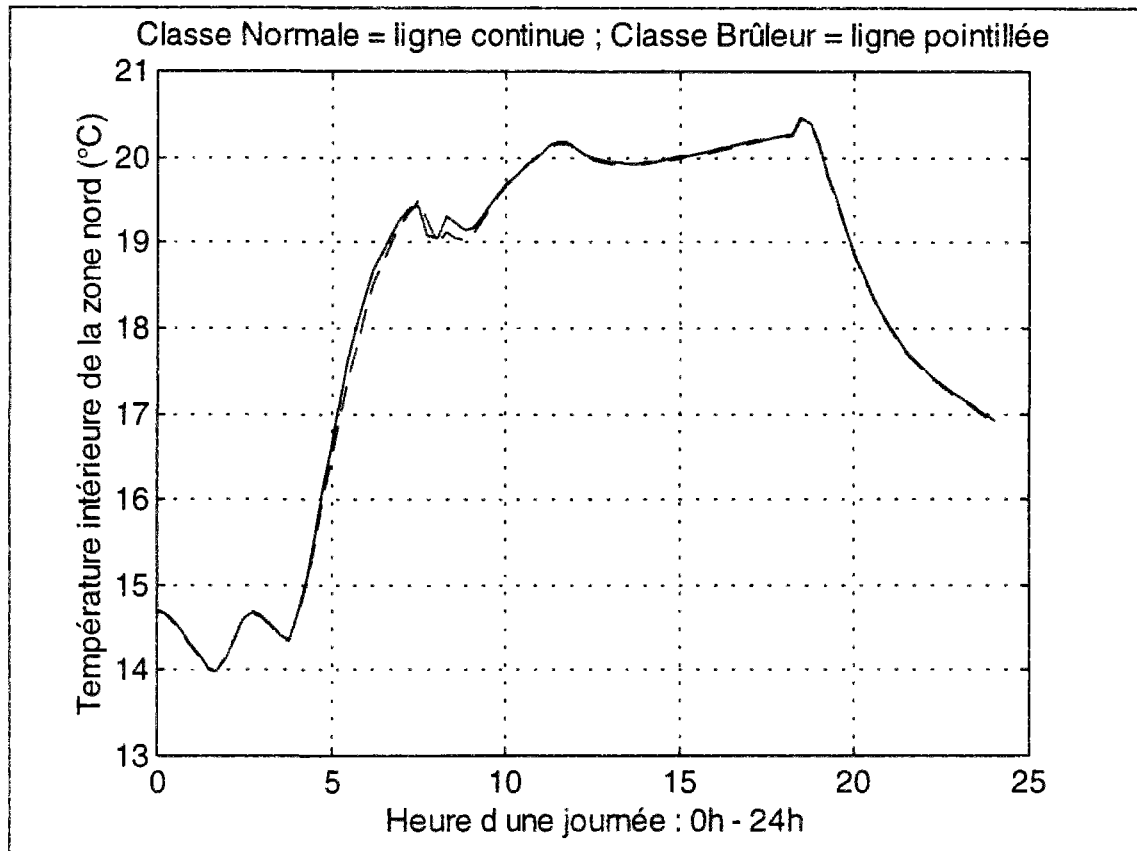


*Figure 2.9 Température intérieure -
Comparaison entre classe normale et classe brûleur (1 journée)*

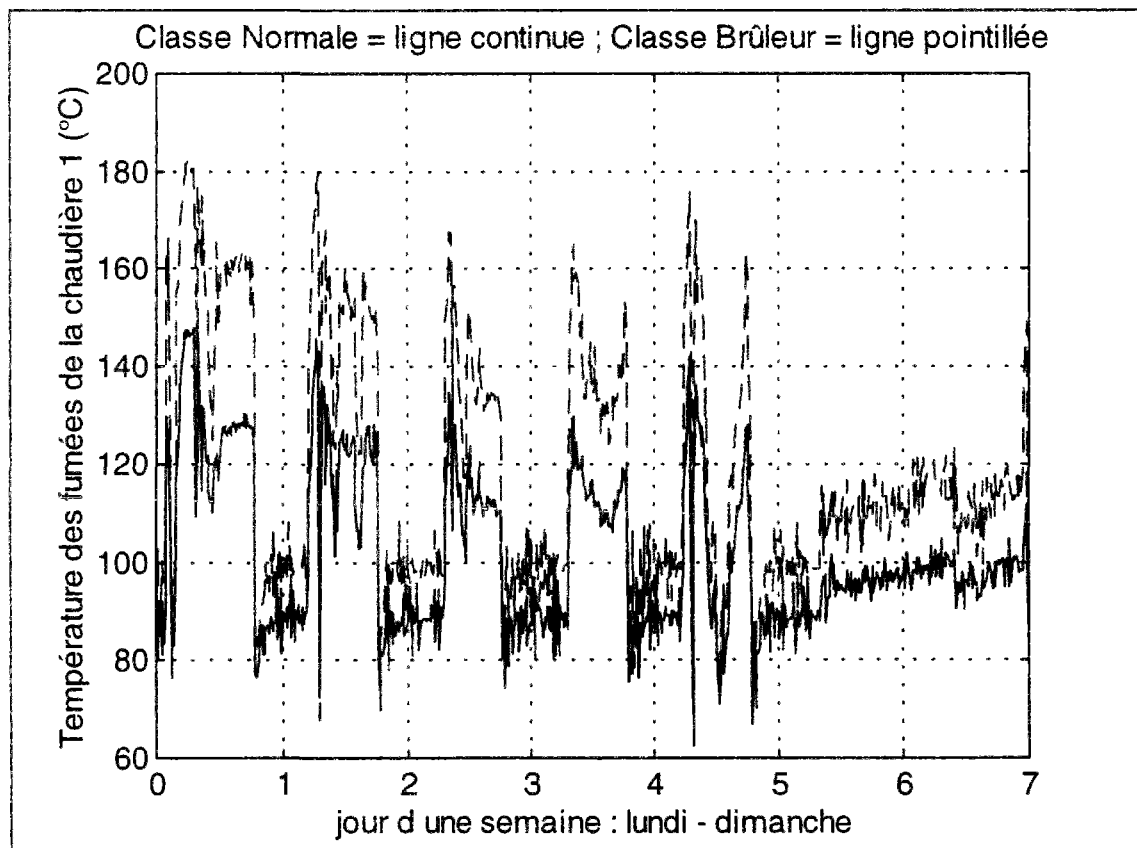


*Figure 2.10 Température des fumées -
Comparaison entre classe normale et classe brûleur (1 semaine)*

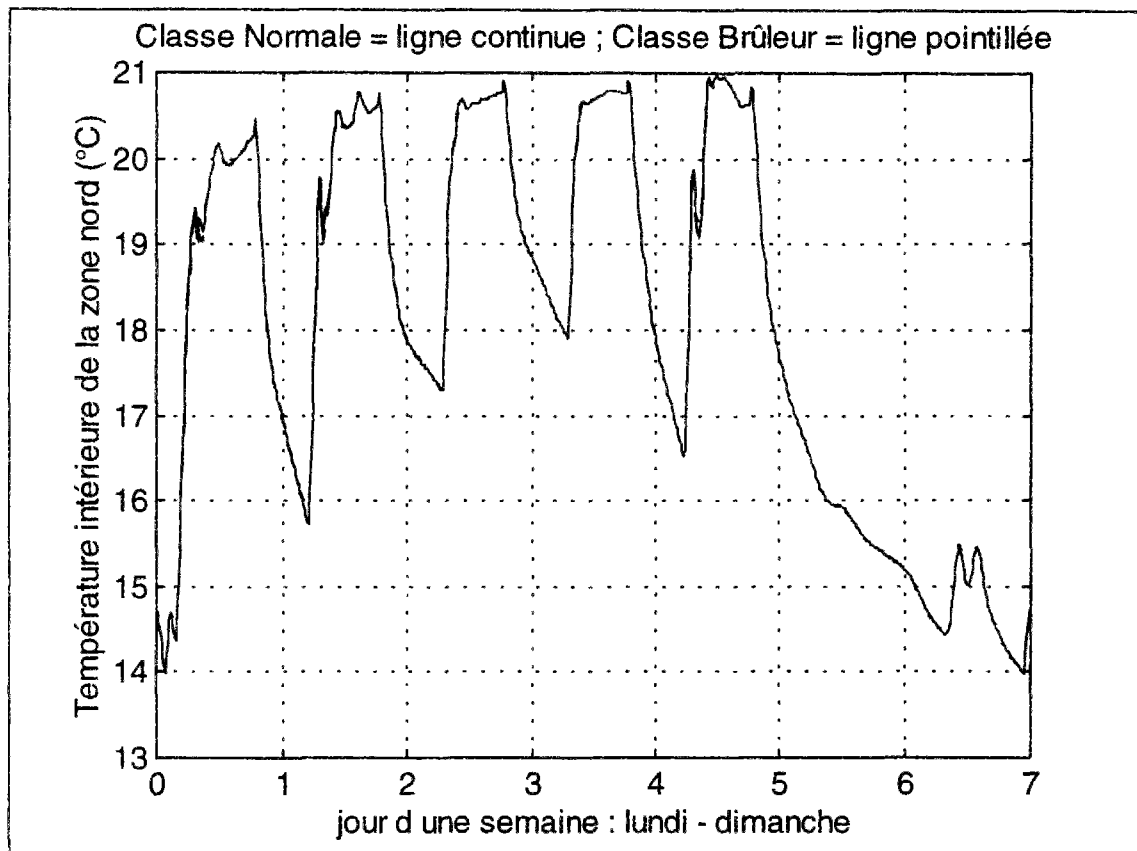


*Figure 2.11 Température intérieure -
Comparaison entre classe normale et classe brûleur (1 semaine)*

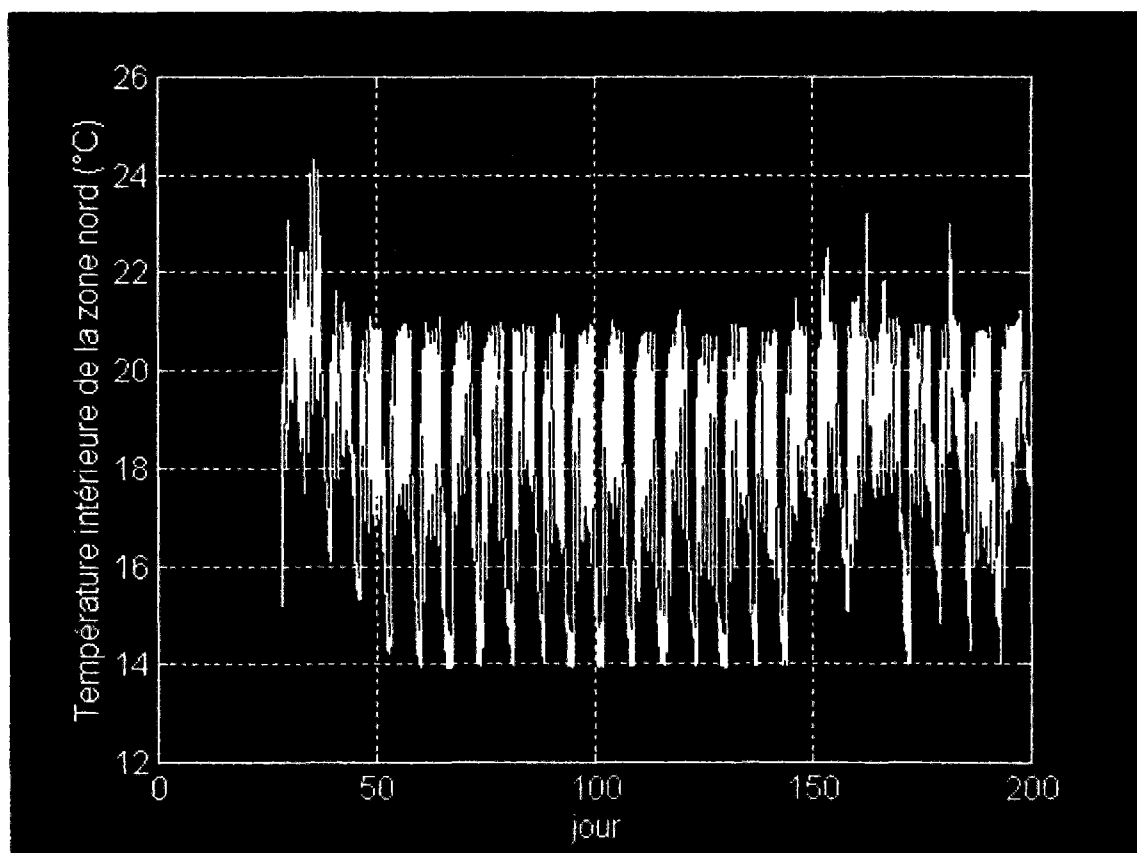


Figure 2.12 Température intérieure de classe normale (une saison de chauffage : 172 jours)

De plus, les Figures 2.8-2.9 montrent les différentes étapes d'une gestion de chauffage intermittent durant une journée entière :

- de 0h à environ 3h40 : arrêt ou chauffage réduit,
- d'environ 3h40 à 8h : relance matinale de chauffage,
- de 8h à 18h : chauffage normal,
- de 18h à 24h : arrêt.

Et les Figures 2.11 et 2.12 montrent respectivement la variation de la température intérieure d'une semaine et d'une saison de chauffage complète. On voit clairement :

- que la température intérieure pendant les périodes d'occupation est située entre 19 et 21°C (effet des deux régulations : régulateur en fonction de température extérieure et vanne thermostatique dans la pièce), excepté quelques jours où la température extérieure est trop élevée, ou les apports solaires sont trop importants (Figure 2.12),
- que le chauffage est arrêté ou en régime réduit pendant les périodes d'inoccupation (nuit, week-end),
- que le seuil bas de température intérieure est de 14°C, qui correspond bien à la consigne donnée dans notre simulation.

CHAPITRE 3

METHODE CONNEXIONNISTE POUR LE DIAGNOSTIC

Dans ce chapitre, nous nous attachons d'abord à rappeler les concepts de base de la Reconnaissance des Formes statistique qui font d'elle une approche puissante pour le diagnostic. Nous présentons ensuite la méthode connexionniste - Réseaux de Neurones Artificiels (RNA) en tant qu'outil pour la reconnaissance des formes, en accentuant sur une architecture particulière - un réseau Perceptron Multi-Couches (PMC).

3.1 - Reconnaissance des formes statistique et diagnostic.....	56
3.1.1 - Généralité sur la reconnaissance des formes statistique.....	56
3.1.1.1 - Phase d'analyse.....	57
3.1.1.2 - Phase décisionnelle	58
3.1.2 - Diagnostic de défauts automatisé	58
3.1.3 - Concept de rejet	60
3.1.3.1 - Règle de Bayes	61
3.1.3.2 - Rejet d'ambiguïté et rejet de distance.....	61
3.1.4 - Prétraitement des données	62
3.1.4.1 - Sélection et extraction des paramètres/caractères.....	63
3.1.4.2 - Création d'une partition de l'espace de représentation	64
3.2 - Méthode connexionniste en tant qu'outil pour la reconnaissance des formes	65
3.2.1 - Introduction.....	65
3.2.2 - Définition d'un système connexionniste	65
3.2.2.1 - Structure.....	66
3.2.2.2 - Dynamique	67
3.2.2.3 - Apprentissage	69
3.2.3 - Un modèle connexionniste - réseaux perceptron multi-couches	70
3.2.3.1 - Présentation du modèle	70
3.2.3.2 - Mise en oeuvre pratique.....	76

3.1 - RECONNAISSANCE DES FORMES STATISTIQUE ET DIAGNOSTIC

3.1.1 - GENERALITE SUR LA RECONNAISSANCE DES FORMES STATISTIQUE

On accorde aux êtres vivants, qualifiés d'évolués, la faculté de reconnaître un objet, une situation, et de réagir en fonction de la classification qui en a été faite. Cette classification a pu être effectuée grâce à une association entre la perception d'une part et un concept général issu d'expériences passées d'autre part. La Reconnaissance des Formes statistique (RdF), dont la naissance et l'essor sont liés au développement des calculateurs modernes, peut être vue comme une famille de méthodes permettant l'automatisation du processus de reconnaissance.

En effet, la Reconnaissance des Formes est considérée comme une science de définition d'algorithmes permettant de classer des objets dont l'aspect a varié par rapport à un objet type, ou en d'autres termes, de pouvoir décider, après avoir observé un objet, à quelle classe d'objets connus il peut être associé. Traditionnellement on distingue deux types de reconnaissance des formes : la reconnaissance structurelle qui s'appuie sur une représentation des formes à l'aide de grammaires et la reconnaissance statistique qui utilise une représentation purement numérique des formes.

L'approche reconnaissance des formes statistique a été appliquée avec succès depuis longtemps dans de nombreux domaines : reconnaissance de caractères, traitement d'images, traitement de la parole, analyse de scènes, etc.. Ayant entre autres la capacité d'apprentissage, elle apparaît également comme un outil puissant de détection et de diagnostic de défauts.

En reconnaissance des formes statistique, on définit la forme par un ensemble de **paramètres**, appelés aussi **caractères**. Constituons un espace à l'aide d'une base dont chaque élément est associé à un paramètre/caractère. Cet espace est appelé **espace de représentation** ou **espace des formes**, noté R^d dans le cas où les paramètres peuvent prendre toute valeur réelle. Une **forme** est donc représentée par un point dans l'espace de dimension d . Dans la suite, nous désignerons une forme par un vecteur noté $\vec{x}$ (d composantes : $x_1, x_2, \dots, x_i, \dots, x_d$) ; x_i représente la $i^{\text{ème}}$ composante ou valeur du $i^{\text{ème}}$ paramètre pour cette forme.

Les prototypes sont donc des points représentatifs dans cet espace R^d et le problème de reconnaissance des formes est d'associer un point quelconque de cet espace à un des prototypes. Si aucune perturbation ne marquait une nouvelle forme observée, celle-ci se confondrait dans R^d avec un prototype et le problème de reconnaissance deviendrait trivial. A cause du bruit, à chaque forme type est donc associée une zone géométrique ; à chaque zone, on associe le nom de **classe**. Le problème de reconnaissance des formes est donc l'obtention de frontières entre les M classes notées $\omega_1, \omega_2, \dots, \omega_M$ (Figure 3.1). Suivant sa position par rapport à l'ensemble des frontières, une nouvelle forme observée c'est-à-dire un nouveau point $\vec{x}$ apparu dans R^d sera associée à une des M classes parmi $\omega_1, \omega_2, \dots, \omega_M$, c'est l'opération de **classification** ou de **discrimination**.

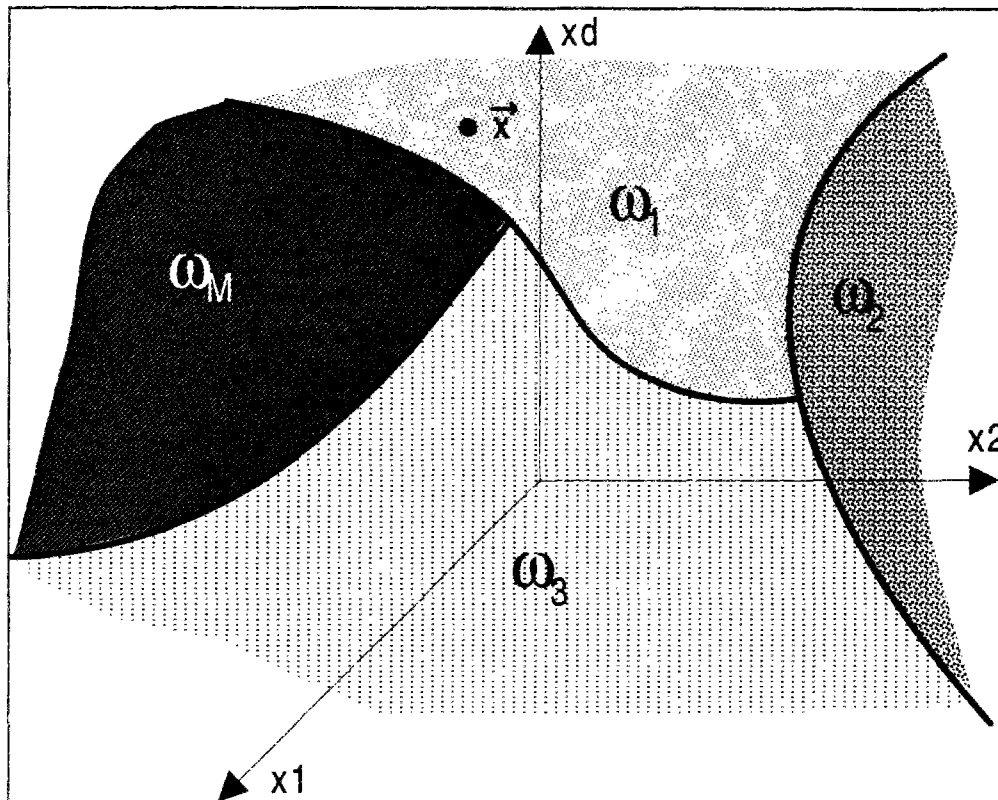


*Figure 3.1 Principe d'un problème de reconnaissance des formes :
à quelle classe associe l'observation $\bar{x}$*

La mise en oeuvre d'une stratégie de reconnaissance des formes se décompose en deux phases distinctes : une phase d'analyse et une phase décisionnelle.

3.1.1.1 - Phase d'analyse

La phase d'analyse, à partir de mesures sur un système réel ou de simulation, comprend trois aspects différents.

D'abord le choix d'un espace de représentation. Il s'agit de construire un vecteur de dimension d - vecteur forme, dont les composantes sont issues d'une extraction ou sélection de d paramètres pertinents (voir paragraphe 3.1.4.1).

Ensuite la création d'une partition de l'espace. Cela consiste à agréger les différentes formes en M classes $\omega_1, \omega_2, \dots, \omega_M$ (voir paragraphe 3.1.4.2).

Enfin le choix d'une méthode de classification et l'élaboration de frontières entre M classes - un **discriminateur** ("classifier" en anglais) pour la reconnaissance de formes. Cela suppose qu'on ait donné une description complète et précise de chacune des classes dans l'espace de représentation (ou des formes) - un ensemble des données numériques par exemple ; cette opération de matérialisation du concept dans l'espace des formes est appelée **apprentissage** (voir aussi paragraphe 3.2.2.3). Celui-ci est réalisé en mémorisant le concept par une structuration de l'espace des formes.

Un autre ensemble des données, dont la statistique est la même que celle de l'ensemble d'apprentissage, est utilisé pour tester la qualité de la règle de décision en termes d'erreur de classification ; cette erreur doit être en général estimée, le calcul en erreur sera difficile.

Notons que la bonne classification n'est pas toujours garantie en dehors de l'ensemble d'apprentissage. La reconnaissance des formes pose de ce point de vue un problème analogue à celui de l'interpolation, et surtout de l'extrapolation. Le choix d'un ensemble d'apprentissage représentatif, donc exhaustif, n'est pas toujours aisé. Lorsque le nombre de classes est connu a priori (reconnaissance de caractères alphanumériques, etc.), le problème est simple. Le cas contraire, un problème de diagnostic par exemple, demeure délicat.

Deux grandes familles de méthodes de classification se distinguent : les **méthodes paramétriques** supposant connues (ou estimées) les lois de probabilité des différentes classes, c'est par exemple la règle de Bayes, [Tou 1974] ; les **méthodes non-paramétriques** n'impliquent aucune hypothèse sur les statistiques des classes, citons par exemple la règle de k plus proches voisins, [Fukunaga 1972], les réseaux de neurones, [Hertz 1991]. On précisera cette dernière méthode dans le paragraphe 3.2.

3.1.1.2 - Phase décisionnelle

S'appuyant sur le discriminateur (ou l'opérateur de reconnaissance) établi lors de l'apprentissage, la phase décisionnelle, appelée aussi la phase de reconnaissance, [Pau 1975], s'attache à l'affectation d'une nouvelle forme dans l'une des classes connues dont les représentants de celle-ci lui sont "les plus semblables".

A la décision de reconnaissance d'une forme sont attachés d'une part le risque de classer cette forme dans une classe qui ne serait pas la bonne, d'autre part le risque de rejeter cette forme si elle est suffisamment dissemblable de toutes les classes d'apprentissage. On appelle **taux de bonne classification/reconnaissance**, dans une classe donnée, la probabilité pour qu'une forme quelconque appartenant à cette classe, soit correctement classée dans celle-ci.

3.1.2 - DIAGNOSTIC DE DEFAUTS AUTOMATISE

De nombreuses définitions sont proposées pour le terme de **diagnostic**. Celle du dictionnaire Robert est rigoureusement étymologique : "action de déterminer une maladie d'après ses symptômes".

On pourrait dans le domaine médical concevoir de réaliser des diagnostics basés exclusivement sur les réponses des malades aux interrogatoires, et aux résultats de certaines analyses, sans examen clinique systématique nécessitant le recours à un médecin.

De la même manière, nous sommes amenés à envisager dans le domaine technique des diagnostics basés sur l'analyse statistique de toutes les informations disponibles sur le fonctionnement et l'environnement d'un système jusqu'à l'instant de la panne compris. Cette analyse statistique sera automatisée, afin que le démontage du système par un technicien qualifié ne soit réalisé qu'à titre de recours ultime au stade de l'identification de la cause de la panne. Ayant déterminé automatiquement la ou les causes suspectées, on se contentera de démontrer le ou les sous-ensembles probablement avariés.

En général, un système présente plusieurs **modes de fonctionnement**. On peut observer des modes de différents types [Dubuisson, 1990 a] :

- (1) modes de fonctionnement normaux qui, s'ils ne sont pas nuisibles, ne correspondent pas forcément tous à ce que l'on attend du point de vue de la qualité de la mission accomplie. Parmi eux, il y a le mode de fonctionnement nominal qui doit être, en principe, en parfaite adéquation avec la qualité de la mission.

(2) modes de fonctionnement anormaux. Lorsque le système est gouverné par un de ces modes, la mission peut être partiellement remplie ou non remplie ; en tout état de cause, le service n'est pas satisfaisant, il peut aussi être critique. Ces modes peuvent d'ailleurs être décomposés en :

- modes interdits : ce sont des modes sous lesquels le système ne doit absolument pas fonctionner pour des raisons de sécurité, danger, etc.,
- modes défaillants qui correspondent à des mauvais fonctionnements du système. Un sous-ensemble de ces modes est constitué des modes de pannes où un composant et/ou un sous-ensemble du système, voire le système complet ne fonctionne plus,
- modes dégradés qui correspondent soit à l'accomplissement partiel de la mission, soit à l'accomplissement de celle-ci avec des performances moindres,
- modes critiques pour lesquels le système présente des caractéristiques de fonctionnement très particulières et souvent non souhaitées.

(3) modes de fonctionnement d'exception qui peuvent être normaux ou anormaux, mais peu tolérés et/ou peu fréquents.

(4) modes d'indisponibilité sous lesquels le système ne peut accomplir sa mission.

(5) modes de fonctionnement évolutifs qui caractérisent des modes, souvent transitoires, de passage d'un mode à un autre, quelle que soit la qualification du mode de départ ou du mode d'arrivée (normal ou anormal).

Bien évidemment, tous les systèmes ne présentent pas tous les modes de fonctionnement : cette classification qui n'est d'ailleurs pas une partition permet de couvrir pratiquement tous les modes susceptibles de se produire et de s'adapter à la variété des systèmes rencontrés.

Le **diagnostic automatisé** consiste à reconnaître des causes/modes de panne probables d'un système à partir des symptômes observés et des antécédents. La liaison entre le diagnostic de défauts et la reconnaissance des formes paraît évidente car la notion de mode de fonctionnement s'apparente clairement à celle de classe en reconnaissance des formes (voir paragraphe 3.1.1). Cette liaison est schématisée par la Figure 3.2. Les d paramètres, issus d'un traitement de mesures recueillies grâce à différents capteurs, vont eux servir à construire des vecteurs ou formes-symptômes. Il va de soi que ces paramètres sont supposés être pertinents vis-à-vis du problème posé (voir paragraphe 3.1.4.1).

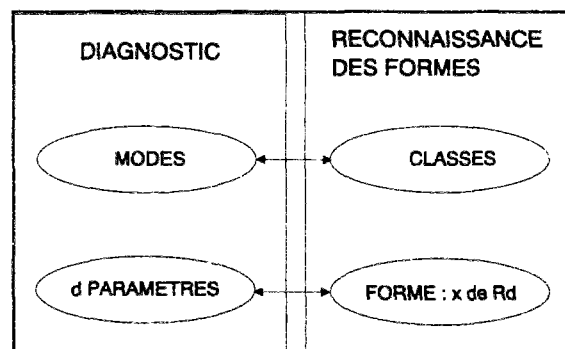


Figure 3.2 *Diagnostic et reconnaissance des formes*

Formuler un diagnostic automatisé, c'est alors prendre la décision de reconnaître automatiquement la forme-symptômes dans l'ensemble des formes observables, classifiées par classes de mode de pannes ou de maladies. Cette reconnaissance est immédiatement interprétée en termes de mode de fonctionnement du système.

Nous avons signalé plus haut (paragraphe 3.1.1.1) que la classification/reconnaissance supposait un apprentissage préalable grâce à une base des données disponible (historique du système, simulation, ...). On constate que si l'apprentissage des règles de décision peut nécessiter un temps de travail important, la classification d'une nouvelle forme-symptôme lors de l'exploitation est fréquemment exécutée en temps réel.

Il est également utile d'évoquer que le schéma présenté ci-dessus (Figure 3.2) pourrait être insuffisant pour une application de la Reconnaissance des Formes (RdF) au diagnostic, à cause de la non-exhaustivité de la reconnaissance des modes de fonctionnement du système au départ. C'est par exemple le cas d'un système évoluant dans le temps ou lorsque seules des données correspondant à un état normal sont disponibles, etc.. Les méthodes de RdF, où le nombre de classes est généralement fixé, doivent donc être étendues pour prendre en compte ce problème. C'est pour cela qu'a été conçu le concept de rejet, [Chow, 1957] et [Dubuisson, 1990 b].

3.1.3 - CONCEPT DE REJET

Chow introduisit pour la première fois le concept de rejet en 1957, [Chow, 1957]. Une trentaine d'années plus tard, Dubuisson, [Dubuisson, 1990 b], montre que ce dernier ne s'intéresse qu'au problème des points ambigus situés près de la frontière de décision (Figure 3.3 : cas 1) : en d'autres termes, les zones de rejet ne correspondent qu'à des régions sises entre les classes, là où plusieurs décisions pourraient être prises.

Cette option de rejet permet au système de différer la décision lorsque l'information pour classer un individu n'est pas suffisante. En revanche elle devient insuffisante vis-à-vis du problème de la non exhaustivité des modes d'apprentissage. C'est pourquoi il est nécessaire d'employer une nouvelle notion de rejet, introduite par Dubuisson et qualifiée de rejet de distance, par opposition au rejet de Chow, qualifié de rejet d'ambiguïté. Ce nouveau concept s'applique donc aux points situés loin de toutes zones d'apprentissage (Figure 3.3 : cas 2), et qui, sans rejet, seraient affectés de manière erronée à l'une des classes connues.

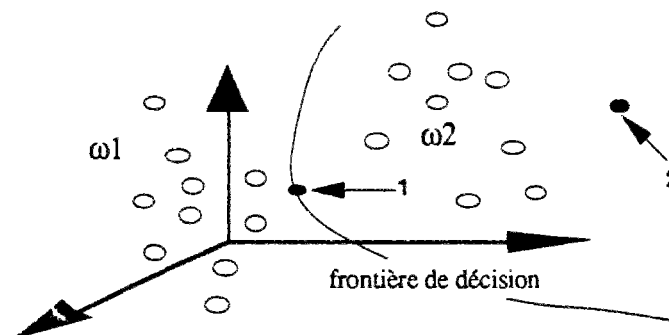


Figure 3.3 Option de rejet

Notons que le rejet de distance est primordial lorsque l'on veut construire un système de décision adaptatif, dans le sens où il est capable de découvrir au fur et à mesure de nouveaux modes de fonctionnement. A la suite d'un certain nombre de rejets en distance, il est possible, par classification automatique (voir paragraphe 3.1.4.2 : méthodes de coalescence, ...), de détecter l'apparition éventuelle de nouvelles classes. Celles-ci doivent être prises en compte lors d'un nouvel apprentissage de façon à adapter la règle de décision. Masson a développé un système de ce type pour la surveillance des structures internes de cuve d'un réacteur nucléaire, [Masson, 1992].

Afin de faciliter la formulation de ces deux types de rejet, on décrit d'abord la règle de Bayes ou du coût minimum. Cette dernière suppose une connaissance totale des lois de probabilité régissant les observations : elles permettent de définir des règles générales qu'il s'agira d'adapter à la réalité lorsque cette connaissance totale n'existera plus.

3.1.3.1 - Règle de Bayes

Soit un vecteur $\bar{x}$ (d composantes : $x_1, x_2, \dots, x_d$) de R^d obéissant, dans la classe ω_i parmi M classes notées $\omega_1, \omega_2, \dots, \omega_M$, à la loi de probabilité multidimensionnelle $f(\bar{x} / \omega_i)$. Les probabilités a priori $P(\omega_i)$ de chaque classe sont aussi données par :

$$\sum_{i=1}^M P(\omega_i) = 1 \quad (3.1)$$

Logiquement, tout vecteur $\bar{x}$ doit être associé à une des M classes : il faut donc trouver une règle de décision $d(\bar{x})$ qui à $\bar{x}$ associe un nombre compris entre 1 et M. Ainsi :

$$d(\bar{x}) = i \quad (3.2)$$

traduit la décision d'associer $\bar{x}$ à la classe ω_i .

La règle de Bayes consiste en effet à associer $\bar{x}$ à la classe pour laquelle la probabilité a posteriori $P(\omega_i / \bar{x})$ est maximale :

$$d(\bar{x}) = i \text{ i.e. } \bar{x} \text{ associé à } \omega_i \text{ si } P(\omega_i / \bar{x}) = \text{Max } P(\omega_j / \bar{x}) \text{ (i, j=1, 2, ..., M)} \quad (3.3)$$

avec :

$$P(\omega_i / \bar{x}) = \frac{f(\bar{x} / \omega_i)P(\omega_i)}{f(\bar{x})} \text{ et} \quad (3.4)$$

$$f(\bar{x}) = \sum_{i=1}^M P(\omega_i)f(\bar{x} / \omega_i) \quad (3.5)$$

$f(\bar{x})$ étant la densité de probabilité non conditionnelle de $\bar{x}$, appelée aussi densité mélange.

3.1.3.2 - Rejet d'ambiguïté et rejet de distance

Notons maintenant la nouvelle formulation du problème : on considère que l'on a M classes et que le vecteur $\bar{x}$ doit appartenir à une de ces M classes ($\omega_1, \omega_2, \dots, \omega_M$) ; mais si l'expression particulière de $\bar{x}$ semble interdire cette affectation a priori, on préfère le placer dans la classe dite de rejet notée ω_0 .

Le système de décision comprend alors M+1 décisions possibles [Devijver, 1982] :

$$\left. \begin{array}{l} d(\bar{x}) = 1 \text{ i.e. } \bar{x} \text{ associé à } \omega_1 \\ d(\bar{x}) = 2 \text{ i.e. } \bar{x} \text{ associé à } \omega_2 \\ \dots\dots\dots \\ d(\bar{x}) = M \text{ i.e. } \bar{x} \text{ associé à } \omega_M \end{array} \right\} : \bar{x} \text{ est classé} \quad (3.6)$$

$$d(\bar{x}) = 0 \text{ i.e. } \bar{x} \text{ associé à } \omega_0 : \bar{x} \text{ est rejeté} \quad (3.7)$$

Nous introduisons maintenant une notion des coûts ou des pénalisations de classification : $C(\omega_i / \omega_j)$ qui représente le coût de classification d'un vecteur appartenant réellement à la classe ω_j dans la classe ω_i (i = 0, 1, ..., M ; j = 1, ..., M).

Dans le cas de la règle dite des coûts (0, 1), nous avons :

$$- C(\omega_i / \omega_i) = 0 \text{ et } C(\omega_i / \omega_j) = 1 \text{ (} i, j = 1, \dots, M ; i \neq j \text{)} \quad (3.8)$$

$$- C(\omega_0 / \omega_j) = C_r \text{ (} j = 1, \dots, M ; C_r \text{ une constante)} \quad (3.9)$$

La règle de rejet d'ambiguïté s'écrit :

$$d(\bar{x}) = \begin{cases} i & \text{i.e. } \bar{x} \text{ associé à } \omega_i \text{ si } P(\omega_i / \bar{x}) = \text{Max } P(\omega_j / \bar{x}) \geq 1 - C_r \text{ (} i, j = 1, 2, \dots, M \text{)} \\ 0 & \text{sinon} \end{cases} \quad (3.10)$$

La règle de rejet de distance, qui s'applique aux points éloignés de toute classe d'apprentissage, peut s'exprimer sous la forme suivante :

$$d(\bar{x}) = 0 \text{ si } f(\bar{x}) \leq C_d \quad (3.11)$$

Cela signifie qu'une observation $\bar{x}$ sera jugée éloignée des classes connues si sa densité de probabilité non conditionnelle est trop faible, c'est-à-dire inférieure à un seuil C_d qu'il convient de fixer. Notons que C_d n'est pas un coût de rejet analogue à C_r .

La figure 3.4 montre bien la position des différentes zones de rejet entre deux classes gaussiennes (ω_1 et ω_2) en dimension 1 ; m_1 la moyenne de la classe ω_1 et m_2 la moyenne de la classe ω_2 avec l'hypothèse $m_1 > m_2$.

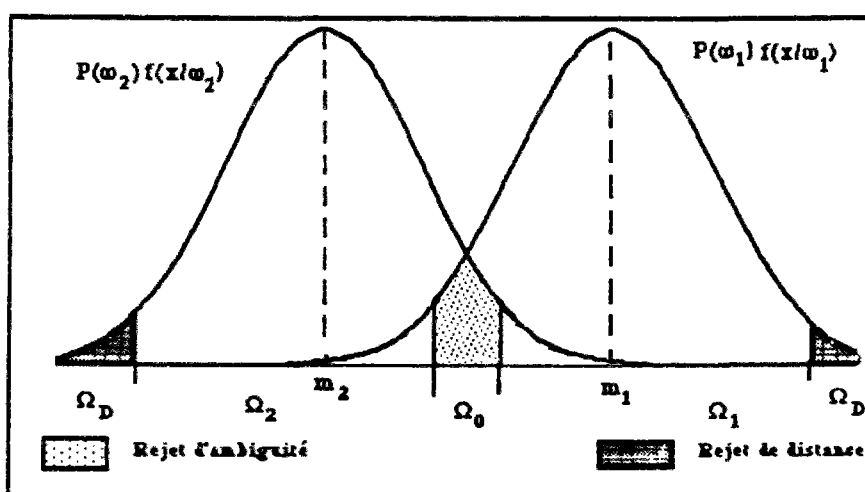


Figure 3.4 Comparaison de zones de rejet de distance et de rejet d'ambiguïté

3.1.4 - PRETRAITEMENT DE DONNEES

Nous avons signalé plus haut que la mise en oeuvre d'une stratégie de reconnaissance des formes nécessitait une phase d'analyse (voir paragraphe 3.1.1.1) qui comprend trois étapes c'est-à-dire le choix d'un espace de représentation, la création d'une partition de l'espace et l'élaboration des frontières entre les classes à discriminer.

Nous présentons dans ce paragraphe les deux premières étapes ; et la troisième étape sera présentée dans le paragraphe 3.2.

On dispose souvent, en début de la phase d'analyse, d'un ensemble de n variables (n supérieur à d) issu des mesures sur le système, dont il importe de faire ressortir des caractéristiques.

C'est l'objectif du prétraitement de données, qui correspond aux deux premières étapes de la phase d'analyse :

- choix d'un espace de représentation, c'est-à-dire sélection et/ou extraction de paramètres/caractères pertinents,
- création d'une partition de l'espace construit par ces d paramètres.

3.1.4.1 - Sélection et extraction des paramètres/caractères

(1) Généralité

Ce point est très important car le choix d'un espace influe sur la qualité du résultat de discrimination. Dans un espace où les classes sont mal séparées, la difficulté de discrimination augmente. Un meilleur choix donne des classes très distinctes, ce qui pose moins de problèmes à la discrimination. Malheureusement il n'y a pas de règle générale pour choisir l'espace de représentation : ce choix ne s'appuie que sur les connaissances a priori existant sur le problème à résoudre, aucune démarche algorithmique ne permet de le formaliser.

En pratique se pose un dilemme : il faut d'une part diminuer le nombre de paramètres/caractères en vue d'alléger le calcul et d'implanter le minimum de capteurs, et d'autre part augmenter ce nombre permettant de mieux décrire et caractériser les formes et améliorer la séparation entre classes. Aussi un compromis est-il nécessaire.

Constituant les composantes des formes d'apprentissage brutes ou initiales, les variables fournies par la mesure sur un système complexe, sont généralement nombreuses et comportent souvent des informations redondantes. Les méthodes de sélection ou extraction de caractères permettent de retenir un sous-ensemble des composantes du vecteur forme initial, sous-ensemble pertinent pour l'objectif fixé. Autrement dit, ce sont les d paramètres, constituant la forme réduite, qui caractérisent les formes observées.

L'extraction (compression) de caractères ("feature extraction" en anglais) utilise un processus de projection (transformation) de variables brutes (dimension élevée d'un espace de représentation) sur un sous espace (dimension réduite). Du point de vue du diagnostic, les méthodes d'extraction de caractères ne diminuent pas le nombre de mesures à effectuer, elles restent au nombre de n mais l'espace de représentation des classes est de dimension plus faible (dimension d).

D'autre part, la sélection de caractères ("feature selection" en anglais) a pour but de chercher d paramètres/caractères parmi les n originels. Par conséquent, le nombre de mesures à effectuer diminue comme la dimension de l'espace de représentation. Pour le diagnostic, la sélection de caractères présente un double intérêt : la réduction à la fois des temps de calcul et du coût des acquisitions en éliminant les informations redondantes ou inintéressantes, [Dubuisson, 1990 a].

Notons par ailleurs qu'il est préférable que la sélection de paramètres s'appuie sur une analyse physique du problème concerné, en vue de trouver des paramètres les plus représentatifs et efficaces vis-à-vis de la discrimination de classes.

(2) Performances d'une méthode de sélection ou extraction de caractères

Les performances d'une méthode de sélection/extraction de paramètres se mesurent à :

- la capacité à mieux représenter la forme observée pour une décision dans un temps minimum (temps réel) ;
- l'élimination de la redondance dans la forme ou, en termes probabilistes, la minimisation de la dépendance stochastique des coordonnées de la forme réduite ;
- l'aptitude à séparer au mieux les formes réduites appartenant à des classes différentes, c'est-à-dire à conserver le maximum de l'information observée rendant possible une bonne classification.

Notons en particulier un critère qui prend de plus en plus d'importance ; c'est celui d'invariance, [Simon, 1984]. Elle découle du fait que la reconnaissance d'une classe revient à mettre en évidence une ou plusieurs propriétés communes à un ensemble de représentations ou formes différentes, donc leurs invariances. Les variants dépendent de l'application considérée. Dans la mesure où l'on souhaite trouver des caractères indépendants de l'application, ce sont ceux vérifiant de tels invariants qu'il faudrait sélectionner.

Cette attitude est à rapprocher de celle observée en physique théorique et qui en constitue un postulat fondamental : les lois de la physique expriment l'invariance (ou la conservation) d'une certaine grandeur.

En pratique, la transposition de cette attitude à la reconnaissance des formes statistique ne se fait pas sans problème. Deux formes, qui ne diffèrent que par une transformation faisant partie des invariants du problème, devraient avoir la même représentation. Mais il y a en pratique des bruits de mesure, de codage, de transmission, etc. qui font que ces représentations ne coïncident pas.

Une démarche commune consiste à faire en sorte que les nuages de points similaires soient aussi compacts que possible, [Diday, 1980], exigence d'autant plus difficile à satisfaire que la dimension de l'espace de représentation est élevée.

3.1.4.2 - Création d'une partition de l'espace de représentation

L'expérience nécessaire pour constituer un ensemble d'apprentissage sera ordinairement constituée par un grand nombre de formes d'apprentissage. Deux circonstances peuvent alors se présenter :

- ou ces formes d'apprentissage ont préalablement été rangées dans des classes d'apprentissage, par un étiquetage manuel aidant par la connaissance de l'application envisagée par exemple. On parle alors d'apprentissage supervisé avec professeur ;
- ou ces formes d'apprentissage n'ont pas été classées a priori et on parle alors d'apprentissage non supervisé sans professeur.

En matière d'apprentissage supervisé, d'abord on se borne à chercher des représentants de chaque classe. Ensuite on doit chercher une règle de décision pour séparer les classes qui ont été définies par le professeur ; cette règle sera présentée dans le chapitre suivant.

Et en matière d'apprentissage non supervisé, il faut faire usage d'algorithmes de coalescence, encore appelés méthodes de "clustering" (en anglais). Sans aucune hypothèse préalable autre que la fonction de similarité utilisée, ces algorithmes heuristiques regroupent les formes d'apprentissage les plus semblables et on construit ainsi séquentiellement un ensemble de classes naturelles, de sorte que l'on est ramené au problème de l'apprentissage supervisé.

Dans la littérature, les algorithmes de coalescence utilisent essentiellement des fonctions de similarité s'exprimant à l'aide des seules composantes des formes et requièrent généralement une connaissance (ou estimation) du nombre maximal de classes naturelles. Citons par exemple les nuées dynamiques, [Diday 1980] ou les k-moyennes, [Duda 1973].

3.2 - METHODE CONNEXIONNISTE EN TANT QU'OUTIL POUR LA RECONNAISSANCE DES FORMES

3.2.1 - INTRODUCTION

Depuis toujours, l'homme s'est attaché à comprendre le fonctionnement du monde vivant et notamment le processus complexe de la cognition ; il est tout à fait logique de considérer le cerveau humain comme l'un des plus performants pour effectuer ce genre de tâches. D'autre part, le neurone ou cellule nerveuse est, on le sait, l'élément fonctionnel de base du cerveau. On comprend alors l'intérêt suscité par les **réseaux connexionnistes**.

C'est en 1943 que Mac Culloch et Pitts, [Mc Culloch, 1943] proposent un modèle qui va servir de base aux réseaux de neurones actuels. S'appuyant sur des connaissances neurobiologiques, ils mettent au point un modèle booléen du neurone humain.

Parallèlement, les travaux de Hebb, [Hebb, 1949], neuro-physiologiste, mettent en avant l'importance des **synapses** dans le processus d'apprentissage. En 1949, il propose une règle de modification synaptique qualitative.

Grâce à ces premiers travaux, des modèles vont se développer. En 1958, le **perceptron** de Rosenblatt [Rosenblatt, 1958], premier modèle de réseau, est capable d'apprendre certaines fonctions simples en modifiant progressivement la force des liaisons synaptiques entre les neurones. Dans le même temps, Widrow et Hoff, [Widrow, 1960], deux électroniciens travaillent sur des structures particulières de réseaux pour le filtrage adaptatif du signal : le modèle **adaline** (ADaptive LInear NEuron) et le modèle **madaline** (Many ADaptive LInear NEuron). L'intérêt du travail de Widrow et Hoff réside surtout dans la règle d'adaptation des poids utilisée dite **règle de Widrow-Hoff**.

Ces travaux vont être violemment controversés en 1969 par Minsky et Papert qui montrent les limites du Perceptron, [Minsky, 1969], entraînant par là même un ralentissement des travaux sur les réseaux. C'est vers les années 80, que l'engouement pour les réseaux reprend. Les recherches de Hopfield issues de la physique, et surtout la mise au point d'un **algorithme d'apprentissage** pour les **réseaux multi-couches : Rétro-Propagation (RP)** ("Back-Propagation" ou "BP" en anglais) du gradient de l'erreur (voir paragraphe 3.2.2.3), [Le Cun, 1985], [Parker, 1985] et [Rumelhart, 1986], ouvrent de nouveaux horizons aux modèles connexionnistes.

Aujourd'hui, le domaine d'application des méthodes connexionnistes est déjà très étendu. Elles ont été introduites en automatique (commande de processus), en traitement du signal (compression, prédiction, filtrage), en synthèse de la parole (NETtalk : apprendre à un réseau à prononcer le texte anglais), dans certains domaines où les méthodes de l'intelligence artificielle sont prédominantes, et aussi en diagnostic (reconnaissance des formes), etc.. Nous nous cantonnerons ici aux notions utiles en reconnaissance des formes.

3.2.2 - DEFINITION D'UN SYSTEME CONNEXIONNISTE

On estime à l'heure actuelle que le cerveau comprend entre 10^{10} et 10^{11} neurones et que chaque neurone peut recevoir ou envoyer de 10 000 à 100 000 signaux. Si la justification du terme neuro-mimétique et autres terminologies comme **réseaux de neurones artificiels**, ou **réseaux de neurones formels**, semble aisée, il va de soi que les modèles informatiques sont pour l'instant bien loin des performances de leur illustre modèle. Ce domaine en plein essor ouvre cependant de nouvelles voies de recherche, la parallélisation et la possibilité de réalisations matérielles dédiées, apparaissant comme l'atout majeur de ces techniques.

Dans un **système connexionniste**, la connaissance réside précisément dans les **connexions**; c'est-à-dire dans des valeurs numériques (appelées des **poids synaptiques**) associées aux arcs reliant les **unités** (appelées aussi **automates**, **cellules**, **neurones**). Chaque unité possède une **activité** caractérisée par un nombre.

Pour définir un réseau connexionniste, il y a trois éléments à fournir :

- la **structure** : la manière dont les unités seront interconnectées et les contraintes liant les poids ;
- la **dynamique** : comment les activités des unités sont couplées entre elles, comment les signaux se propagent dans le réseau ;
- l'**apprentissage** : les règles qui tentent de déterminer les valeurs "optimales" des poids synaptiques pour que le réseau accomplisse une certaine tâche. Cette tâche est caractérisée par une relation prédéfinie (en général par une liste d'exemples) entre les activités de certaines unités.

Les trois aspects sont présentés successivement ci-après.

3.2.2.1 - Structure

Les structures utilisées sont très variées : réseaux totalement connectés, localement connectés, organisés en couches successives. Le choix des connexions facilite néanmoins l'utilisation du réseau et l'interprétation des poids. Différentes structures de réseaux sont illustrées par la Figure 3.5.

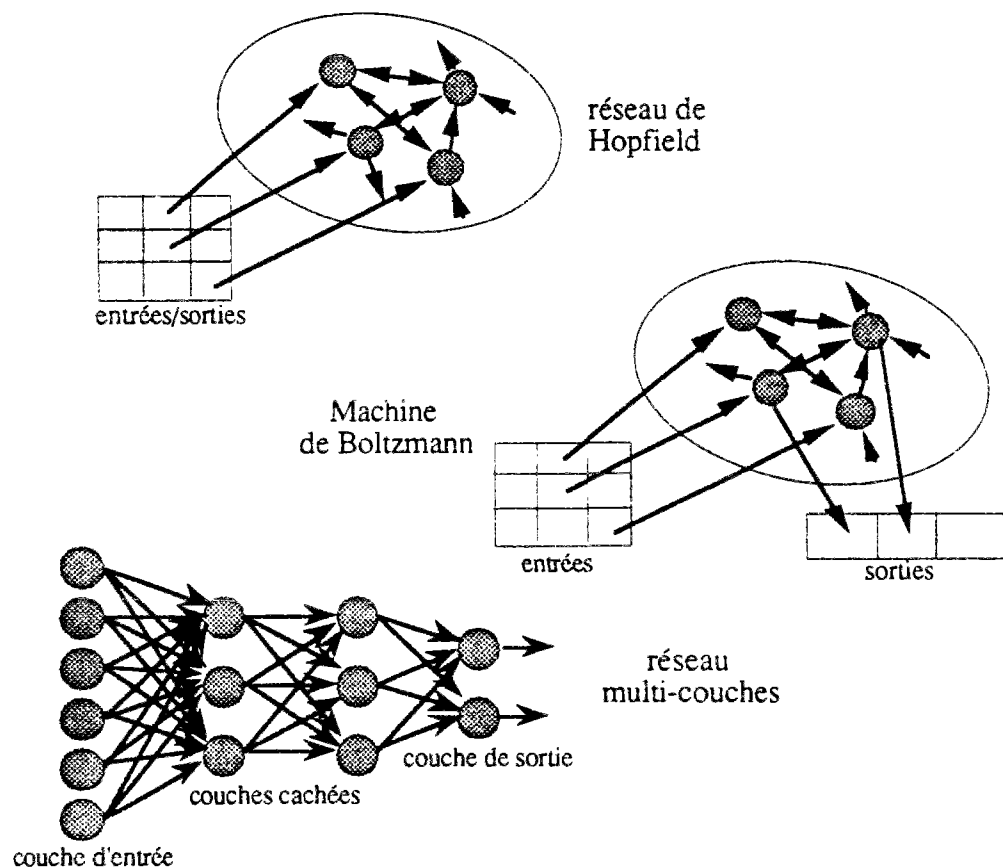


Figure 3.5 Structures de réseaux

3.2.2.2 - Dynamique

Selon le type d'entrées et le type de fonction de transfert ou fonction d'activation, on peut distinguer plusieurs types de neurones. Nous indiquons quelques exemples parmi les fonctions de transfert les plus employées dans le Tableau 3.1.

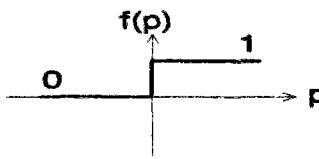
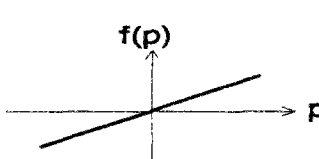
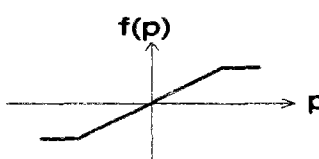
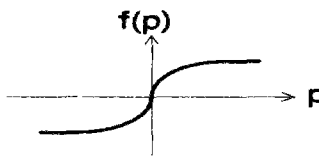
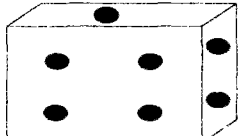
	entrées / sorties	fonction de transfert
automate à seuil (Mac Culloch & Pitts)	- entrées binaires ou continues - sorties binaires	échelon : 
automate linéaire	- entrées réelles - sorties réelles	linéaire : 
automate à saturation	- entrées réelles - sorties bornées	linéaire bornée : 
automate continu	- entrées réelles - sorties réelles	sigmoïde : 
automate probabiliste	- entrées quelconques - sorties binaires	fonction stochastique : 

Tableau 3.1 Exemples de fonctions de transfert

Notons en particulier que :

- la **fonction de transfert à échelon** est utilisée pour créer des **neurones perceptron**. Les réseaux perceptron (une couche d'entrées et une couche de sortie) sont utilisés comme discriminateurs en termes de reconnaissance des formes. Ils sont efficaces seulement dans le cas où les formes d'entrées de réseau sont linéairement séparables, ce qui est très rare dans le monde réel. Une solution pour ce problème est d'utiliser un processus de prétraitement conduisant à un ensemble de formes linéairement séparables. Notons également que la sortie d'un perceptron ne peut prendre qu'une de deux valeurs (0 ou 1) et cela limite évidemment l'application de ce type de réseaux.
- la **fonction de transfert linéaire** est utilisée pour créer des réseaux de **neurones linéaires** (une couche d'entrées et une couche de sortie). Ce type de réseaux peut classifier des formes d'entrées non-linéairement séparables. Les sorties des neurones peuvent prendre n'importe quelle valeur. Un réseau linéaire peut réaliser une approximation de fonctions linéaires ; mais il ne peut pas présenter correctement la non-linéarité entre une forme d'entrée et sa sortie cible de réseau.
- l'introduction de la **fonction de transfert sigmoïde** permet de créer des **réseaux Perceptron Multi-Couches** (PMC) (voir paragraphe 3.2.3). Ce type de réseau peut s'affranchir des problèmes de formes non-linéairement séparables et des problèmes de la non-linéarité entre une forme d'entrée et sa sortie cible de réseau.

On présente ci-après, comme exemple, la formulation d'un neurone artificiel multi-entrées.

Un **neurone artificiel multi-entrées** (Figure 3.6) reçoit de ses semblables ou de l'extérieur un nombre n de **signaux activateurs** x_j par l'intermédiaire de synapses ; il effectue ensuite une sommation pondérée (par w_{ij} associé à chaque synapse) de ces signaux pour calculer un **potentiel d'activation** p_i :

$$p_i = \sum_j w_{ij} x_j \quad (j = 1, \dots, n) \quad (3.12)$$

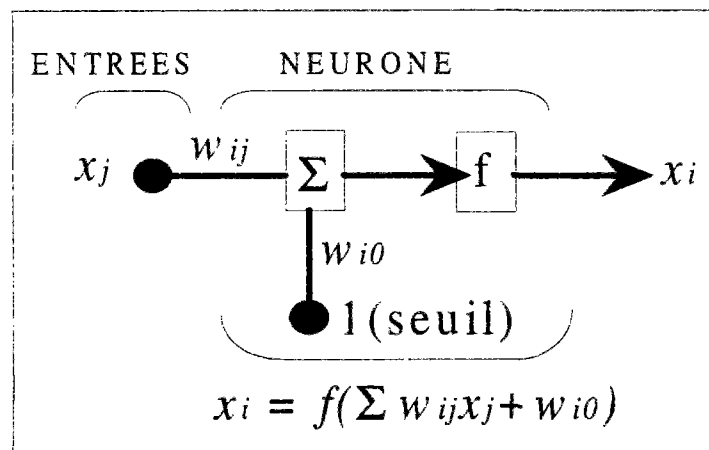


Figure 3.6 *Un neurone artificiel multi-entrées*

Une fonction de transfert ou fonction d'activation f produit enfin, par comparaison avec un seuil w_{i0} , un signal de sortie x_i lui-même envoyé vers d'autres neurones :

$$x_i = f(p) \quad (3.13)$$

$$\text{où : } p = p_i + w_{i0} = \sum_j w_{ij} x_j + w_{i0} \quad (3.14)$$

La dynamique d'un réseau constitué par ce type de neurones sera donc du type **propagation vers l'avant** ("feedforward" en anglais).

Notons que les poids w_{ij} et les seuils w_{i0} sont tous des paramètres scalaires ajustables. L'idée centrale des réseaux de neurones est que ces paramètres peuvent être ajustés de façon à ce que ces réseaux présentent certains comportements désirés et intéressants. Par conséquent, on peut apprendre aux réseaux à accomplir une tâche particulière en réglant les poids et les seuils des réseaux, ou les réseaux pourraient eux-mêmes modifier ces paramètres pour atteindre des objectifs désirés.

Un neurone dispose en général d'un seuil, mais on peut omettre ce seuil dans certains cas. En effet, celui-ci est fréquemment considéré comme un poids synaptique ajustage associé à une entrée supplémentaire fixée à 1. La formule (3.14) peut donc s'écrire :

$$p = \sum_j w_{ij} x_j \quad (j = 0, 1, \dots, n; x_0 = 1) \quad (3.15)$$

3.2.2.3 - Apprentissage

L'apprentissage, c'est-à-dire la capacité d'acquérir de nouveaux comportements ou d'en modifier d'anciens par expérience, est une propriété de presque tous les animaux, y compris ceux que l'on peut difficilement qualifier d'intelligents [Le Un, 1987].

Les travaux sur l'apprentissage ont débuté très tôt dans l'histoire de l'intelligence artificielle, mais ses succès sont restés très marginaux, excepté dans le domaine de la reconnaissance des formes, [Le Un, 1987]. Deux tendances se sont dessinées dès le départ : les méthodes plutôt heuristiques et/ou symboliques d'une part, et les méthodes plutôt numériques d'autre part.

Les méthodes connexionnistes sont les héritières des travaux de la deuxième tendance. En effet, la propriété la plus intéressante des réseaux connexionnistes est leur faculté d'apprentissage, c'est-à-dire leur capacité à mémoriser des connaissances grâce à l'expérience. Cette information est entièrement contenue dans les poids synaptiques w_{ij} (valeurs numériques). On distingue généralement deux types d'apprentissage :

- l'apprentissage supervisé qui consiste à établir les poids de façon à imposer pour chaque entrée une sortie désirée - utilisation d'un "professeur" ;
- l'apprentissage non-supervisé, réalisé sans "professeur" grâce à un processus d'auto-organisation basé sur une compétition engagée entre les neurones.

Le premier (l'apprentissage associatif) s'apparente aux méthodes non-paramétriques de reconnaissance des formes ; et le deuxième (l'apprentissage compétitif) relève de son côté de méthodes de coalescence.

On n'a pas l'intention de lister ici tous les algorithmes d'apprentissage. En revanche, on présentera, dans le paragraphe suivant, l'algorithme de Rétro-Propagation du gradient qui sera utilisé pour notre application : apprentissage des réseaux Perceptron Multi-Couches (apprentissage associatif).

3.2.3 - UN MODELE CONNEXIONNISTE - RESEAUX PERCEPTRON MULTI-COUCHES

3.2.3.1 - Présentation du modèle

Un réseau PMC (Perceptron Multi-Couches) ("multi-layers perceptron" en anglais) désigne un réseau connexionniste ayant une structure particulière, une dynamique propagation vers l'avant (sans circuit) et une règle d'apprentissage du type RP.

Un PMC est, comme son nom l'indique, organisé en couches successives (une couche d'entrée, une ou plusieurs couches cachées, et une couche de sortie) ; un neurone ne reçoit des signaux que des neurones en amont (couche immédiatement précédente).

La **structure** d'un réseau à deux couches (la première couche, n'exécutant rien, n'est pas comptée dans le nombre des couches) est schématisé par la Figure 3.7. Il faut se souvenir qu'en absence de couche cachée, un PMC devient un réseau adaline et que l'apprentissage des poids peut être effectué en utilisant la règle de Widrow-Hoff.

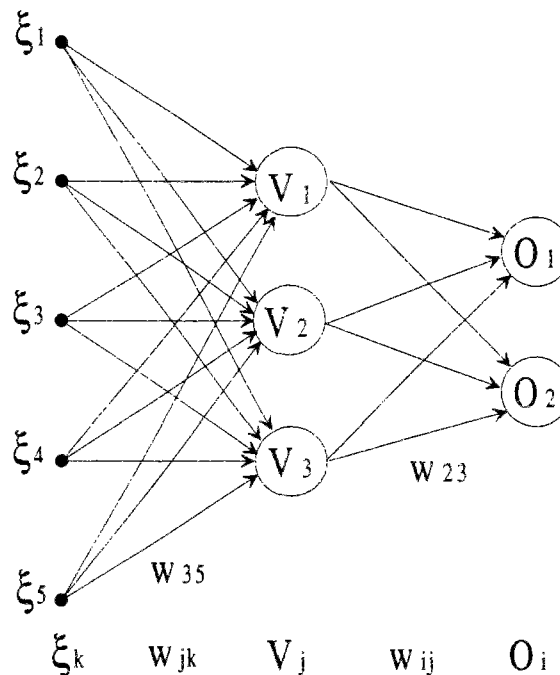


Figure 3.7 Exemple d'un réseau à deux couches et notations

Les réseaux PMC peuvent être vus d'une façon générale comme un outil de modélisation numérique de la fonction réalisant le passage d'un espace quelconque dans un autre. Ceci peut être réalisé sans information a priori sur la forme de la fonction en utilisant une base d'exemples, ou base d'apprentissage, c'est-à-dire un ensemble de couples de données (espace d'entrée - espace de sortie) qui vont permettre l'apprentissage du réseau.

Pour des raisons de clarté, la formulation de la dynamique et de l'algorithme d'apprentissage se réalise avec un réseau à deux couches (Figure 3.6). L'extension à un nombre de couches plus élevé se déduit aisément. Les notations utilisées sont les suivantes :

O_i : neurones de sortie

V_j : neurones cachés

ξ_k : neurones d'entrée

O_i^μ : sortie produite par un neurone de sortie O_i en présence de chaque forme d'entrée μ

V_j^μ : sortie produite par un neurone caché V_j en présence de chaque forme d'entrée μ

ξ_k^μ : une des composantes d'une forme d'entrée μ , ou d'un mode de fonctionnement en termes de diagnostic (notons que chaque neurone d'entrée ξ_k , excepté le neurone seuil qui reçoit une valeur fixée à 1, reçoit une valeur ξ_k^μ à chaque présentation d'une forme μ)

w_{ij} : poids reliant les neurones cachés aux neurones de sortie

w_{jk} : poids reliant les neurones d'entrée aux neurones cachés

N_{ns} : nombre de neurones de sortie (pas de neurone seuil) ($i = 1, \dots, N_{ns}$)

N_{nc} : nombre de neurones cachés (le neurone seuil de la couche cachée compris) ($j = 1, \dots, N_{nc}$)

N_{ne} : nombre de neurones d'entrée (le neurone seuil de la couche d'entrée compris) ($k = 1, \dots, N_{ne}$)

N_{fe} : nombre de formes d'entrée ($\mu = 1, \dots, N_{fe}$)

Pour une forme donnée μ , un neurone caché V_j reçoit une entrée (un potentiel d'activation + un seuil) :

$$p_j^\mu = \sum_k w_{jk} \xi_k^\mu \quad (3.16)$$

et produit avec une fonction de transfert f :

$$V_j^\mu = f(p_j^\mu) = f\left(\sum_k w_{jk} \xi_k^\mu\right) \quad (3.17)$$

Ensuite un neurone de sortie O_i reçoit :

$$p_i^\mu = \sum_j w_{ij} V_j^\mu = \sum_j w_{ij} f\left(\sum_k w_{jk} \xi_k^\mu\right) \quad (3.18)$$

et produit (toujours avec une fonction de transfert) :

$$O_i^\mu = f(p_i^\mu) = f\left(\sum_j w_{ij} V_j^\mu\right) = f\left(\sum_j w_{ij} f\left(\sum_k w_{jk} \xi_k^\mu\right)\right) \quad (3.19)$$

La **dynamique** d'un réseau PMC est donc décrite par les équations (3.16), (3.17), (3.18) et (3.19) d'où l'appellation : propagation vers l'avant (antérograde).

La fonction de transfert f (formules (3.17), (3.18) et (3.19)) doit être dérivable pour permettre la mise au point d'un algorithme d'apprentissage utilisant les propriétés de dérivation des fonctions composées (voir les formules (3.22), (3.23), (3.24) et (3.25)) ; ce qui exclut la fonction Signe. La fonction f la plus utilisée est la sigmoïde définie par :

$$f(p) = h \frac{e^{rp} - 1}{e^{rp} + 1} \quad (3.20)$$

où : h : paramètre qui indique l'amplitude de la sigmoïde

r : paramètre (raideur) qui règle la pente de la partie quasi linéaire de la sigmoïde.

Notons que si $r \rightarrow +\infty$, la fonction tend vers la fonction Signe.

Notons également que pour les neurones seuils, dont la sortie est fixée à +1 et qui n'ont pas de neurones amonts, il faudrait considérer un potentiel d'activité p égal à $+\infty$. En pratique, on tient compte, pour chaque neurone seuil (un par couche), du fait que sa sortie est fixée à +1 sans faire intervenir son activité.

Après les descriptions de la structure et de la dynamique du réseau PMC, il faut maintenant définir la règle d'**apprentissage**.

La Rétro-Propagation est certainement l'un des plus simples et des plus efficaces algorithmes d'apprentissage pour réseaux à neurones cachés ; elle est simplement une généralisation de la règle de Widrow-Hoff à une fonction coût non-linéaire.

En effet, pour une forme d'entrée μ , on fixe un vecteur cible de sortie $\tilde{\zeta}^\mu$ (N_{ns} composantes : $\zeta_1^\mu, \dots, \zeta_i^\mu, \dots, \zeta_{N_{ns}}^\mu$) et on apprend le réseau de telle manière que :

$$O_i^\mu = \zeta_i^\mu \text{ (résultat désiré pour chaque } i \text{ et } \mu) \quad (3.21)$$

Pour ce faire, il faudrait d'abord trouver une configuration de poids qui minimise un certain critère. En l'occurrence, ce critère ne doit directement dépendre que des états des neurones de sortie, et pas des neurones cachés (c'est-à-dire pas de sortie désirée associée aux neurones cachés). C'est pour cela que l'on introduit une fonction coût du type :

$$E(w) = \frac{1}{2} \sum_{\mu i} [\zeta_i^\mu - O_i^\mu]^2 \text{ ou encore} \quad (3.22)$$

$$E(w) = \frac{1}{2} \sum_{\mu i} \left[\zeta_i^\mu - f\left(\sum_j w_{ij} f\left(\sum_k w_{jk} \xi_k^\mu\right)\right) \right]^2 \quad (3.23)$$

Notons que d'autres fonctions coût sont possibles, mais celle-ci est très utilisée.

Ensuite, il s'agit de calculer les dérivées partielles de E par rapport à tous les w_{ij} et w_{jk} de manière à être en mesure d'appliquer l'algorithme du gradient (les formules (3.24) et (3.25)) afin de modifier ces poids jusqu'à $E(w) \rightarrow 0$:

$$\left. \begin{aligned} w_{ij, t+1} &= w_{ij, t} - \eta \frac{\partial E}{\partial w_{ij, t}} \\ w_{jk, t+1} &= w_{jk, t} - \eta \frac{\partial E}{\partial w_{jk, t}} \end{aligned} \right\} \text{ ou} \quad (3.24)$$

$$\left. \begin{aligned} \Delta w_{ij, t+1} &= w_{ij, t+1} - w_{ij, t} = -\eta \frac{\partial E}{\partial w_{ij, t}} \\ \Delta w_{jk, t+1} &= w_{jk, t+1} - w_{jk, t} = -\eta \frac{\partial E}{\partial w_{jk, t}} \end{aligned} \right\} \quad (3.25)$$

où : t : itération t

η : pas d'adaptation

Il faut maintenant calculer les gradients des erreurs sur la couche de sortie et la couche cachée pour formuler Δw_{ij} et Δw_{jk} (pour ne pas alourdir la formulation, les indices t et $t+1$ sont omis dans les formules suivantes)

D'une part, pour chaque poids liant un neurone caché et un neurone de sortie w_{ij} , la règle du gradient (voir les formules (3.24) ou (3.25)) peut s'écrire :

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \sum_{\mu} [\zeta_i^\mu - O_i^\mu] f'(p_i^\mu) V_j^\mu = \eta \sum_{\mu} \delta_i^\mu V_j^\mu \quad (3.26)$$

$$\text{où : } \delta_i^\mu = f'(p_i^\mu) [\zeta_i^\mu - O_i^\mu] \quad (3.27)$$

D'autre part, pour chaque poids liant un neurone d'entrée et un neurone caché w_{jk} , la relation de dépendance avec l'état d'un neurone de sortie est indirecte et totalement non-linéaire ; elle dépend des poids et des états de presque tous les autres neurones du réseau. L'expression de la règle du gradient est donc plus complexe. En effet, par une technique de chaînage, la règle du gradient (3.24) ou (3.25) pour w_{jk} s'écrit :

$$\begin{aligned}
 \Delta w_{jk} &= -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \frac{\partial E}{\partial V_j^\mu} \frac{\partial V_j^\mu}{\partial w_{jk}} \\
 &= \eta \sum_{\mu} [\xi_i^\mu - O_i^\mu] f'(p_i^\mu) w_{ij} f'(p_j^\mu) \xi_k^\mu \\
 &= \eta \sum_{\mu} \delta_i^\mu w_{ij} f'(p_j^\mu) \xi_k^\mu \\
 &= \eta \sum_{\mu} \delta_j^\mu \xi_k^\mu
 \end{aligned} \tag{3.28}$$

$$\text{avec : } \delta_j^\mu = f'(p_j^\mu) \sum_i w_{ij} \delta_i^\mu \tag{3.29}$$

Les formules (3.26) et (3.28) permettent donc de calculer les gradients de tous les poids par propagation de l'erreur des sorties vers les entrées, d'où le nom de **Rétro-Propagation**.

Bien que les formules (3.26) et (3.28) soient écrites par la sommation de toutes les formes μ ($\mu = 1, 2, \dots, N_l$), il y a en pratique deux approches différentes pour l'adaptation des poids w_{ij} et w_{jk} .

La première est une **méthode incrémentale** qui consiste à adapter les poids après chaque présentation d'une forme. Cette approche fait décroître sensiblement la fonction coût (même pour un η assez petit) à chaque itération, et peut conduire à des minima locaux de la fonction coût (ceci s'expliquera dans le paragraphe 3.2.3.2). L'ordre aléatoire de la présentation de l'ensemble des formes d'entrée permet une exploration plus large de la surface coût, et donc une résolution partielle des problèmes des minima locaux.

La deuxième est une **méthode globale** ("alternative batch mode" en anglais) qui consiste à n'adapter les poids qu'après la présentation de tous les exemples (les formules (3.26) et (3.28)) ; il demande un mémoire de calcul plus important.

L'efficacité de ces deux approches dépend du problème traité. Cependant l'approche incrémentale semble meilleure dans beaucoup de cas, en particulier dans le cas où les exemples d'apprentissage sont réguliers ou redondants.

On présente, comme exemple, l'algorithme de RP de la première approche comme suit :

Initialiser tous les poids w_{ij} et w_{jk} .

Tant que le critère d'arrêt non atteint⁽¹⁾

Pour toutes les formes μ ($\mu = 1, 2, \dots, N_r$) de la base d'apprentissage

Faire

- Présentation d'une forme μ
- Calcul des sorties des neurones cachés et des neurones de sortie (propagation vers l'avant) (formules (3.16)-(3.19))
- Calcul de l'erreur commise par rapport aux cibles (formules (3.22)-(3.23))
- Application de la règle du gradient (formules (3.26)-(3.28)) et modification des poids (Rétro-Propagation)

Fait

Fin, Tant que

⁽¹⁾ Plusieurs critères d'arrêt sont possibles, citons parmi les plus employés :

- arrêt si le nombre d'itérations est supérieur à une certaine borne,
- arrêt si la somme des erreurs en sortie du réseau pour toutes les formes d'apprentissage est inférieure à un certain seuil,
- arrêt si la progression de l'erreur entre deux itérations est inférieure à un certain seuil.

Ces deux derniers critères sont plus employés en pratique : dans le cas où se fixe la valeur minimale de la fonction coût, on utilise plutôt le deuxième critère ; dans le cas contraire on préfère utiliser le troisième.

3.2.3.2 - Mise en oeuvre pratique

(1) Choix de fonctions de transfert

Les fonctions de transfert les plus employées sont : la sigmoïde logarithmique (sortie de neurone entre 0 et 1), la sigmoïde tangentielle (sortie de neurone entre -1 et 1) et la fonction linéaire non-bornée (sortie de neurone quelconque) seulement utilisée pour les neurones de la couche de sortie.

L'expérience montre qu'un réseau muni de sigmoïde tangentielle converge mieux qu'un réseau à sigmoïde logarithmique. L'explication de ce phénomène est simple : en effet, lorsqu'un réseau commence à se structurer, un état 0 désiré aura tendance à beaucoup moins influencer une variation des poids qu'un état 1 désiré d'où un certain déséquilibre de l'apprentissage.

(2) Initialisation des poids synaptiques

Il existe un lien entre les valeurs des poids et les autres paramètres. On initialise généralement les poids avec des valeurs faibles tirées aléatoirement et calculées en fonction du nombre des formes d'apprentissage, la taille des couches cachées (nombre des neurones cachés) et la ou les fonctions de transfert utilisées. Ceci doit permettre au réseau de se situer au départ dans une zone dynamique de la sigmoïde, où la dérivée de la fonction, qui intervient dans la modification des poids, est significative.

(3) Minima locaux

Peut-on garantir la convergence de la Rétro-Propagation ? la réponse est malheureusement « non ».

En effet, la surface de la fonction coût n'est généralement pas convexe et ses minima ne sont pas forcément des solutions « exactes », c'est-à-dire avec une erreur nulle. La solution finale d'un réseau peut donc être piégée par ses minima locaux au fur et à mesure de l'application de la règle du gradient ; ceci dépend par ailleurs des conditions initiales (poids initiaux).

La Figure 3.8 montre la surface de la fonction coût d'un réseau très simplifié (un seul neurone caché). On y remarque que la fonction coût a un minimum global situé au centre de la surface. En revanche, les vallées de deux côtés conduisent à des minima locaux.

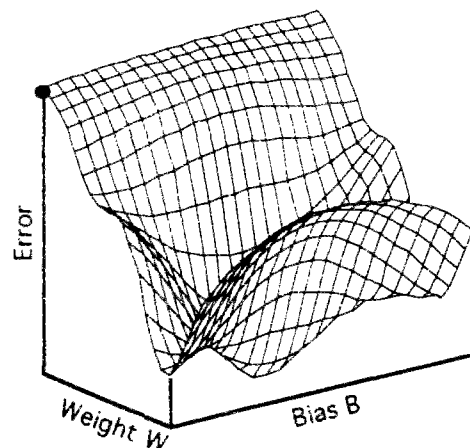
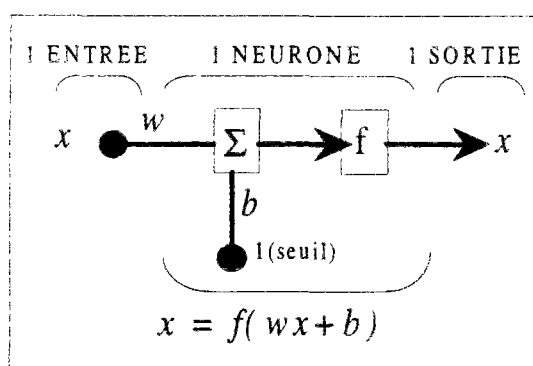


Figure 3.8 Surface de l'erreur (ou surface de la fonction coût) du réseau avec 1 neurone caché

La convergence vers un minimum local peut être suffisante si celui-ci est proche du minimum global ou/et satisfait à l'erreur désirée.

Comment faire face aux problèmes des minima locaux ? C'est un problème général d'optimisation.

Les tentatives pour créer des algorithmes afin d'éviter ces minima locaux n'ont pas complètement abouti. Une meilleure approche serait, semble-t-il, de garantir que les minima locaux produisent des erreurs suffisamment petites pour être acceptables vis-à-vis du problème. En conséquence, même si le réseau est piégé dans une vallée peu profonde (un minimum local), la solution peut être satisfaisante.

Pour cela, une méthode souvent utilisée est de donner au réseau plus de degrés de liberté, c'est-à-dire augmenter le nombre des couches cachées ou/et la taille de celles-ci. Notons que cette méthode doit être utilisée avec précaution car trop de neurones cachés entraînent un apprentissage par coeur ("overfitting" en anglais). On reviendra plus loin sur ce point.

D'autres mesures sont également prises pour réduire la possibilité de piéger un réseau dans les minima locaux. Citons comme exemples :

- initialiser les poids synaptiques de façon à situer au départ le réseau dans une zone dynamique de la sigmoïde,
- utiliser la méthode incrémentale d'apprentissage en présentant toute les formes d'apprentissage dans un ordre aléatoire,
- ajouter un terme de moment ("momentum" en anglais) à la formule classique de modification des poids synaptiques,
- utiliser des méthodes d'optimisation plus puissantes,
- etc.

(4) Amélioration (ou variation) de l'algorithme de Rétro-Propagation

L'algorithme Rétro-Propagation peut être amélioré avec deux approches différentes : l'approche heuristique (l'apprentissage avec moment par exemple), et l'approche utilisant des méthodes d'optimisation plus puissantes (l'optimisation Levenberg-Marquardt par exemple). En pratique, l'algorithme Rétro-Propagation classique est rarement utilisé.

Moment et pas d'adaptation

Un moment permet au réseau non seulement d'essayer de répondre aux problèmes des minima locaux mais aussi d'orienter la direction de descente sur la surface coût. Fonctionnant comme un filtre passe-bas, un moment permet au réseau d'ignorer les petits obstacles sur la trajectoire de l'erreur, et par conséquent de s'affranchir de certains minima locaux.

En ajoutant un terme de moment à la formule classique de modification des poids (la formule (3.25)), la règle du gradient s'écrit :

$$\left. \begin{aligned} \Delta w_{ij, t+1} &= w_{ij, t+1} - w_{ij, t} = -\eta \frac{\partial E}{\partial w_{ij, t}} + m_c \Delta w_{ij, t} \\ \Delta w_{jk, t+1} &= w_{jk, t+1} - w_{jk, t} = -\eta \frac{\partial E}{\partial w_{jk, t}} + m_c \Delta w_{jk, t} \end{aligned} \right\} \quad (3.30)$$

avec le paramètre moment m_c qui varie entre 0 et 1 ; la valeur 0.9 est souvent utilisée.

Un terme de moment est utile dans les deux approches d'apprentissage : incrémentale ou globale. Il a été proposé pour l'apprentissage incrémental, où chaque évolution des poids n'est liée qu'à la forme présentée (apprentissage partiel) et ceci entraîne certaines fluctuations sur la trajectoire de la surface coût. Dans le cas de la méthode globale, l'évolution s'appuie sur toutes les formes, c'est-à-dire qu'il y a intégration de toute l'information.

D'autre part, le pas d'adaptation η (voir la formule 3.30) doit être étudié. En effet, choisir le pas d'adaptation est un vrai défi pour un réseau non-linéaire ; ce choix conditionne la convergence de l'apprentissage. Un pas trop faible a pour effet de rendre la phase d'apprentissage très lente, un pas trop important peut entraîner des oscillations des valeurs des poids. D'autre part une valeur de pas peut être bien adaptée au début de l'apprentissage et plus du tout en fin d'apprentissage. Une approche usuelle, [Hertz, 1991], consiste à modifier η en fonction des progrès sur le trajectoire de l'erreur. Si l'adaptation des poids ne produit pas une décroissance de l'erreur, alors le pas doit être réduit et l'on annule la modification apportée aux poids. Dans le cas contraire, on peut essayer d'être moins "timide" et augmenter η . Le schéma général d'adaptation de η est donc le suivant :

$$\begin{aligned} \Delta\eta &= -b\eta & \text{si } \Delta E > 0 \\ a & & \text{si } \Delta E < 0 \quad (\text{consécutivement}) \\ 0 & & \text{si non} \end{aligned} \quad (3.31)$$

avec: $\Delta\eta$: variation de η

ΔE : variation de la fonction coût

a et b : constantes (>0)

Le terme « consécutivement » signifie les k dernières itérations.

Par ailleurs, l'utilisation d'un processus d'adaptation de η (voir la formule (3.31)) peut accélérer l'apprentissage car ce processus essaie de garder un pas η le plus grand possible tout en garantissant la stabilité d'apprentissage.

Certains auteurs, [Jacobs, 1988], ont également préconisé l'usage d'une telle procédure et d'un pas d'adaptation différent pour chaque poids du réseau ou pour chaque forme d'apprentissage.

Optimisation Levenberg-Marquardt

L'optimisation proposée par Levenberg-Marquardt utilise une extension de la méthode Newton et présente des meilleures performances que la Rétro-Propagation classique. En revanche, cette méthode demande beaucoup plus de mémoire pour l'apprentissage ; par conséquent, elle est plutôt employée pour l'apprentissage « hors ligne » pour une implantation ultérieure dans un système comme un réseau physique avec les poids synaptiques prédéterminés.

L'algorithme de Levenberg-Marquardt s'écrit⁽²⁾ :

$$\Delta w(a) = (D + \eta I)^{-1} \nabla E(a) \text{ et} \quad (3.32)$$

$$D = J^T J \quad (3.33)$$

où :

a : vecteur qui représente l'espace des poids synaptiques

J : matrice Jacobienne

Notons que si la valeur de η est très grand, l'expression ci-dessus s'approche de la règle du gradient ; par contre si η est trop petit, cette expression devient la méthode Gauss-Newton. Cette dernière est plus rapide, mais devient moins précise quand la solution est proche d'une erreur minimale ; le paramètre est aussi ajusté avec un schéma du type de la formule (3.31).

⁽²⁾ Pour simplifier les notations, on utilise un vecteur a pour l'espace des poids synaptiques que l'on cherche ; spécifier un a correspond à spécifier tous les poids synaptiques.

(5) Structure du réseau Perceptron Multi-Couches

Chaque utilisateur de réseau PMC est naturellement amené à se poser la question délicate de la structure du réseau quel que soit l'algorithme d'apprentissage employé (les formules (3.25), (3.30), (3.32), ...) : nombre de couches, nombre de neurones par couche.

En ce qui concerne le nombre de couches, un résultat important a été apporté par Funahashi, [Funahashi, 1989], qui montre que toute fonction continue peut être approchée par un réseau à une seule couche cachée, à fonctions de transfert linéaires pour les couches d'entrée et de sortie et à fonction de transfert monotone croissante bornée (c'est le cas de la sigmoïde) pour la couche cachée ; ceci pourvu que le nombre de neurones sur la couche cachée soit suffisant.

Tous les algorithmes d'apprentissage sont sensibles au nombre de neurones sur la ou les couches cachées. Le choix du nombre de neurones sur la ou les couches cachées est guidé par la constatation suivante : plus ce nombre de neurones est élevé, plus l'erreur globale du réseau sur les formes apprises diminue, mais si il est trop important, alors le pouvoir de généralisation du réseau décroît (Figure 3.9). Cette remarque est également valable pour le nombre d'itérations à effectuer (Figure 3.10) : après un certain nombre d'itérations, l'erreur d'apprentissage diminue toujours, alors que l'erreur de généralisation, qui diminuait de façon très sensible au début de l'apprentissage, se met à augmenter. On assiste, dans ces deux cas, à un phénomène d'apprentissage « par coeur », où le réseau est même capable d'apprendre du bruit.

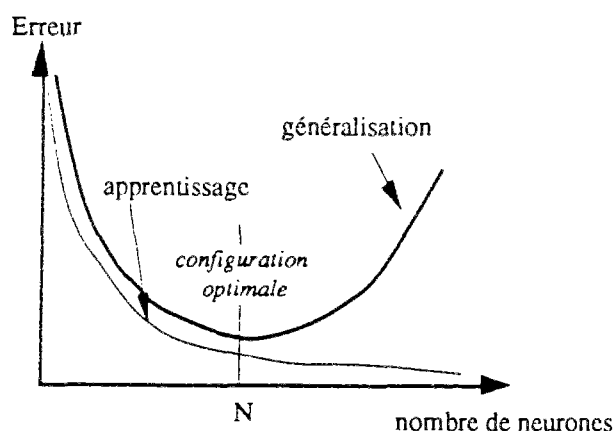


Figure 3.9 : Nombre de neurones optimal

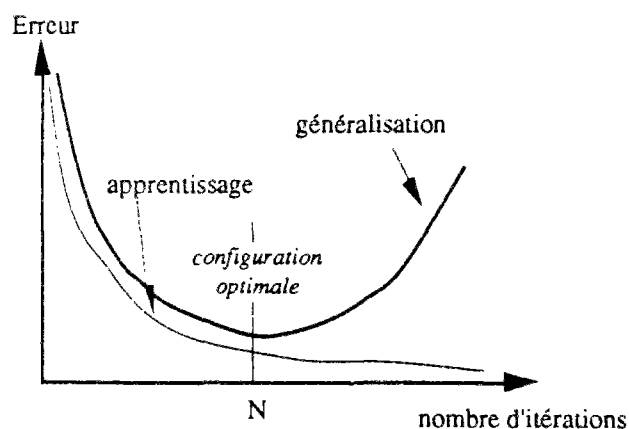


Figure 3.10 : Nombre d'itérations optimal

On peut donc édicter les règles suivantes :

- si apprendre longtemps à un réseau donne toujours une erreur globale importante, le problème est probablement un manque de neurones sur la ou les couches cachées ;
- si un réseau est appris avec un très bon résultat d'apprentissage mais ne peut pas engendrer un résultat correct pour les formes non apprises, celui-ci a probablement trop de neurones sur la ou les couches cachées.

(6) Réseau Perceptron Multi-Couches et rejet (pour le diagnostic)

Un réseau PMC peut facilement implanter le rejet d'ambiguïté qui correspond à la région intermédiaire entre différentes classes.

Soit $\tilde{\zeta}(\omega_i)$ le vecteur cible dédié à la classe ω_i parmi M classes $\omega_1, \dots, \omega_i, \dots, \omega_M$; $\vec{O}^\mu$ le vecteur des sorties réelles du réseau pour la forme d'entrée μ . Une règle de décision avec rejet d'ambiguïté peut se formuler par exemple, [Masson, 1992], comme suit :

$$d(\mu) = i \text{ si et seulement si } \begin{cases} \|\vec{O}^\mu - \tilde{\zeta}(\omega_i)\|^2 = \min_j \|\vec{O}^\mu - \tilde{\zeta}(\omega_j)\|^2 & (j = 1, \dots, M) \text{ et} \\ \|\vec{O}^\mu - \tilde{\zeta}(\omega_i)\|^2 \leq \text{seuil} \end{cases} \quad (3.34)$$

$d(\mu) = 0$ si non (rejet d'ambiguïté)

Cette règle signifie que le point doit être rejeté si ses sorties ne correspondent pas suffisamment avec l'une des sorties cibles. De plus, si l'on adopte un codage cohérent, on est à même de déterminer quelles sont les classes pour lesquelles le système hésite. Un codage efficace consiste par exemple à choisir un neurone de sortie par classe, avec 1 comme la valeur cible si la forme appartient à la classe et -1 si non.

Le rejet de distance ne peut quant à lui être intégré directement dans une telle règle de décision. Il faut utiliser pour ce rejet une structure des réseaux du type RCE (Restricted Coulomb Energy), [Azcarra, 1991] et [Barschdorff, 1991].

Les réseaux Perceptron Multi-Couches, en tant qu'outil pour la reconnaissance des formes, seront appliqués pour les travaux de recherche présentés dans le chapitre suivant ; on utilise ces réseaux d'une part pour l'élaboration de frontières (non-linéaires en général) entre les classes à discriminer (les modes de fonctionnement à diagnostiquer en termes de diagnostic) dans la phase d'analyse et d'autre part pour la classification de celles-ci dans la phase décisionnelle.

CHAPITRE 4

DEVELOPPEMENT D'UN RESEAU CONNEXIONNISTE POUR LE DIAGNOSTIC DE SYSTEMES DE CHAUFFAGE COLLECTIF

Nous présentons dans ce chapitre le développement d'une méthode automatique de diagnostic de défauts en appliquant une méthode connexionniste (les réseaux Perceptron Multi-Couches). Ce travail comporte deux étapes successives :

- *mise au point des réseaux de neurones pour la discrimination de 7 modes de fonctionnement (sans et avec défauts) avec une base de données d'un système de chauffage simulé,*
- *test de la généricité de ces réseaux avec les données des 4 autres systèmes.*

4.1 - Introduction	82
4.2 - Mise au point des structures de réseaux de neurones	84
4.2.1 - Prétraitement de données	84
4.2.1.1 - Présentation des classes (modes de fonctionnement) à discriminer	84
4.2.1.2 - Sélection des paramètres pertinents	85
4.2.2 - Mise au point des structures de réseaux de neurones	93
4.2.2.1 - Introduction	93
4.2.2.2 - Réseaux Uniques	96
4.2.2.2.1 - Apprentissage sans la classe normale (RU11 et RU12)	96
4.2.2.2.2 - Apprentissage avec la classe normale (RU2)	102
4.2.2.2.3 - Conclusion sur les réseaux uniques (RU11, RU12 et RU2)	105
4.2.2.3 - Réseaux Multiples	106
4.2.2.3.1 - Réseaux Multiples 1 (RM1)	106
4.2.2.3.2 - Réseaux Multiples 2 (RM2)	111
4.2.2.3.3 - Conclusion sur les réseaux multiples (RM1+RM2)	118
4.2.2.4 - Conclusion	122
4.3 - Test de la généricité des réseaux développés	123
4.3.1 - Influence d'une modification de la consigne de température intérieure de la période d'occupation	125
4.3.2 - Influence d'un changement du dimensionnement des radiateurs	128
4.3.3 - Influence d'un changement de l'inertie du bâtiment	131
4.3.4 - Influence d'une modification de la consigne de température intérieure de la période d'inoccupation	134
4.3.5 - Conclusion	137

4.1 - INTRODUCTION

Le développement de la méthode s'appuie sur le principe de la reconnaissance des formes. Cette dernière nécessite une phase d'analyse qui comporte trois étapes distinctes (voir paragraphe 3.1.1.1) : (1) le choix d'un espace de représentation (i.e. choix des paramètres pertinents), (2) la création d'une partition de cet espace par les différentes classes (les différents modes de fonctionnement en termes de diagnostic) à discriminer, (3) l'élaboration de frontières entre ces classes.

Les deux premières étapes constituent la partie **prétraitement des données** qui ont pour objectif de créer une base de données utilisable pour la troisième étape. L'ordre dans lequel on effectue ces deux premières étapes peut être différent suivant le problème traité. Dans le cas où l'ensemble des classes à discriminer est a priori connu, la création d'une partition de l'espace est simple. Par conséquent, le problème clé du traitement de données se concentrera sur le choix des paramètres pertinents.

La troisième étape nécessite une phase d'**apprentissage** et une phase de **généralisation**.

Ces trois étapes peuvent s'effectuer de façon successive et répétitive. Autrement dit, si les résultats d'apprentissage et de test ne sont pas satisfaisants vis-à-vis de la performance de diagnostic prédéterminée, il faut réviser le prétraitement de données.

Prétraitement de données

En ce qui concerne notre étude, le prétraitement des données (en particulier choix des paramètres pertinents) s'appuie essentiellement sur une analyse physique des systèmes à diagnostiquer puisqu'on sait a priori que telles variables à tels moments sont des reflets des signatures de tel défaut. Cette analyse devrait donc nous permettre de choisir les paramètres les plus pertinents vis-à-vis des défauts à détecter en éliminant les informations redondantes et inintéressantes, et par conséquent, de mieux comprendre le fonctionnement du système en présence de ces défauts.

Le choix des paramètres pertinents est crucial pour mener à bien la détection et/ou le diagnostic ultérieurs. En effet, à mesure que les techniques de représentation des connaissances se font plus nombreuses et efficaces, le goulot d'étranglement dans la réalisation d'un système basé sur la connaissance est principalement l'acquisition de cette connaissance ; la construction d'une base de connaissance nécessite la mise en oeuvre d'importants moyens humains et matériels puisqu'elle s'appuie essentiellement sur la compréhension du fonctionnement du système à diagnostiquer.

Apprentissage

L'apprentissage, destiné à élaborer les frontières (ou discriminateur) entre les différentes classes à discriminer, utilise quant à lui les réseaux PMC (Perceptron Multi-Couches) présentés précédemment. Pour ce faire, deux algorithmes d'apprentissage, Rétro-Propagation (RP) classique et la RP améliorée du type Levenberg-Marquardt, seront appliqués et ensuite comparés afin de tester leurs performances [Demuth, 1992].

Généralisation

Un des problèmes clés de l'apprentissage dans les domaines de l'intelligence artificielle est celui de la généralisation, c'est-à-dire la capacité à étendre une compétence à des exemples non appris. La généralisation est conditionnée par la capacité de l'apprentissage à élaborer des **représentations internes** adéquates (les représentations internes pour un réseau PMC correspondent aux poids synaptiques reliant les neurones). Par exemple, on espère qu'un système de décision basé sur la reconnaissance des formes classera correctement des formes n'appartenant pas à l'ensemble d'apprentissage.

Or une question essentielle se pose : comment tester la qualité d'une représentation interne générée par apprentissage c'est-à-dire la généricité d'un système de décision ?

Deux méthodes principales sont possibles. La première, surtout utilisée en reconnaissance des formes, est de **séparer l'ensemble des exemples en un ensemble d'apprentissage et un ensemble de test, et de mesurer la performance du système sur l'ensemble de test après apprentissage**. La seconde est d'interpréter la représentation extraite par l'apprentissage (si tant est qu'elle soit interprétable) et de la comparer à un modèle connu par avance comme étant le "bon modèle" ; cette méthode est souvent utilisée dans les domaines relevant classiquement de l'intelligence Artificielle.

Il semble très difficile de formaliser la notion de généralisation. D'un point de vue objectif, aucune généralisation n'est meilleure qu'une autre puisque seules des données extérieures au problème (des exemples supplémentaires) permettront de mesurer sa qualité [Le Cun, 1985]. En ce sens, la généralisation est un problème "mal posé", puisque la donnée des exemples ne suffit pas à le résoudre. On est donc amené à utiliser des critères subjectifs "pertinents" pour qualifier les "bonnes représentations" supposées favoriser la généralisation.

Des études théoriques sur les réseaux de neurones multi-couches, [Denker, 1987], [Le Cun, 1989] et [Patarrello, 1987], montrent que la qualité d'une représentation interne (i.e. généralisation) dépend principalement des trois points suivants :

- le volume total de l'espace de représentation interne (nombre total des réseaux possibles).
Soit $\bar{w}$ un point dans un espace abstrait de poids, le volume total est défini comme :
$$V_0 = \int d\bar{w} \times \rho(\bar{w})$$
 avec $\rho(\bar{w})$ la densité a priori pour chaque $\bar{w}$,
- la partie du volume de l'espace de représentation qui engendre "la bonne généralisation" (i.e. les réseaux de solution),
- le nombre des exemples (ou formes) d'apprentissage.

Si ce volume total V_0 est trop grand et/ou le nombre des exemples d'apprentissage est insuffisant, il y aura un nombre important des représentations internes (ou réseaux), qui sont consistantes avec ces exemples d'apprentissage. En revanche, seulement un petit volume de l'espace (ou un petit nombre de réseaux) donne des bonnes solutions de généralisation.

Par conséquent, la probabilité d'avoir une mauvaise représentation interne est plus grande, celle-ci engendrant une mauvaise généralisation. Pour cette raison, **quand la demande de généricité croît, il faut en principe augmenter la base d'apprentissage afin de pouvoir espérer une bonne généralisation**.

La réflexion ci-dessus nous a conduits à une stratégie de recherche en deux étapes successives.

La première consiste à mettre au point des structures de réseaux de neurones en utilisant une base de données recueillies sur un système de chauffage spécifique simulé pour la détection des 7 modes de fonctionnement (un sans défaut et six avec défauts), c'est-à-dire la discrimination des 7 classes en termes de reconnaissance des formes.

La deuxième est de tester la généricité de ces réseaux en utilisant plusieurs systèmes de chauffage "similaires" simulés afin de déceler si on peut **apprendre une structure de réseaux de neurones avec un seul système de chauffage et étendre cette compétence apprise vers d'autres systèmes "similaires" mais non appris**.

Ce point est très important en terme d'application sur le terrain, car il est souhaitable qu'un outil de détection diagnostic soit appris par un seul système de chauffage (de préférence simulé), mais diagnostique correctement les autres systèmes sans réapprendre cet outil.

4.2 - MISE AU POINT DES STRUCTURES DE RESEAUX DE NEURONES

La mise au point des structures de réseaux pour la discrimination des classes (modes de fonctionnement en diagnostic) nécessite une base de données représentatives et pertinentes. Le travail présenté dans ce paragraphe comprend :

- d'abord l'obtention d'une base de données utilisable pour le développement de réseaux par le prétraitement des données recueillies sur un système de chauffage simulé,
- ensuite l'apprentissage et le test de différentes structures de réseaux Perceptron Multi-Couches avec cette base de données afin d'obtenir une ou des structures de réseaux adéquates à la discrimination des classes apprises.

4.2.1 - PRETRAITEMENT DE DONNEES

En ce qui concerne notre étude, le choix d'un espace de représentation (i.e. choix des paramètres pertinents) constitue le coeur du prétraitement de données puisque les classes à discriminer ont été déjà sélectionnées grâce au travail de recensement et de classification des défauts présenté dans le Chapitre 1. Pour la commodité de l'exposé, on présentera ci-dessous d'abord la création de la partition de l'espace, ensuite la sélection des paramètres pertinents.

4.2.1.1 - Présentation des classes (modes de fonctionnement) à discriminer

Disposant d'une série des défauts typiques de systèmes de chauffage, une base de données, qui représente 7 modes de fonctionnement (sans et avec défauts) d'un système de chauffage spécifique, a été créée par simulation (voir Chapitre 2).

Rapellons que le mode de fonctionnement sans défaut - mode normal de référence - n'est pas forcément optimal mais représentatif vis-à-vis de systèmes de chauffage réels. En particulier le fonctionnement éventuellement non-optimal du régulateur-programmateur n'est pas considéré comme défaut dans la simulation comme dans la réalité. Cela augmente probablement la difficulté de détection des défauts liés au régulateur-programmateur (on le verra plus loin dans le développement et la généralisation des réseaux de neurones).

Les classes (en termes de reconnaissance des formes) correspondant à ces modes de fonctionnement sont reprises du Chapitre 2, et ensuite numérotées et présentées ci-dessous :

- la classe n°0 - classe **normale**,
- la classe n°1 - classe **brûleur**,
- la classe n°2 - classe **échangeur**,
- la classe n°32 - classe **courbe de chauffe**,
- la classe n°41 - classe **relance précoce**,
- la classe n°42 - classe **relance tardive**,
- la classe n°5 - classe **vanne de régulation**.

Nous avons donc effectué, en termes de reconnaissance des formes, une partition de l'espace de représentation pour ces 7 classes à discriminer, cet espace étant divisé en 7 zones : ω_0 , ω_1 , ω_2 , ω_{32} , ω_{41} , ω_{42} , ω_5 . Il faut maintenant construire cet espace de représentation, c'est-à-dire sélectionner des paramètres pertinents.

4.2.1.2 - Sélection des paramètres pertinents

Il est incontestable que le choix des paramètres⁽¹⁾ pertinents du système à surveiller grâce à une connaissance physique du domaine est un point très important pour une bonne performance d'une méthode de diagnostic basée sur la reconnaissance des formes.

Nous avons effectué pour cela trois approches mutuellement complémentaires du traitement de données. Ces approches sont basées sur des connaissances d'expert, l'analyse physique de modes de fonctionnement du système et l'examen graphique de données de simulation.

La première approche consiste à sélectionner les variables "indépendantes" et pertinentes pour éliminer les informations redondantes et inutiles.

La seconde approche consiste à choisir des périodes de mesure dont les informations sur les différents défauts à diagnostiquer sont significatives.

Et la troisième approche est de normaliser les paramètres sélectionnés afin d'obtenir des paramètres numériquement comparables et de faciliter la généralisation de la méthode de diagnostic développée.

(1) Sélection des variables "indépendantes"

Après une phase d'analyse physique défaut par défaut, nous avons sélectionné, parmi les variables mesurables (voir Annexe A2 et Annexe A3), 11 variables supposées porter le plus d'informations sur les défauts à diagnostiquer ; ces variables sont les suivantes :

Boucle Primaire (BP)

Température des Fumées de la Chaudière 1 (TFC1)	(°C)
Température des Fumées de la Chaudière 2 (TFC2)	(°C)
Consommation de Gaz Naturel de la Chaudière 1 (CGNC1)	(m ³ /s)
Consommation de Gaz Naturel de la Chaudière 2 (CGNC2)	(m ³ /s)
Température de Départ Primaire Commun (TDPC)	(°C)
Température de Retour Primaire Commun (TRPC)	(°C)

Boucle Secondaire de Chauffage (BSC)

Température de Départ du Circuit Nord (TDCN)	(°C)
Température de Départ du Circuit Sud (TDCS)	(°C)

Ventilation Mécanique Contrôlée (VMC)

Température Intérieure zone Nord (TIN)	(°C)
Température Intérieure zone Sud (TIS)	(°C)

Environnement

Température EXTérieure (TEXT)	(°C)
-------------------------------	------

Ces variables sont échantillonnées avec un pas de 15 minutes, le pas de simulation du système étant de 1 minute. L'analyse qui suit s'appuiera essentiellement sur ces variables.

⁽¹⁾ Un paramètre est, pour notre étude, défini comme le résultat de filtrage (moyennage par exemple) sur une durée déterminée d'une variable brute issue d'une mesure d'un système réel ou de simulation.

(2) Choix des périodes de mesure optimales (filtrage des données)

Vis-à-vis des défauts à détecter, on ne peut pas utiliser toute la durée (i.e. 200 jours ou 4800 heures) de simulation pour deux raisons principales :

- d'une part, toutes les classes (sans ou avec défauts) sont liées au fonctionnement du système. Il faut donc utiliser des informations (ici des variables mesurables) représentant le fonctionnement du système pour bien détecter ces classes. Pour cette raison, les périodes d'inoccupation (les week-ends et les nuits), où le système de chauffage s'arrête ou marche au régime réduit seulement pour maintenir une température intérieure minimale, doivent être éliminées puisque ces périodes contiennent a priori très peu d'indications (signatures de défaut) de ces classes.
- d'autre part, la période de relance contient a priori les informations les plus significatives pour les défauts de brûleurs (combustion), d'échangeurs (entartrage et encrassement) et de programmeurs d'intermittence à condition que cette période soit suffisamment longue. En effet, si cette période est trop courte, le système peut ne pas atteindre son régime permanent et par conséquent ces informations risquent d'être insuffisantes pour bien détecter ces défauts.

Comment obtenir une période de relance matinale de chauffage suffisamment longue ? Plusieurs éléments entrent en jeu :

D'abord la température intérieure de la période d'occupation du jour précédent, qui conditionne la durée de relance matinale de chauffage de la manière suivante : la relance ne commence que si la température intérieure descend en-dessous d'un certain seuil bas avant la période d'occupation ; par conséquent plus élevée est la température de la période d'occupation du jour précédent, plus longue est la durée nécessaire pour atteindre ce seuil, et donc plus courte est la période de relance matinale de chauffage.

Ensuite la température extérieure et l'ensoleillement, qui conditionnent à la fois la température intérieure et la durée de relance matinale de chauffage ; c'est-à-dire que plus élevée est la température extérieure, plus lentement décroît la température intérieure de la période d'inoccupation et donc plus courte est la durée de relance matinale.

Suivant cette réflexion une phase d'analyse physique et graphique des données (variables mesurables) de simulation a été effectuée ; et nous avons retenu **les jours de la semaine (du lundi au vendredi) où la température extérieure moyenne journalière est inférieure à 10°C** en supposant que les relances pendant ces jours-là sont suffisamment longues. Ceci nous a donc permis d'éliminer les jours pour lesquels la durée de relance matinale de chauffage est trop courte pour la détection ou le diagnostic.

Finalement nous n'avons retenu que **70 jours (i.e. 14 semaines du lundi au vendredi)** pour une saison de chauffage de simulation de 200 jours au total.

Pour ce type de système, un **diagnostic journalier** du mode de fonctionnement est raisonnablement suffisant. Cela signifie que chaque jour constitue un exemple (ou forme) pour chaque classe (mode de fonctionnement). Nous avons donc une base de données utilisables qui contient au total 490 exemples - $70(\text{exemples / classe}) \times 7(\text{classes})$ - pour l'ensemble des 7 classes.

Suivant toujours ce principe de diagnostic journalier, les **paramètres** dit "**journaliers**", qui représentent les composantes des formes (ou exemples) de l'ensemble des classes, seront obtenus par le filtrage de 11 variables (voir ci-dessus) sur les périodes sélectionnées suivant une analyse physique du fonctionnement des systèmes de chauffage.

L'idée directrice est la suivante : le paramètre comportant le plus d'informations sur un défaut quelconque doit être retenu ; autrement dit un défaut devrait être bien discriminé seulement par un paramètre de ce type.

Une journée (i.e. 24 heures) est, en ce sens, divisée en quatre périodes caractéristiques pour un système de chauffage d'un bâtiment de bureaux :

- la première période correspond à la période d'occupation - de deux heures après le début d'occupation jusqu'à la fin d'occupation (de 10h à 18h dans le cas de l'occupation type bureau). Notons que le début d'occupation (8h- 10h) correspond à la période transitoire entre deux régimes de chauffage (i.e. relance matinale et régime normal) et, doit être enlevé pour obtenir une période d'occupation dite de "régime établi". Cette période d'occupation (10h-18h) comporte en particulier les informations liées aux défauts de courbe de chauffe (par exemple la classe n°32 - classe courbe de chauffe).
- la deuxième période correspond à la période d'arrêt - période de fermeture de la vanne de régulation (de 18h à 23h). Cette période, qui débute dès la fin de la période d'occupation, correspond à l'arrêt (ou maintien) du système de chauffage de la période d'inoccupation. Pour éviter la relance matinale dans cette période, on a fixé la fin de cette période à 23h. Cette période comporte principalement les informations liées aux défauts d'étanchéité de la vanne de régulation (par exemple la classe n°5 - classe vanne de régulation).
- la troisième période correspond à la période de relance matinale. Durant cette période, les chaudières fonctionnent en pleine puissance. On peut espérer que les défauts liés au fonctionnement du brûleur ou de l'échangeur peuvent être plus facilement détectables (par exemple la classe n°1 - classe brûleur, la classe n°2 - classe échangeur). D'autre part, il est logique de penser que la période de relance matinale doit être utilisée pour la détection de défauts (par exemple la classe n°41 - classe relance précoce, la classe n°42 - relance tardive) liés au programmeur d'intermittence.
- enfin la quatrième période correspond à une journée entière (24 heures). Cette période est utilisée surtout pour calculer la moyenne journalière de la température extérieure bien que celle-ci ne varie pas en fonction de la présence de défaut. En effet, ce type de système de chauffage est réglé principalement en fonction de la température extérieure ; on sélectionne dans un premier temps cette variable pour étudier son influence sur la performance du diagnostic malgré le fait que celle-ci ne discrimine aucune classe.

(3) Normalisation des paramètres

Les paramètres sont normalisés en fonction de leurs valeurs nominales (moyenne, minimum, etc.) et de leurs variations afin d'obtenir des paramètres numériquement comparables ; la formule employée s'écrit :

$$P_{normalisé} = \frac{P_{filtré} - P_{réf}}{\Delta P} \quad (4.1)$$

Avec :

$P_{filtré}$: moyenne d'une variable sur une période spécifique (période d'occupation, période d'arrêt, etc.)

$P_{réf}$: moyenne ou minimum d'une variable

ΔP : écart entre le maximum et le minimum (ou moyenne) d'une variable

L'objectif en choisissant telle ou telle valeur des paramètres de normalisation $P_{réf}$ et ΔP , est de rendre les paramètres normalisés aussi indépendants que possible du système (bâtiment) à diagnostiquer. Par exemple, on leur donne une moyenne proche de zéro et un écart type proche de 1 quel que soit le système (bâtiment). On espère qu'une telle approche permettra d'utiliser pour différents systèmes le même réseau de neurones.

Pour ce faire, l'adaptation de l'outil de détection diagnostique (réseau de neurones) à différents systèmes de chauffage devrait se limiter à une modification des paramètres de normalisation $P_{\text{réf}}$ et ΔP . Cela impose que les valeurs des paramètres de normalisation peuvent être obtenues simplement à partir de la description du système (valeurs nominales, valeurs de référence, etc.) ou de mesures ponctuelles effectuées sur celui-ci.

Notons que cette approche nécessite souvent plusieurs essais des paramètres de normalisation afin de trouver les valeurs désirées de celles-ci.

(4) Obtention des paramètres pertinents

Après une phase d'analyse graphique de 10 variables sélectionnées (température extérieure non comprise) sur un système de chauffage spécifique et une phase de normalisation, nous avons construit deux espaces de représentation différents afin de comparer leur qualité de classification. Notons que la qualité de ces espaces de représentation sera vérifiée par le test de la performance de généralisation des réseaux après apprentissage ; et on doit, si cette performance n'est pas satisfaisante, créer d'autres espaces pour mieux discriminer les classes sélectionnées.

L'Espace de Représentation 1 (ER1) est construit avec les 5 paramètres suivants :

- T_{fre} (Température moyenne normalisée des fumées de la chaudière 1 durant toute la période de **relance**). On constate sur la Figure 4.1 que ce paramètre aide à la discrimination des classes de **brûleur** et d'**échangeur** des autres classes,
- T_{ioc} (Température intérieure moyenne normalisée de la zone nord de la période d'**occupation** : 10h à 18h). On constate sur la Figure 4.2 que ce paramètre aide à la discrimination de la classe de **courbe de chauffe** des autres classes,
- T_{ies} (Température intérieure moyenne normalisée de la zone nord de 6h à 8h ; 8h signifie la fin de relance). On constate sur la Figure 4.3 que ce paramètre aide à la discrimination des classes de **relance précoce** et de **relance tardive** des autres classes,
- T_{dep} (Température moyenne normalisée de l'eau de **départ** du circuit secondaire nord de 18h à 23h ; cette période correspond à la période d'arrêt ou de réduction de chauffage), qui aide à la discrimination de la classe de **vanne de régulation** des autres classes (Figure 4.4),
- T_{ext} (Température **extérieure** moyenne normalisée de 0h à 24h, c'est-à-dire la valeur quotidienne moyenne) (Figure 4.5).

Et l'Espace de Représentation 2 (ER2) est construit en remplaçant le paramètre T_{ies} par les deux paramètres T_{ie} et T_{is} ; il comprend les 6 paramètres suivants :

- T_{fre} (Température moyenne normalisée des fumées de la chaudière 1 durant toute la période de **relance**), qui aide à la discrimination des classes de **brûleur** et d'**échangeur** des autres classes (Figure 4.1),
- T_{ioc} (Température intérieure moyenne normalisée de la zone nord de la période d'**occupation** : 10h à 18h), qui aide à la discrimination de la classe de **courbe de chauffe** des autres classes (Figure 4.2),
- T_{ie} (Température intérieure moyenne normalisée de la zone nord à 6 h, 2 heures avant la fin de relance), qui aide à la discrimination de la classe de **relance précoce** des autres classes (Figure 4.6),
- T_{is} (Température intérieure moyenne normalisée de la zone nord à 8h, fin de relance), qui aide à la discrimination de la classe de **relance tardive** des autres classes (Figure 4.7),
- T_{dep} (Température moyenne normalisée de l'eau de **départ** du circuit secondaire nord de 18h à 23h), qui aide à la discrimination de la classe de **vanne de régulation** des autres classes (Figure 4.4),

- T_{ext} (Température extérieure moyenne normalisée de 0h à 24h) (Figure 4.5).

En résumé, on dispose d'une base de données comportant au total 490 formes (ou exemples) pour l'ensemble des 7 classes issues de la simulation d'un système de chauffage spécifique, une forme étant représentée soit par 5 composantes (T_{frel} , T_{ioc} , T_{i68} , T_{dep} , T_{ext}), soit par 6 composantes (T_{frel} , T_{ioc} , T_{i6} , T_{i8} , T_{dep} , T_{ext}). En d'autres termes, chacune de ces formes est codée par un point $\bar{x}$ d'un espace de représentation de dimension 5 ou 6.

Deux remarques sur cette base de données :

D'abord, il n'y a aucun paramètre, dans cet espace de représentation, qui discrimine à lui seul la classe normale des autres classes. La question se pose donc : est-il plus difficile de détecter la classe normale que les classes avec défauts ?

Ensuite, cette base de données contient les exemples de lundi. Or, en réalité, il y a souvent des problèmes de régulation-programmation liés aux lundis. On se pose donc la deuxième question : est-il plus difficile de détecter les défauts liés à la régulation-programmation (défauts courbe de chauffe ou défauts relance matinale) le lundi que les autres jours de la semaine ? Là encore le programmeur-régulateur utilisé dans notre simulation fonctionne-t-il correctement ?

On prendra soin de ces questions dans le travail présenté au paragraphe 4.2.2.2.

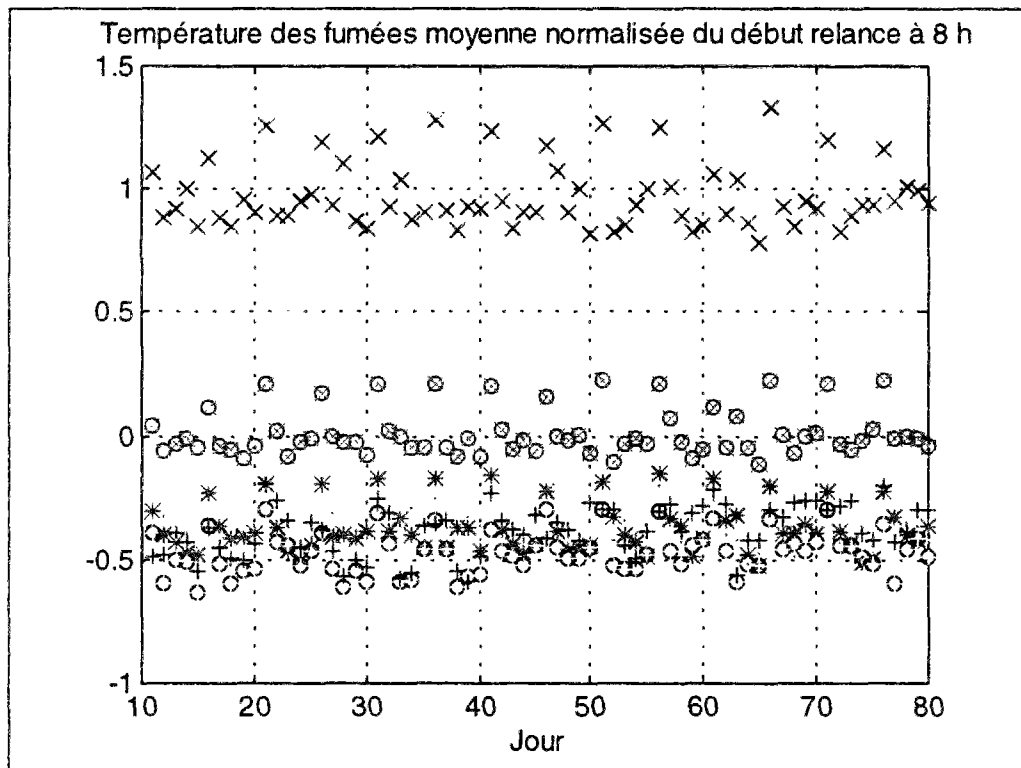


Figure 4.1 Discrimination des classes brûleur et échangeur

Note : Signification des signes des figures 4.1 - 4.7 :

○ => les formes de la classe normale

⊗ => les formes de la classe brûleur

× => les formes de la classe échangeur

* => les formes de la classe courbe de chauffe

⊕ => les formes de la classe relance précoce

+ => les formes de la classe relance tardive

• => les formes de la classe vanne de régulation

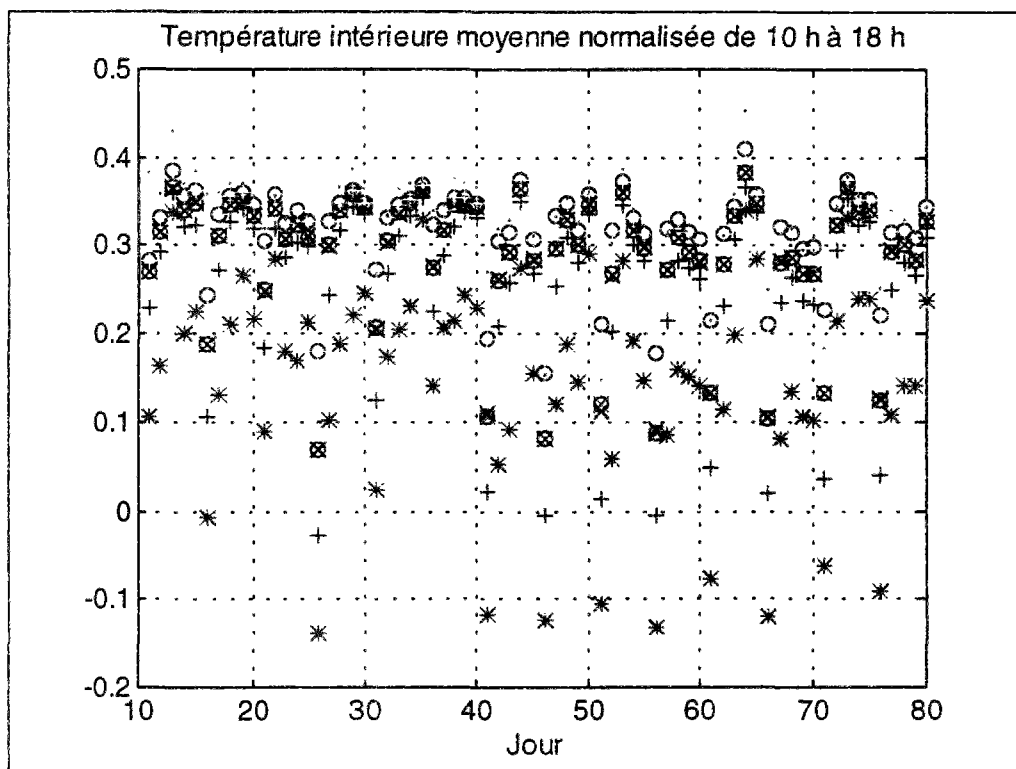


Figure 4.2 Discrimination de la classe courbe de chauffe

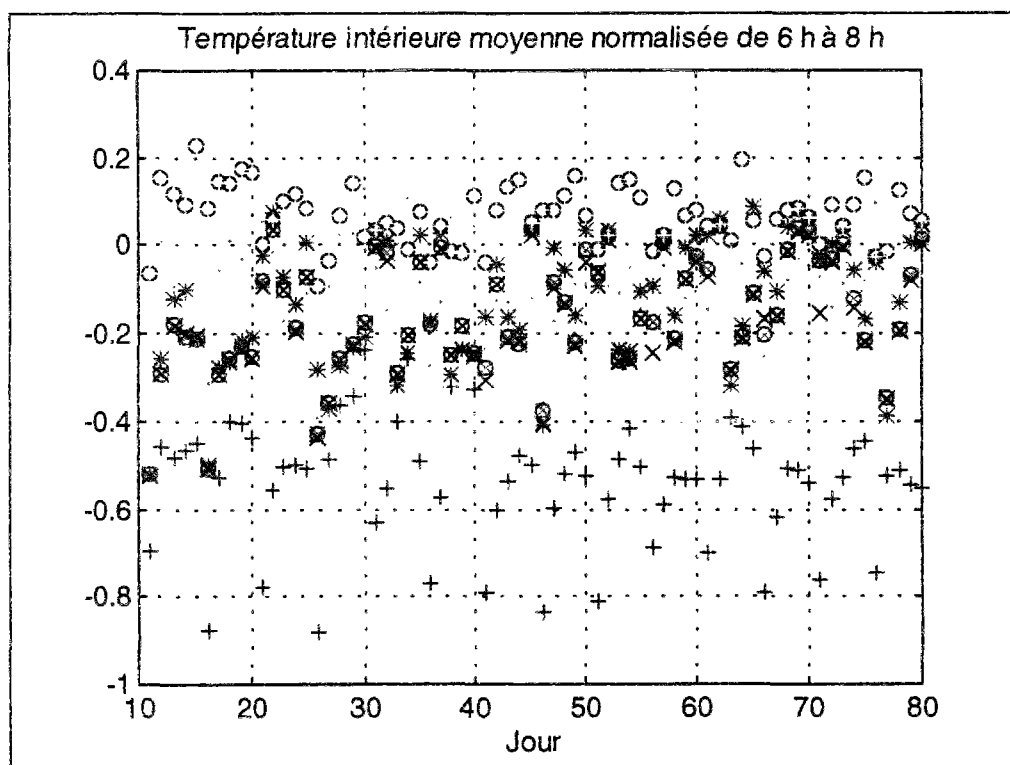


Figure 4.3 Discrimination des classes relance précoce et relance tardive

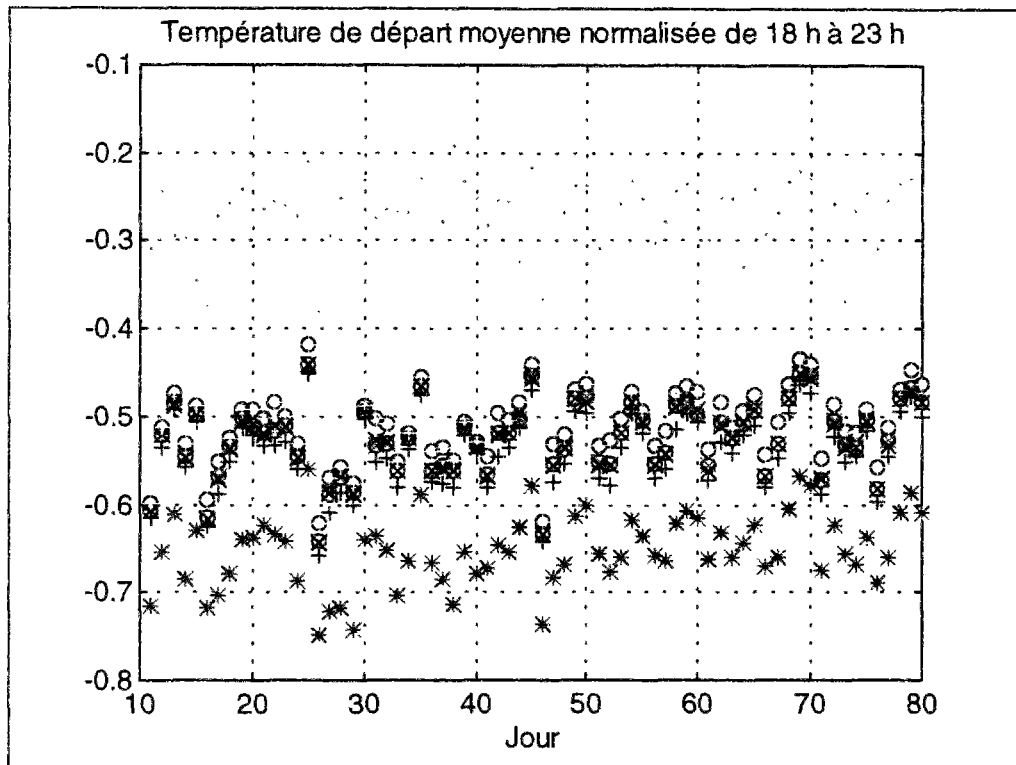


Figure 4.4 Discrimination de la classe vanne de régulation

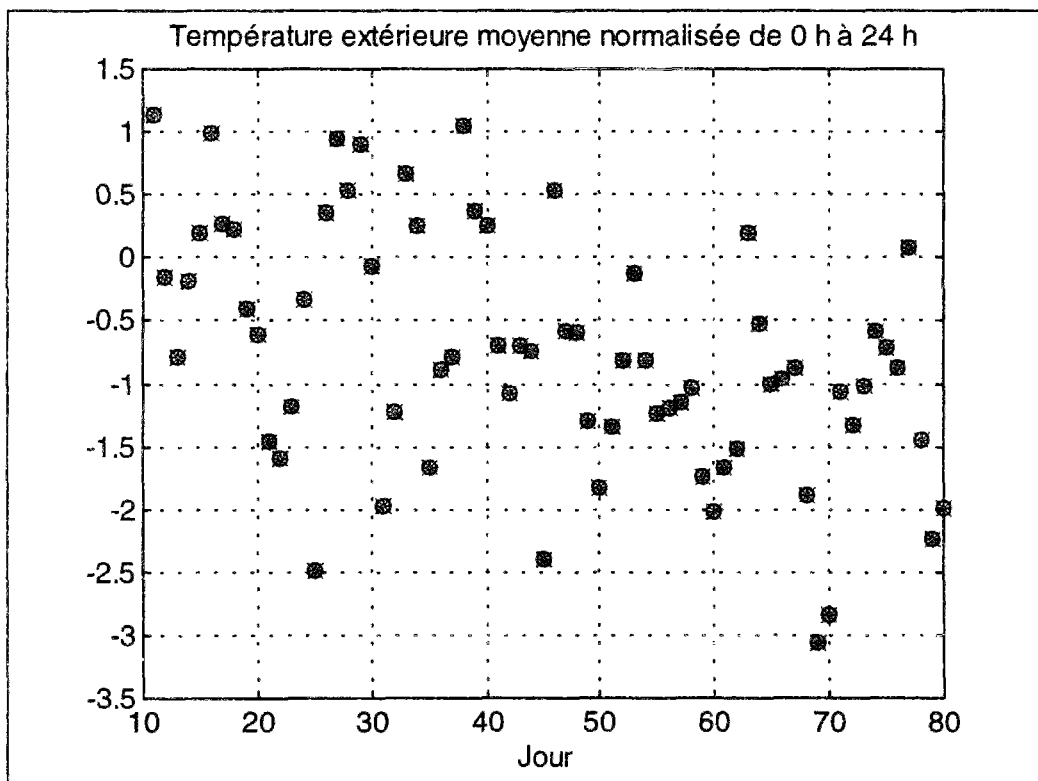


Figure 4.5 Température extérieure moyenne normalisée de 0h à 24h

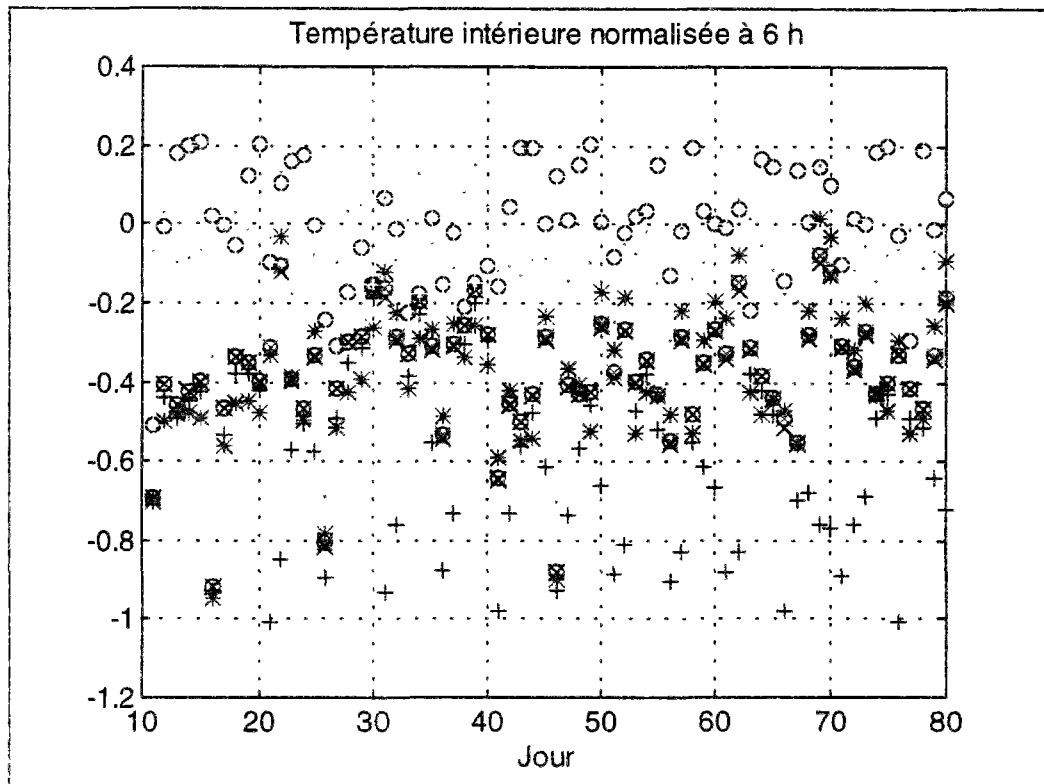


Figure 4.6 Discrimination de la classe relance précoce

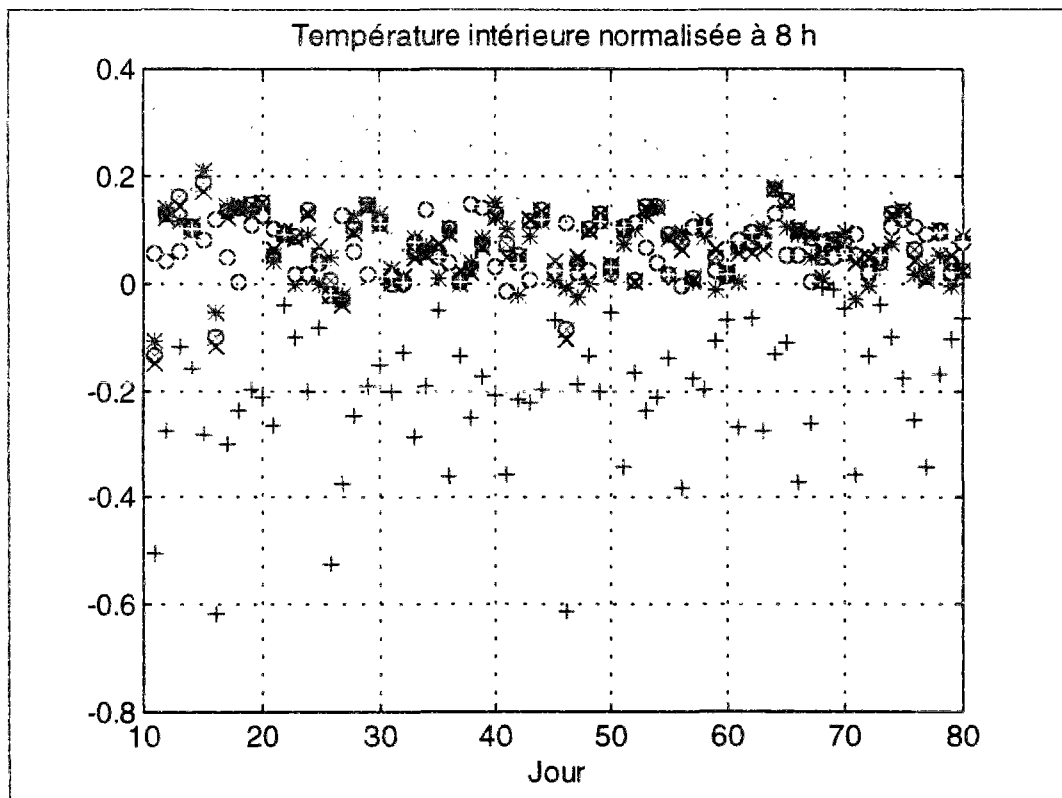


Figure 4.7 Discrimination de la classe relance tardive

4.2 2 - MISE AU POINT DES STRUCTURES DE RESEAUX DE NEURONES

4.2.2.1 - Introduction

Ayant pour objectif de discriminer le mieux possible les 7 classes (une classe sans défaut et six classes avec défauts), nous procéderons aux étapes suivantes afin de mettre au point différentes structures de réseaux et de comparer leur performance de généralisation, [Li, 1995 a], [Li, 1995 b] :

- séparer la base de données (490 formes) en deux parties : un ensemble pour apprendre les réseaux et un ensemble pour tester la performance de généralisation de ceux-ci,
- déterminer la structure de réseaux,
- apprendre les réseaux avec un processus supervisé (apprentissage avec professeur),
- analyser les résultats d'apprentissage et de test.

(1) Séparation de la base de données en un ensemble d'apprentissage et un ensemble de test

La base de données, qui a été créée précédemment et comporte 490 exemples de 7 classes, est divisée en deux ensembles :

- un ensemble d'apprentissage de 245 exemples ($35(\text{exemples} / \text{classe}) \times 7(\text{classes})$),
- un ensemble de test de 245 exemples ($35(\text{exemples} / \text{classe}) \times 7(\text{classes})$).

Notons que ces deux ensembles seront utilisés pour apprentissage et test soit en partie soit en totalité selon la structure de réseau.

Chaque exemple (forme) est composé soit de 5 paramètres ($T_{\text{fre}}, T_{\text{ioc}}, T_{\text{i68}}, T_{\text{dep}}, T_{\text{ext}}$) soit de 6 paramètres ($T_{\text{fre}}, T_{\text{ioc}}, T_{\text{i6}}, T_{\text{i8}}, T_{\text{dep}}, T_{\text{ext}}$).

(2) Détermination de la structure de réseaux

Les réseaux employés ci-après sont tous du type PMC (Perceptron Multi-Couches) et comprennent trois couches successives.

La première couche est la couche d'entrée de réseau. Le nombre des neurones de cette couche est égal au nombre de composantes d'un vecteur $\vec{x}$ qui code une forme augmenté d'un neurone seuil. Par exemple, un vecteur $\vec{x}$ composé de 5 composantes correspond à 6 ($= 5+1$) neurones dans la couche d'entrée.

La deuxième couche désigne la couche cachée. La procédure de détermination du nombre des neurones cachés consiste à :

- d'abord apprendre différents réseaux avec un nombre des neurones décroissant ou croissant,
- ensuite calculer la somme des erreurs quadratiques de l'ensemble d'apprentissage entre les sorties cibles et les sorties réelles de réseau ; cette somme décroît en principe avec l'augmentation du nombre des neurones cachés,
- enfin appliquer un ensemble de données (non appris par le réseau) pour tester la généralisation du réseau. Un compromis entre les erreurs d'apprentissage et la performance de la généralisation devra être trouvé ; ceci devrait nous permettre de trouver un nombre optimal de neurones cachés.

La troisième couche correspond à la couche de sortie. La détermination de la taille (nombre des neurones) de la couche de sortie consiste à choisir un neurone de sortie par classe d'apprentissage, avec +1 comme valeur cible si l'exemple (forme) appartient à la classe et -1 sinon (Tableau 4.1). Ce codage des formes de sortie du réseau devra nous permettre de déterminer, pour un exemple non appris, quelles sont les classes pour lesquelles le système de décision hésite, donc un rejet d'ambiguïté.

Les structures des réseaux étudiées sont les suivantes :

- Réseaux Uniques (RU), c'est-à-dire un seul réseau pour la discrimination de l'ensemble de 7 classes
- Réseaux Multiples (RM) tels que des classes (ou sous-ensembles des classes) sont discriminées par une combinaison des réseaux spécifiques.

(3) Apprentissage de réseaux avec un processus supervisé (apprentissage avec professeur)

Deux algorithmes d'apprentissage (Rétro-Propagation classique, formules (3.26) - (3.29) ; Optimisation Levenberg-Marquardt, formules (3.32) et (3.33)) seront appliqués et ensuite comparés afin de tester leur performance.

Le codage des vecteurs de sortie désirés est réalisé en fonction du nombre des classes d'apprentissage utilisées. Dans le cas de M classes d'apprentissage ($\omega_1, \dots, \omega_M$), le codage pour les formes de chaque classe est présenté dans le Tableau 4.1.

Classe (ω_i)	Vecteur de sortie désiré ($\bar{\zeta}_{\omega_i}$ avec M composantes)
ω_1	$\bar{\zeta}_{\omega_1} = [+1 \ -1, \dots, -1]$
.	.
.	.
.	.
ω_M	$\bar{\zeta}_{\omega_M} = [-1, \dots, -1 \ +1]$

Tableau 4.1 Vecteurs de sortie désirés pour M classes d'apprentissage

(4) Présentation et analyse des résultats d'apprentissage et de test

Les résultats d'apprentissage seront analysés et présentés sous forme de SEQ (Somme des Erreurs Quadratiques entre les sorties désirées et les sorties réelles de l'ensemble d'apprentissage).

Deux méthodes de présentation

Les résultats de test de réseaux seront présentés :

- d'une part sous forme de tableau indiquant les taux de bonne classification, de mauvaise classification et de non-classification (rejet d'ambiguïté) d'une classe ou un groupe de classes (voir comme exemple le Tableau 4.3),
- d'autre part sous forme de matrice de confusion indiquant, pour chaque classe, à la fois la probabilité de classer dans telle ou telle autre classe et la probabilité de non-classification (voir comme exemple le Tableau 4.4).

Deux règles d'analyse

Deux règles seront appliquées pour choisir un réseau "optimal" (un apprentissage "optimal" ou une représentation interne "optimale").

La première est la règle de décision avec rejet d'ambiguïté. Cette règle, ayant pour objectif d'analyser les vecteurs de sortie réels de réseau, s'écrit :

$$\forall J = 1, \dots, M \quad \begin{cases} \text{si } O_J^\mu \geq s \text{ alors } O_J^\mu = O_J'^\mu = +1 \\ \text{si } O_J^\mu < s \text{ alors } O_J^\mu = O_J'^\mu = -1 \end{cases}$$

$$d(\mu) = i \text{ si et seulement si } \|\tilde{O}'^\mu - \tilde{\zeta}_{\omega_i}\|^2 \equiv 0 \quad (i = 1, \dots, M) \quad (4.2)$$

$$d(\mu) = 0 \text{ si non (rejet d'ambiguïté)}$$

Avec :

$\tilde{\zeta}_{\omega_i}$: vecteur de sortie désiré dédié à la classe ω_i parmi les M classes d'apprentissage (voir le Tableau 4.1)

$\tilde{O}^\mu$: vecteur de sortie réel (M composantes - $O_1^\mu, \dots, O_M^\mu$ - qui correspondent à M classes d'apprentissage) d'une forme d'entrée μ (i.e. un exemple)

s : seuil qui sera fixé en fonction de vecteurs de sortie réels

Notons que $d(\mu)=i$ traduit la décision d'associer une forme μ à la classe ω_i et que $d(\mu)=0$ traduit la décision de non-classification (rejet d'ambiguïté) d'une forme μ .

La seconde est employée pour trouver la "meilleure" représentation interne de réseau parmi différents apprentissages (i.e. différentes initialisations des poids synaptiques) vis-à-vis d'un nombre quelconque de neurones cachés par une comparaison du résultat d'apprentissage avec celui de test.

Cette règle⁽²⁾ est édictée, pour une structure donnée, comme suit :

- si deux apprentissages (c'est-à-dire deux différentes initialisations des poids synaptiques) donnent la même SEQ mais que le premier engendre un meilleur taux de reconnaissance de test, on retiendra ce premier comme représentation interne de réseau,
- si le premier donne une SEQ un peu plus faible mais un taux de reconnaissance de test nettement plus bas, ce sera le deuxième qui devra être retenu.

Deux probabilités

Deux probabilités seront également employées pour analyser et comparer la performance de généralisation de différentes structures de réseaux :

- Probabilité de Reconnaissance (PR) qui est définie comme la probabilité de bonne classification de l'ensemble des classes comportant ou non des défauts (7 classes au total)
- Probabilité de Fausse Alarme (PFA) qui est définie comme la probabilité de détecter, pour l'ensemble des classes comportant ou non des défauts, un défaut qui ne se présente pas.

Pour des raisons de clarté et pour démontrer la méthodologie de travail, la présentation de ces différentes structures se fait de manière chronologique.

⁽²⁾ Au cours de l'apprentissage et du test de différents réseaux, on a remarqué qu'un meilleur apprentissage (ici la SEQ) n'engendrait pas forcément un meilleur score de généralisation (ici le taux de reconnaissance de test), c'est pourquoi on applique, tout au long du travail qui suit, cette règle pour le choix d'une représentation interne optimale.

4.2.2.2 - Réseaux Uniques

Comme on a pu remarquer, à la fin du paragraphe 4.2.1.2 (voir la première remarque), qu'il n'y a pas de paramètre qui discrimine la classe normale des autres classes, l'apprentissage de réseau se fait de deux manières différentes afin de trouver une représentation interne adéquate de réseau :

- apprentissage sans la classe normale mais test avec la classe normale,
- apprentissage et test tous deux avec la classe normale.

4.2.2.2.1 - Apprentissage sans la classe normale (RU11 et RU12)

En utilisant une stratégie d'apprentissage sans la classe normale, deux espaces de représentation des formes (ER1 et ER2) sont employés pour la construction de réseaux de neurones.

(1) Présentation des réseaux utilisant l'espace ER1 (5 paramètres : T_{fre} , T_{ioc} , T_{i68} , T_{dep} , T_{ext})

Base d'apprentissage et base de test

Un ensemble d'apprentissage est composé de 6 classes avec défauts (brûleur, échangeur, courbe de chauffe, relance précoce, relance tardive, vanne de régulation) soit 210 exemples ($35(\text{exemples} / \text{classe}) \times 6(\text{classes})$). Un ensemble de test est composé de 7 classes (6 classes avec défauts + 1 classe normale) soit 245 exemples ($35(\text{exemples} / \text{classe}) \times 7(\text{classes})$).

Structure de réseaux

La structure de ces réseaux est présentée schématiquement par la Figure 4.8.

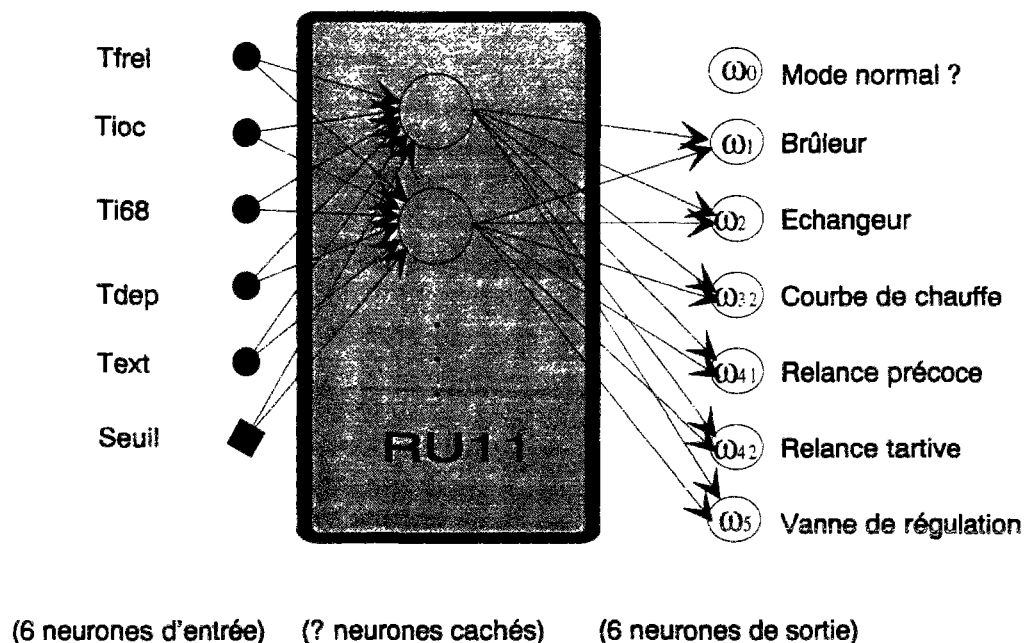


Figure 4.8 Structure des réseaux uniques 11 (RU11)

Le nombre des neurones d'entrée est égal 6 (1 neurone seuil + 5 neurones correspondant à 5 composantes d'une forme). Le nombre des neurones de sortie est de 6 qui correspondent à 6 classes d'apprentissage ; les vecteurs forme de sortie pour ces 6 classes d'apprentissage sont donc de dimension 6.

Nous avons fait varier le nombre des neurones cachés de 3 à 9 en utilisant un algorithme d'apprentissage du type Rétro-Propagation classique, le nombre d'itérations de chaque apprentissage est d'environ 400. Les résultats d'apprentissage et de test sont présentés ci-après.

Résultat d'apprentissage

La SEQ (Somme des Erreurs Quadratiques) de l'ensemble d'apprentissage entre les vecteurs de sortie désirés et les vecteurs de sortie réels, pour chaque nombre des neurones cachés, est présentée par le Tableau 4.2 :

Nombre des neurones cachés	3	4	5	6	7	8	9
SEQ	30.0	8.2	8.2	8.2	8.2	8.1	8.1

Tableau 4.2 SEQ de 210 formes (6 classes) d'apprentissage - RU11

Commentaires sur l'apprentissage :

- un réseau du type RU11 avec 3 neurones cachés donne un mauvais score d'apprentissage ; ceci est dû très probablement à une insuffisance du nombre des neurones cachés.
- quand le nombre des neurones cachés augmente de 4 à 9, la SEQ ne change pas, sachant que la valeur de 8.2 (ou 8.1) signifie qu'un seul exemple parmi 210 est mal classé (la règle de classification est la formule (4.2) avec le seuil $s=0.5$). Est-il utile d'accroître ce nombre des neurones cachés de 4 à 9 ? On répondra avec le résultat de test.

Résultat de test

Le Tableau 4.3 présente les différents taux (%) de classification (bonne, mauvaise et non) pour la classe normale d'une part et pour l'ensemble de 6 classes avec défauts d'autre part.

Commentaires sur le test (voir le Tableau 4.3) :

- le taux de bonne classification de 6 classes avec défauts, qui croît légèrement avec l'accroissement du nombre des neurones cachés de 3 à 8, commence à diminuer si ce nombre dépasse la borne de 8 (phénomène de l'apprentissage par coeur ?),

- le taux de bonne classification de la classe normale, qui décroît avec l'augmentation du nombre des neurones cachés de 3 à 5, se met à augmenter jusqu'à 8 neurones cachés et de nouveau à diminuer si ce nombre dépasse la borne de 8. Le meilleur score de test est donné, semble-t-il, par un réseau avec 8 neurones cachés. En tout cas, ces taux de classification sont tous loin d'être satisfaisants ; le meilleur taux étant de 40 % seulement dans le cas de 8 neurones cachés.

Classe	Taux de classification	Nombre des neurones cachés						
		3	4	5	6	7	8	9
la classe normale (35 formes)	Bonne (%)	34.3	31.4	14.3	20.0	37.1	40.0	34.3
	Mauvaise (%)	65.7	68.6	85.7	80.0	62.9	60.0	65.7
	Non (%)	0	0	0	0	0	0	0
6 classes avec défauts (210 formes)	Bonne (%)	91.0	93.3	94.8	95.2	96.2	96.7	93.3
	Mauvaise (%)	2.9	4.3	2.9	1.4	1.9	1.0	1.9
	Non (%)	6.1	2.4	2.4	3.3	1.9	2.4	4.8

Tableau 4.3 Taux de classification de 245 formes (7 classes) de test - RU11

Notes :

Après avoir examiné de nombreux vecteurs de sortie réels, le seuil dans la règle de décision, voir la formule (4.2), a été fixé à 0.5 ; ce dernier sera utilisé tout au long de l'analyse de résultats qui suit, afin d'avoir un standard comparable entre les différentes structures de réseaux.

Comme les formes (i.e. exemples) de la classe normale (ω_0) n'ont pas été utilisées pour l'apprentissage, quand on les applique aux réseaux (RU11), leurs vecteurs de sortie doivent être classés dans une des classes avec défaut ($\omega_1, \omega_2, \omega_{32}, \omega_{41}, \omega_{42}, \omega_5$) (i.e. les formes mal classées) ou rejetés (i.e. les formes non classées). Or, ces formes de la classe normale rejetées ne devraient pas être considérées comme des formes non classées, mais au contraire, comme des formes bien classées dans la classe normale. Cela signifie également, dans ce cas particulier, qu'il n'y a pas de forme non classée pour la classe normale.

La présentation du résultat de test est divisée en deux parties : la première est pour la classe normale (ω_0), la seconde est pour les 6 classes avec défauts ($\omega_1, \omega_2, \omega_{32}, \omega_{41}, \omega_{42}, \omega_5$), afin de mieux analyser les deux types de résultat qui ont des caractères différents.

Pour mieux analyser ce test, On a étudié en détail la classification de l'ensemble de test avec un réseau comportant 8 neurones cachés. Pour chaque classe, les probabilités de classification dans telle ou telle autre classe et la probabilité de non-classification sont présentées par une matrice dite matrice de confusion (voir le Tableau 4.4). La probabilité de reconnaissance (PR) et la Probabilité de Fausse Alarme (PFA) sont également présentées dans ce tableau.

Commentaires sur le test du réseau RU11 avec 8 neurones cachés (voir le Tableau 4.4) :

- on voit clairement sur ce tableau que la mauvaise classification de la classe normale est entièrement liée aux classes représentant des défauts du régulateur-programmateur. Par ailleurs, la classe de relance précoce est parfois classée dans la classe de courbe de chauffe. Ces mauvais résultats, liés non seulement à la classe normale mais aussi à la classe de relance précoce, nous ont fait penser à la qualité de l'espace de représentation (ER1) qui code l'ensemble des formes, c'est-à-dire la représentabilité de ces paramètres vis-à-vis des classes (modes de fonctionnement sans ou avec défauts) à discriminer.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.40	0	0	0.11	0.32	0.17	0	0
Brûleur	0	0.97	0	0	0	0	0	0.03
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0	0	0	1.00	0	0	0	0
Relance précoce	0	0	0	0.06	0.91	0	0	0.03
Relance tardive	0	0	0	0	0	0.97	0	0.03
Vanne de Régulation	0	0	0	0	0	0	0.94	0.06
PR (*)	0.88							
PFA (**)	0.09							

Tableau 4.4 Probabilités de classification de chaque classe (35 formes) de test - RU11 avec 8 neurones cachés

Note : (*) PR (Probabilité de Reconnaissance) = $\frac{0.40 + 0.97 + 1.00 + 1.00 + 0.91 + 0.97 + 0.94}{7} = 0.88$

(**) PFA (Probabilité de Fausse Alarme) = $\frac{(0.11 + 0.32 + 0.17) + 0 + 0 + 0 + 0.06 + 0 + 0}{7} = 0.09$

- en effet, dans cet espace de représentation, on a utilisé la température intérieure moyenne durant la période de relance matinale (i.e. du début à la fin de la période de relance) pour la détection des deux classes différentes (relance précoce et relance tardive) (voir la Figure 4.3). Or la signature du défaut de relance précoce se manifeste probablement plutôt au début de relance matinale (6h du matin par exemple), et celle du défaut de relance tardive se manifeste plutôt à la fin de relance (8h du matin par exemple). Pour cette raison, on utilisera ci-après deux paramètres différents (température intérieure à 6h et température intérieure à 8h) (voir les Figures 4.6 et 4.7), au lieu d'un seul paramètre (température de 6h à 8h), pour la classification des deux classes (relance précoce et relance tardive) ; on aura donc le deuxième espace de représentation (ER2) pour la construction des réseaux dits RU12 mais l'apprentissage toujours sans la classe normale.
- les fausses alarmes sont principalement liées à la classe normale, c'est-à-dire ce système de diagnostic signale souvent un défaut de régulation (problèmes de courbe de chauffe ou de relance matinale) qui n'existe pas.
- la détection des classes avec défauts est satisfaisante (taux de reconnaissance de 91 % à 100 %).

(2)Présentation des réseaux construits par l'ER2 (6 paramètres : T_{fre} , T_{ioc} , T_{i6} , T_{i8} , T_{dep} , T_{ext})

Base d'apprentissage et base de test

Comme dans le cas de l'ER1, nous avons un ensemble d'apprentissage de 210 exemples ($35(exemples / classe) \times 6(classes)$) et un ensemble de test de 245 exemples ($35(exemples / classe) \times 7(classes)$).

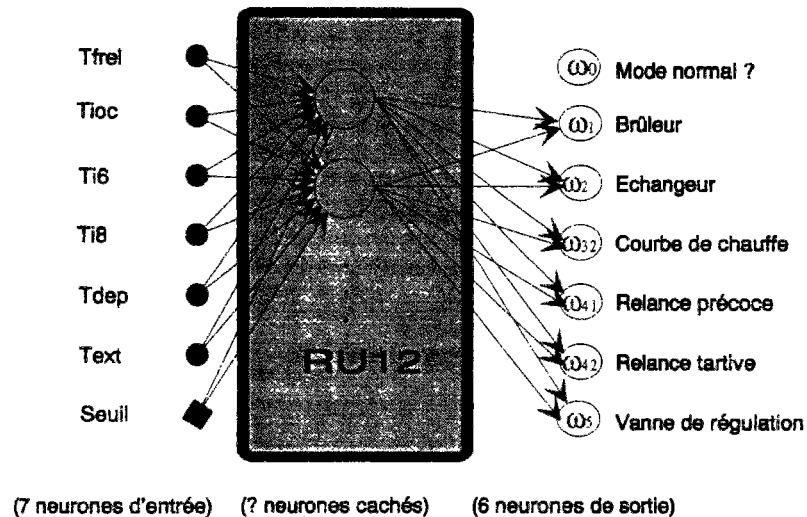


Figure 4.9 Structure des réseaux uniques 12 (RU12)

Structure de réseaux

La structure de ces réseaux est présentée schématiquement par la Figure 4.9.

Le nombre des neurones d'entrée est égal 7 (1 neurone seuil + 6 neurones correspondant à 6 composantes d'une forme). Le nombre des neurones de sortie, ainsi que les vecteurs formes de sortie sont les mêmes que pour les réseaux RU11.

Nous avons fait varier le nombre des neurones cachés de 5 à 8 mais toujours avec un algorithme d'apprentissage du type Rétro-Propagation classique, le nombre d'itérations nécessaire pour la convergence de chaque apprentissage est, cette fois-ci, d'environ 150, c'est-à-dire presque le tiers de celui nécessaire aux réseaux RU11. Autrement dit le temps d'apprentissage est sensiblement réduit par rapport aux réseaux RU11. Les résultats d'apprentissage et de test sont présentés ci-dessous.

Résultat d'apprentissage

La Somme des Erreurs Quadratiques (SEQ) de l'ensemble d'apprentissage entre les vecteurs de sortie désirés et les vecteurs de sortie réels, pour chaque nombre des neurones cachés, est présentée dans le Tableau 4.5 :

Nombre des neurones cachés	5	6	7	8
SEQ	4.3	6.8	4.3	5.3

Tableau 4.5 SEQ de 210 formes (6 classes) d'apprentissage - RU12

Commentaires sur l'apprentissage :

- la valeur de SEQ diminue par rapport à celle des réseaux RU11 (voir le Tableau 4.2) et varie peu avec le changement du nombre des neurones cachés de 5 à 8. En effet, la valeur d'environ 4 de SEQ signifie qu'un seul exemple parmi 210 est non classé (la règle de classification est toujours la formule (4.2) que ce soit pour l'apprentissage ou le test).

Résultat de test

Le résultat de test est présenté dans le Tableau 4.6.

Classe (s)	Taux de classification	Nombre des neurones cachés			
		5	6	7	8
Classe normale (35 formes)	Bonne (%)	54.3	65.7	74.3	62.9
	Mauvaise (%)	45.7	34.3	25.7	37.1
	Non (%)	0	0	0	0
6 classes avec défauts (210 formes)	Bonne (%)	96.7	96.7	97.1	97.1
	Mauvaise (%)	1.4	0.5	1.0	1.0
	Non (%)	1.9	2.8	1.9	1.9

Tableau 4.6 Taux de classification de 245 formes (7 classes) de test - RU12

Commentaires sur le test :

- le score de test pour les classes comportant des défauts ne varie pratiquement pas avec le changement du nombre des neurones cachés, et est très satisfaisant quel que soit le nombre des neurones cachés. En revanche le score de test pour la classe normale est toujours très médiocre bien que celui-ci soit presque doublé par rapport aux résultats des réseaux RU11 (comparaison du Tableau 4.6 avec le Tableau 4.3).

Comme dans le cas des réseaux RU11, on a étudié en détail (matrice de confusion) la classification de l'ensemble de test avec un réseau RU12 qui comporte 7 neurones cachés. Pour chaque classe, les probabilités de classification dans telle ou telle autre classe et la probabilité de non-classification, ainsi que les probabilités PR et PFA sont présentées dans le Tableau 4.7.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.74	0.06	0	0.09	0.11	0	0	0
Brûleur	0	0.91	0	0.03	0	0	0	0.06
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0	0	0	0.97	0.03	0	0	0
Relance précoce	0	0	0	0	1.00	0	0	0
Relance tardive	0	0	0	0	0	1.00	0	0
Vanne de Régulation	0	0	0	0	0	0	0.94	0.06
PR	0.94							
PFA	0.05							

Tableau 4.7 : Probabilités de classification de chaque classe (35 formes) de test - RU12 avec 7 neurones cachés

Commentaires sur le test d'un réseau RU12 avec 7 neurones cachés :

- la mauvaise classification de la classe normale est toujours liée à des classes du régulateur-programmateur (courbe de chauffe et relance précoce mais non relance tardive cette fois ci) bien que la bonne classification de cette classe soit sensiblement meilleure (74 % contre 40 % dans le cas du RU11 avec 8 neurones cachés - voir également le Tableau 4.4).

(3) Conclusion sur les réseaux uniques appris sans la classe normale

En comparant les résultats d'apprentissage avec ceux de test (voir les tableaux 4.2-4.3, 4.5-4.6), on peut remarquer que même si les résultats d'apprentissage (ici les SEQ) ne changent pratiquement pas avec la variation du nombre des neurones cachés, les résultats de test peuvent être très différents selon le nombre des neurones cachés ; ceci est également vrai pour différentes initialisations des poids avec le même nombre des neurones cachés. Une question s'est posée dès le début : le résultat d'apprentissage est-il le seul critère pour choisir un réseau, c'est-à-dire choisir une représentation interne de réseau (poids synaptiques) ? Après de nombreux essais, la réponse est non ! On a donc choisi les réseaux RU11 et RU12 en faisant non seulement une analyse de résultats d'apprentissage mais aussi une analyse de résultats de test.

Les résultats de classification pour l'ensemble de test des classes comportant des défauts sont satisfaisants quel que soit l'espace de représentation utilisé (l'ER1 ou l'ER2) (voir aussi les tableaux 4.3 et 4.6).

Les réseaux RU12, qui utilisent l'espace de représentation ER2, engendrent un score (résultat de test) nettement meilleur pour la classification de la classe normale, par rapport aux réseaux RU11 qui utilisent l'ER1. Cela confirme l'importance de la qualité de l'espace de représentation, c'est-à-dire que l'utilisation des deux paramètres (T_{i6} et T_{i8}) au lieu d'un seul paramètre (T_{i68}) engendre un espace de représentation qui sépare mieux les 7 classes à discriminer. En conséquence, on utilisera, tout le long du travail qui suit, les paramètres de l'ER2 pour développer d'autres structures de réseaux. Tout d'abord, on l'appliquera pour faire un apprentissage des réseaux uniques avec la classe normale.

Jusqu'à maintenant, les différentes structures de réseaux neuronaux ont été apprises sans utilisation de la classe normale. On voit à travers les tableaux 4.4 et 4.7 que la reconnaissance de la classe normale n'est pas satisfaisante (le meilleur taux est de 74 %). Toujours dans l'objectif de rechercher la meilleure représentation interne de réseau (i.e. les poids synaptiques appris) pour la classification, on essaiera dans le paragraphe 4.2.2.2 une stratégie d'apprentissage avec la classe normale.

4.2.2.2 - Apprentissage avec la classe normale (RU2)

Base d'apprentissage et base de test

En introduisant la classe normale pour l'apprentissage, l'ensemble d'apprentissage et l'ensemble de test sont tous deux composés de 7 classes (normale, brûleur, échangeur, courbe de chauffe, relance précoce, relance tardive, vanne de régulation) de 245 exemples ($35(\text{exemples} / \text{classe}) \times 7(\text{classes})$).

Structure de réseaux

La structure de ces réseaux dits Réseaux Uniques 2 (RU2) est présentée schématiquement par la Figure 4.10.

Le nombre des neurones d'entrée, pour les réseaux RU2, est égal à 7 (1 neurone seuil + 6 neurones correspondant à 6 composantes d'une forme) comme dans le cas des réseaux RU12. Le nombre des neurones de sortie est de 7 qui correspondent à 7 classes d'apprentissage ; les vecteurs forme de sortie pour ces 7 classes d'apprentissage sont donc de dimension 7.

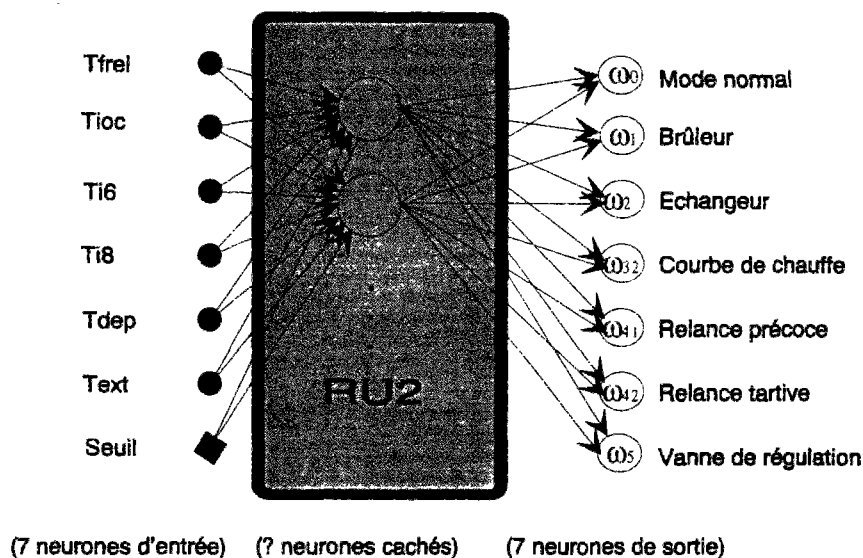


Figure 4.10 Structure des réseaux uniques 2 (RU2)

Nous avons fait varier le nombre des neurones cachés de 3 à 7 avec deux algorithmes d'apprentissage (Rétro-Propagation classique et optimisation de Levenberg-Marquardt). Le nombre d'itérations est d'environ 400 avec le premier algorithme, et compris entre 20 et 70 avec le second. Pour le même nombre de neurones cachés, le rapport du temps de calcul d'un apprentissage ($\frac{\text{RP classique}}{\text{Optimisation LM}}$) est d'environ 10. Comme l'algorithme Levenberg-Marquardt présente une meilleure performance, on ne présentera ci-après que les réseaux (apprentissage et test) appris avec cet algorithme.

Résultat d'apprentissage

La SEQ de l'ensemble d'apprentissage entre les vecteurs de sortie désirés et les vecteurs de sortie réels, pour chaque nombre des neurones cachés, est présentée par le Tableau 4.8 :

Nombre des neurones cachés	3	4	5	6	7
SEQ	54.7	4.9	4.4	4.4	2.0

Tableau 4.8 SEQ de 245 formes (7 classes) d'apprentissage - RU2

Commentaires sur l'apprentissage :

- un réseau RU2 avec 3 neurones cachés donne toujours un mauvais score d'apprentissage. Quand le nombre de neurones cachés augmente de 4 à 7, la SEQ ne change pratiquement pas, et reste autour de 4 ; cette valeur signifie toujours, comme dans le cas de réseaux RU12, un seul exemple (parmi 245 exemples d'apprentissage) non classé.

Résultat de test

Le résultat de test est toujours divisé en deux parties : la première est pour la classe normale (ω_0), la seconde est pour les 6 classes avec défauts (ω_1 , ω_2 , ω_{32} , ω_{41} , ω_{42} , ω_5) ; elles sont présentées sous forme de taux de classification dans le Tableau 4.9.

Classe (s)	Taux de classification	Nombre des neurones cachés				
		3	4	5	6	7
Classe normale (35 formes)	Bonne (%)	71.4	71.4	91.4	91.4	85.7
	Mauvaise (%)	17.2	8.6	5.7	0	8.6
	Non (%)	11.4	20	2.9	8.6	5.7
6 classes avec défauts (210 formes)	Bonne (%)	87.6	91.9	96.2	94.8	94.8
	Mauvaise (%)	5.7	2.9	2.4	1.4	1.0
	Non (%)	6.7	5.2	1.4	3.8	4.2

Tableau 4.9 Taux de classification de 245 formes (7 classes) de test- RU2

Comme dans les cas précédents (RU11 et RU12), on présente (voir le Tableau 4.10) la matrice de confusion (probabilités de classification) pour le réseau RU2 avec 5 neurones cachés qui donne, semble-t-il, le meilleur taux de classification.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.91	0	0	0	~~0.06~~	0	0	0.03
Brûleur	0.03	0.91	0	0	0	~~0.03~~	0	0.03
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0	0	0	1.00	0	0	0	0
Relance précoce	0.03	0	0	0	0.91	0	0	0.06
Relance tardive	0	0	0	0	0	1.00	0	0
Vanne de Régulation	0.06	0	0	0	0	0	0.94	0
PR (*)	0.95							
PFA (**)	0.01							

Tableau 4.10 Probabilités de classification de chaque classe (35 formes) de test - RU2 avec 5 neurones cachés

Commentaires sur le test d'un réseau RU2 avec 5 neurones cachés (Tableau 4.10) :

- le taux de reconnaissance de 95 % signifie une classification très satisfaisante de l'ensemble de 7 classes en utilisant un réseau du type RU2 avec seulement 5 neurones.
- le taux de fausse alarme de 1 % pour l'ensemble de 7 classes représente approximativement, pour le diagnostic journalier, une seule fausse alarme tous les 5 mois⁽³⁾, donc très peu de fausses alarmes.

4.2.2.2.3 - Conclusion sur les réseaux uniques (RU11, RU12 et RU2)

On a étudié trois structures de réseaux uniques : RU11 et RU12 utilisant une stratégie d'apprentissage sans la classe normale, et RU2 utilisant une stratégie d'apprentissage avec la classe normale. Basé à la fois sur le résultat d'apprentissage (i.e. SEQ) et sur celui de test (taux ou probabilité de classification), un réseau du type **RU2 avec 5 neurones cachés** est considéré **comme la meilleure représentation interne des classes apprises (PR=0.95, PFA=0.01)**. Cela montre que : d'une part la qualité de l'espace de représentation a une influence très sensible sur la discrimination des classes ; d'autre part l'apprentissage avec la classe normale se réalise mieux, bien qu'il n'existe aucun paramètre discriminant à lui seul la classe normale des classes avec défauts.

Le choix du nombre des neurones cachés est un travail scrupuleux et demande un temps important pour deux raisons principales :

- manque de critère objectif car ce nombre des neurones cachés conditionne à la fois l'erreur d'apprentissage du réseau et sa généralisation (ou test) ; par ailleurs cette généralisation n'a été réalisée qu'avec un nombre d'exemples limité et est par conséquent insuffisante.
- nécessité d'initialiser plusieurs fois les poids synaptiques (10-15 fois pour notre étude) pour un nombre quelconque de neurones cachés, afin de trouver la "meilleure" représentation interne des classes apprises.

Répondons maintenant aux deux questions posées à la fin du paragraphe 4.2.1.2. Il semble que la première a été résolue avec un réseau du type RU2 et que la deuxième n'existe pas ; mais est-ce toujours vrai quand on généralise ces réseaux vers d'autres systèmes de chauffage non appris ?

En ce qui concerne la poursuite des travaux de recherche, le problème de choix entre deux approches différentes se pose immédiatement :

- fait-on d'abord l'analyse de généralité de ces réseaux uniques au risque de devoir se trouver sans solution si cette généralité n'est pas satisfaisante puisque d'autres types de réseaux n'ont pas été développés ?
- ou développe-t-on d'autres types de réseaux dont on pourra comparer la performance de généralisation et par conséquent choisir le meilleur sachant que l'apprentissage d'un réseau est assez rapide si on utilise un algorithme d'apprentissage puissant du type Levenberg-Marquardt ?

⁽³⁾ 5 mois $\approx$ [100 (jours)] / [5 (jours/semaine) x 4 (semaines/mois)]

Ces deux approches sont envisageables, mais on a choisi la deuxième pour la raison suivante :

- utilisé a priori pour discriminer la classe de vanne de régulation des autres classes, le paramètre T_{dep} (Température de l'eau de départ du circuit secondaire nord de 18 h à 23 h) distingue également assez bien la classe n°32 des autres classes (voir Figure 4.4). Cela signifie que la température de l'eau de départ du circuit secondaire nord sert également à détecter les défauts liés à la courbe de chauffe. Or, ce paramètre est fortement influencé d'une part par le modèle de courbe de chauffe utilisé dans l'algorithme et d'autre part par le type de bâtiment. Il est donc très difficile de trouver T_{dep} de référence qui représente la courbe de chauffe bien réglée pour tous les types de système. La question se pose donc : la généralisation de ces réseaux uniques vers d'autres systèmes de chauffage non appris est-elle possible à cause de la présence de T_{dep} dans ces réseaux ?

Pour que le paramètre T_{dep} ne soit pas présenté dans un réseau dont on détecte les défauts liés à la courbe de chauffe, on a développé les réseaux dits Réseaux Multiples (RM) qui seront présentés dans le paragraphe suivant.

4.2.2.3 - Réseaux Multiples

L'idée des Réseaux Multiples (RM) est de construire une structure imbriquée de réseaux neuronaux à deux niveaux :

- le premier - Réseau Multiple 1 (RM1) - pour discriminer la classe courbe de chauffe (ω_{32}) des autres classes sans utilisation du paramètre T_{dep} (la température de l'eau de départ secondaire) en tant qu'entrée de réseau. Notons que le réglage de la courbe de chauffe et le diagnostic de défauts liés à celle-ci sont, on le sait, généralement difficiles pour un opérateur moyen. Il serait donc intéressant de construire un système de diagnostic de deux niveaux dont le premier peut détecter et puis corriger ce type de défaut, et ensuite le deuxième peut diagnostiquer plus aisément, on l'espère, les autres défauts.
- le deuxième - Réseau Multiple 2 (RM2) - pour discriminer les autres classes (normale ω_0 , brûleur ω_1 , échangeur ω_2 , relance précoce ω_{41} , relance tardive ω_{42} , vanne de régulation ω_5).

4.2.2.3.1 - Réseaux Multiples 1 (RM1)

Bases d'apprentissage et bases de test

On utilise ci-après trois différentes bases des données pour apprendre les réseaux RM1 afin de trouver le "meilleur" apprentissage discriminant le mieux possible la classe courbe de chauffe ω_{32} des autres classes (ω_0 , ω_1 , ω_2 , ω_{41} , ω_{42} , ω_5).

(1) Base de données 1 (BD1)

L'apprentissage est réalisé avec 2 classes (ω_0 et ω_{32}) qui ne comportent pas d'exemple de lundi.

Comme la courbe de chauffe est régulée en fonction de la température extérieure (T_{ext}), celle-ci est utilisée comme entrée de réseau. D'autre part, la température intérieure de la période d'occupation (T_{ioc}), régulée principalement par la courbe de chauffe, contient a priori les informations les plus significatives sur les défauts liés à la courbe de chauffe. Or ce paramètre (T_{ioc}) ne discrimine pas la classe normale des autres classes avec défauts, exceptée la classe courbe de chauffe (voir la Figure 4.2). Pour cette raison, on n'utilise que deux classes (ω_0 et ω_{32}) pour apprendre les réseaux neuronaux mais tester ceux-ci avec toutes les 7 classes.

Pour un bâtiment fonctionnant avec un système de chauffage intermittent, la mise en régime (période d'occupation) est généralement très lente après la longue coupure du chauffage pendant le week-end. Par conséquent, la température intérieure de la période d'occupation est souvent plus basse les lundis que les autres jours de la semaine. Pour ne pas apprendre les réseaux avec un comportement indépendant des défauts de courbe de chauffe (par exemple phénomène de lundi), nous avons éliminé durant l'apprentissage les exemples des lundis.

Toujours avec 70 jours de simulation sélectionnés (14 semaines du lundi au vendredi), nous avons obtenu un ensemble d'apprentissage composé de 56 exemples ($28(\text{exemples / classe}) \times 2(\text{classes})$), 28 exemples correspondant à 7 semaines du mardi au vendredi, et un ensemble de test composé de 245 exemples ($35(\text{exemples / classe}) \times 7(\text{classes})$). Chaque exemple (apprentissage ou test) est codé simplement par deux paramètres : T_{ioc} (Figure 4.2) et T_{ext} (Figure 4.5).

(2) Base de données 2 (BD2)

L'apprentissage est effectué avec 2 classes (normale ω_0 et courbe de chauffe ω_{32}) qui comportent cette fois-ci les exemples de lundi ; ces exemples sont distingués des autres exemples (mardi-vendredi) par la Durée de Non-Occupation (DNO) avant la période d'occupation. Pour un bâtiment de bureaux, la DNO est de 62 heures pour un lundi, de 14 heures pour les autres jours de la semaine.

Nous avons obtenu, dans ce cas-là, un ensemble d'apprentissage de 70 exemples ($35(\text{exemples / classe}) \times 2(\text{classes})$) et un ensemble de test toujours de 245 exemples ; chaque exemple est codé par trois paramètres : T_{ioc} , T_{ext} , et DNO.

(3) Base de données 3 (BD3)

L'apprentissage se réalise avec l'ensemble de 7 classes (y compris les exemples de lundis) mais les sorties des réseaux sont codées, comme les deux bases précédentes, en 2 classes (l'une pour la classe courbe de chauffe et l'autre pour l'ensemble des autres classes). Un ensemble d'apprentissage et un ensemble de test sont tous deux composés de 245 exemples, chaque exemple étant également codé par trois paramètres : T_{ioc} , T_{ext} , et DNO.

Structure de réseaux

La structure des réseaux RM1 est présentée schématiquement par la Figure 4.11. Le nombre de neurones d'entrée est égal à 3 (1 neurone seuil + 2 neurones correspondant à 2 composantes d'une forme) pour la base de données BD1, et égal à 4 (1 neurone seuil + 3 neurones correspondant à 3 composantes d'une forme) pour les bases des données BD2 et BD3. Le nombre des neurones de sortie est de 2 quelle que soit la base de données, qui correspond à deux groupes de classes : le premier groupe représentant la classe courbe de chauffe ; le deuxième groupe représentant toutes les autres classes (normale, brûleur, échangeur, relance précoce, relance tardive, vanne de régulation). Les vecteurs de sortie désirés pour ces classes d'apprentissage sont donc de dimension 2.

Nous avons fait varier, durant l'apprentissage, le nombre des neurones cachés de 1 à 3 pour les bases des données BD1 et BD2, de 2 à 4 pour la base de données BD3. L'algorithme d'apprentissage est la méthode d'optimisation Levenberg-Marquardt.

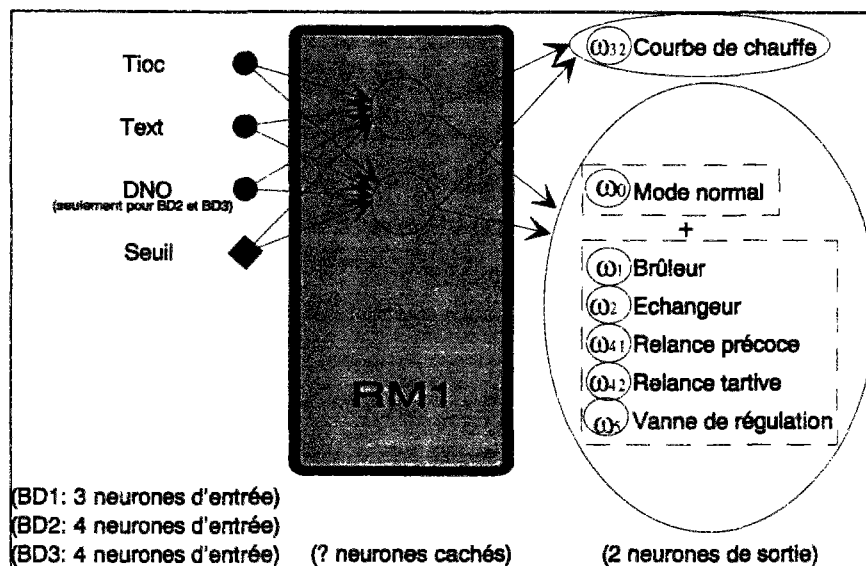


Figure 4.11 Structure des Réseaux Multiples 1 (RM1)

Les résultats d'apprentissage et de test sont présentés ci-après.

Résultat d'apprentissage

Les Sommes des Erreurs Quadratiques (SEQ) des 3 bases d'apprentissage entre les vecteurs de sortie désirés et les vecteurs de sortie réels, avec la variation du nombre des neurones cachés de 1 à 4, sont présentées dans le Tableau 4.11 :

Nombre des neurones cachés		1	2	3	4
SEQ	BD1	28	28	25	-
	BD2	29	24	26	-
	BD3	-	76	48	54

*Tableau 4.11 SEQ de 3 bases d'apprentissage - RM1
 (BD1 - 56 formes ; BD2 - 70 formes ; BD3 - 245 formes)*

Commentaire sur l'apprentissage : la SEQ ne diminue pas toujours avec l'accroissement du nombre des neurones cachés, mais la tendance est décroissante.

Résultat de test

Quelle que soit la base de données employée, on utilise toujours l'ensemble de 7 classes pour le test. Pour chaque base de données, le résultat (probabilités de classification) avec un seul nombre de neurones cachés et pour un seul apprentissage, considéré comme le meilleur, sera présenté (Tableaux 4.12, 4.13 et 4.14).

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.91	0.09	0
Normale	0.23	0.77	0
Brûleur	0.23	0.77	0
Echangeur	0.23	0.77	0
Relance précoce	0.17	0.83	0
Relance tardive	0.31	0.69	0
Vanne de Régulation	0.14	0.86	0

Tableau 4.12 Probabilités de classification de chaque classe (35 formes) de test - RM1 avec la BD1 (2 neurones cachés)

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.88	0.12	0
Normale	0.11	0.89	0
Brûleur	0.09	0.91	0
Echangeur	0.09	0.91	0
Relance précoce	0	1.00	0
Relance tardive	0.26	0.71	0.03
Vanne de Régulation	0	1.00	0

Tableau 4.13 Probabilités de classification de chaque classe (35 formes) de test - RM1 avec la BD2 (2 neurones cachés)

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.77	0.23	0
Normale	0	1.00	0
Brûleur	0	1.00	0
Echangeur	0	1.00	0
Relance précoce	0	1.00	0
Relance tardive	0.14	0.86	0
Vanne de Régulation	0	1.00	0

Tableau 4.14 Probabilités de classification de chaque classe (35 formes) de test - RM1 avec la BD3 (3 neurones cachés)

Commentaires sur le test (trois tableaux ci-dessus) [Li, 1996]:

- deux neurones cachés pour les BD1, BD2 et trois neurones pour la BD3 semblent le meilleur choix ; et cela signifie que l'augmentation des exemples d'apprentissage nécessite éventuellement plus de neurones cachés.
- la performance de reconnaissance est également conditionnée par la base d'apprentissage. La BD1 engendre un très bon score de reconnaissance pour la classe courbe de chauffe au détriment de toutes les autres classes en particulier de la classe relance tardive (i.e. fausses alarmes nombreuses). En revanche pour les classes autres que la classe courbe de chauffe, la BD3 entraîne un taux de reconnaissance tout à fait remarquable exceptée la classe relance tardive ; mais la détection de défauts liés à la courbe de chauffe n'est pas très bonne (23 % des exemples non détectés!) bien que la non-détection de ce type de défaut soit souvent liée à des jours où les apports solaires sont importants.
- la mauvaise classification de la **classe courbe de chauffe** est généralement liée à des jours où les apports solaires sont relativement importants. Prenons comme exemple, le test avec la base de données BD2 : 3 parmi les 4 exemples mal classés (12 %, voir le Tableau 4.13) étant liés à ce type de jour (voir également la Figure 4.12). Notons que la non-détection, pour des jours dont l'ensoleillement est fort, de la courbe de chauffe trop basse n'est pas très préoccupante car d'une part les usagers de bâtiment ne peuvent pas remarquer ce type de défaut puisque la température intérieure due à l'ensoleillement est normale ; d'autre part il n'y a pas de fausse alarme envoyée à l'équipe de maintenance par ce réseau. **En ce sens, on dirait que la base de données BD3 engendre un meilleur résultat de reconnaissance, mais sera-ce toujours vrai avec des exemples de test plus génériques ? On le verra dans le paragraphe 4.3.**

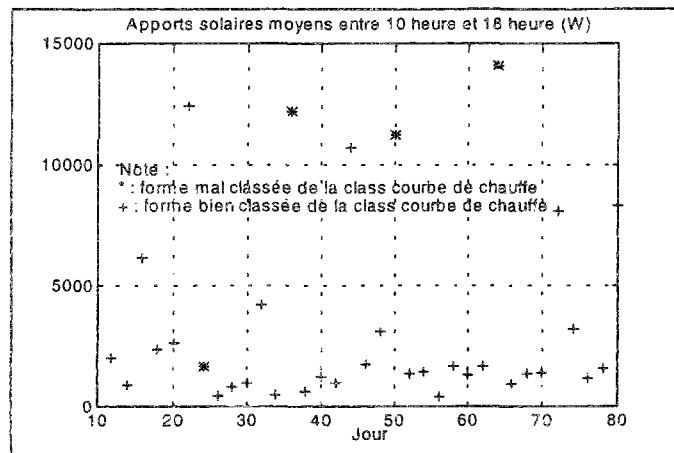


Figure 4.12 Effet des apports solaires sur la classification de la classe courbe de chauffe

- une des difficultés majeures est la détection de la **classe relance tardive** (selon la base de données utilisée, 14 à 31 % de la classe relance tardive classés dans la classe courbe de chauffe). En effet, quand la relance matinale est trop tardive, la température intérieure au début d'occupation sera trop basse. Par conséquent, la température intérieure durant l'occupation (T_{ioc}) pourrait également être trop basse ; cela ressemble à un effet de courbe de chauffe trop basse. Comme le paramètre T_{ioc} a été utilisé en tant qu'entrée des réseaux RM1, il est compréhensible qu'un réseau du type RM1 classe la classe relance tardive dans la classe courbe de chauffe, donc émette des fausses alarmes. Notons d'ailleurs que ce phénomène se produit beaucoup plus fréquemment au début de la semaine c'est-à-dire après un arrêt de chauffage relativement long (les lundis par exemple), qu'à la fin de la semaine (les vendredis par exemple).

- en ce qui concerne les **autres classes (normale, brûleur, échangeur, relance précoce, vanne de régulation)**, la mauvaise détection en moyenne ne dépasse pas 6 % ($= \frac{0.11 + 0.09 + 0.09 + 0 + 0}{5}$, voir le Tableau 4.13) avec la BD2, et se réduit à zéro avec la base BD3, restant à savoir si le seuil de 6 % est acceptable ou non pour l'équipe de maintenance.
- si on utilise un seul réseau RM1, les probabilités de reconnaissance (PR) et les probabilités de fausse alarme (PFA) pour les trois bases d'apprentissage sont récapitulées dans le tableau suivant :

Base d'apprentissage	PR_{M1}	PFA_{M1}
RM1 avec BD1	0.80 (*)	0.20 (**)
RM1 avec BD2	0.90	0.09
RM1 avec BD3	0.95	0.05

Tableau 4.15 PR_{M1} et PFA_{M1} des Réseaux Multiples 1 (RM1)

Note : (*) $\frac{0.91 + 0.77 + 0.77 + 0.77 + 0.83 + 0.66 + 0.86}{7} = 0.80$ (voir Tableau 4.12)

(**) $\frac{0.09 + 0.23 + 0.23 + 0.23 + 0.17 + 0.31 + 0.14}{7} = 0.20$ (voir Tableau 4.12)

La question se pose : la base d'apprentissage BD3 est-elle la meilleure parmi les trois bases de données (BD1, BD2 et BD3) ?

On saura mieux répondre à cette question en combinant les réseaux RM1 avec les réseaux RM2 présentés dans le paragraphe 4.2.2.3.2.

4.2.2.3.2 - Réseaux Multiples 2 (RM2)

Nous avons utilisé trois bases de données pour apprendre les réseaux RM1 permettant la discrimination de la classe courbe de chauffe des autres classes. Il faut maintenant développer des réseaux (RM2) afin de discriminer toutes les autres classes que la classe courbe de chauffe, c'est-à-dire la classe normale ω_0 , la classe brûleur ω_1 , la classe échangeur ω_2 , la classe relance précoce ω_{41} , la classe relance tardive ω_{42} , la classe vanne de régulation ω_5 .

Base d'apprentissage et base de test

En utilisant les réseaux RM2 pour discriminer les autres classes que la classe courbe de chauffe, on suppose que la dernière a déjà été bien détectée et corrigée. En conséquence il est raisonnable d'**apprendre les réseaux RM2 sans utiliser la classe courbe de chauffe**.

Cela signifie que la base d'apprentissage est composée de 210 formes (6 classes dont chaque classe comporte 35 formes).

Pour la base de test, on applique les hypothèses suivantes :

- si une forme de la classe courbe de chauffe est classée par le réseau RM1, dans le premier groupe (classe courbe de chauffe) (voir la Figure 4.11), on considère que celle-ci est bien classée et n'entre plus dans le réseau RM2. En revanche, si une forme de la classe courbe de chauffe est classée par le RM1, dans le deuxième groupe (les 6 autres classes) (voir la Figure 4.11), celle-ci sera testée par le réseau RM2 et ensuite soit classée dans une des 6 autres classes soit rejetée car ambiguë.

- si une forme de 6 autres classes est classée par le RM1 dans le premier groupe, celle-ci est considérée comme mal classée et n'entre plus dans le RM2, ce qui entraîne une fausse alarme. En revanche si une de ces formes est classée par le RM1 dans le deuxième groupe, elle sera testée par le RM2 et ensuite soit classée dans une des 6 autres classes soit rejetée car ambiguë.

Ces hypothèses nous ont conduits à avoir, pour l'ensemble des 7 classes, une base de test comportant moins de 245 formes, chaque classe ayant un nombre généralement différent des formes testées par le réseau RM2. Par exemple, en ce qui concerne le réseau RM1 avec la base de données BD3 (voir le Tableau 4.14), il n'a que 23 % (i.e. 8 formes) des formes de la classe courbe de chauffe, mais 86 % (i.e. 30 formes) des formes de la classe relance tardive qui entreront dans le réseau RM2.

Structure de réseaux

La structure des réseaux RM2 est présentée schématiquement par la Figure 4.13.

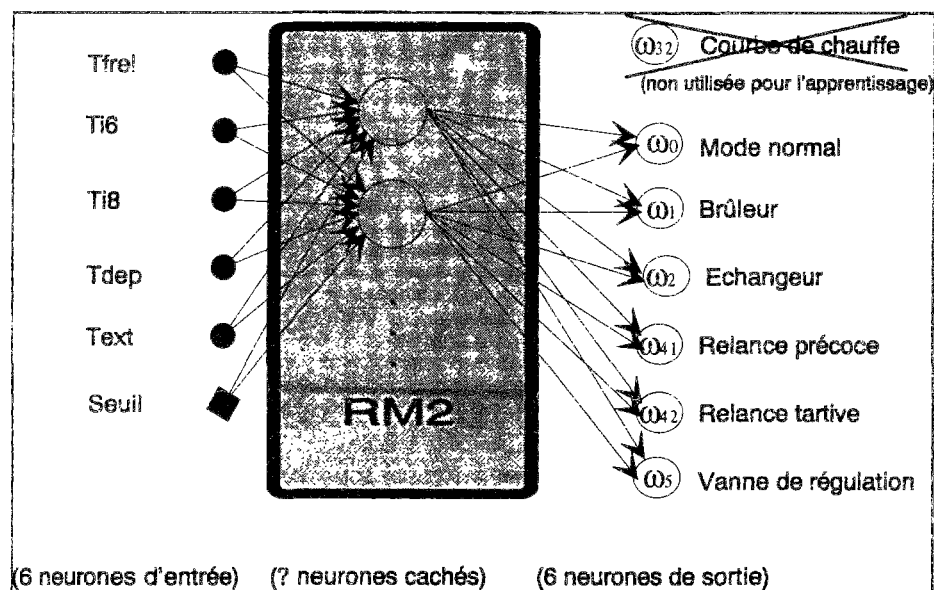


Figure 4.13 Structure des Réseaux Multiples 2 (RM2)

Du fait que la classe courbe de chauffe n'est pas utilisée pour l'apprentissage, il n'est pas nécessaire de prendre la température intérieure durant la période d'occupation T_{ioc} comme entrée de réseau. Par conséquent, nous n'avons que 5 paramètres (Figure 4.13) pour coder les exemples d'apprentissage et de test, c'est-à-dire que le nombre des neurones d'entrée est de 6 (1 neurone seuil + 5 neurones correspondant à 5 composantes d'une forme). Et le nombre des neurones de sortie est de 6 correspondant à 6 classes d'apprentissage. Les vecteurs forme de sortie pour ces 6 classes d'apprentissage sont donc de dimension 6.

Le nombre des neurones cachés varie de 3 à 6 durant l'apprentissage en utilisant toujours l'algorithme optimisation de Levenberg-Marquardt. Les résultats d'apprentissage et de test sont présentés ci-dessous.

Résultats d'apprentissage

Pour trouver la meilleure représentation interne des réseaux RM2, on a appris les RM2 avec 4 nombres différents de neurones cachés (voir le Tableau 4.16).

La SEQ de l'ensemble d'apprentissage entre les vecteurs de sortie désirés et les vecteurs de sortie réels en fonction du nombre des neurones cachés est présentée dans le Tableau 4.16 :

Nombre des neurones cachés	3	4	5	6
SEQ	8.1	4.4	6.4	3.3

Tableau 4.16 SEQ de 210 formes (6 classes) d'apprentissage - RM2

Commentaire sur l'apprentissage :

- comme avec les réseaux RM1, on peut noter que la SEQ ne diminue pas toujours avec le nombre des neurones cachés, mais la tendance est décroissante (la valeur de 8 de SEQ signifie qu'un seul exemple d'apprentissage est mal classé).

Résultat de test

On a testé les RM2 dans un premier temps avec l'ensemble des formes (210 formes dont chaque classe en comportant 35) des 6 classes (normale, brûleur, échangeur, relance précoce, relance tardive, vanne de régulation), sans tenir compte de la classification de celles-ci par les RM1.

Le résultat de test (ici taux de classification) est présenté sous deux formes différentes (voir le Tableau 4.17) :

- taux de classification de chaque classe afin de mieux cibler les problèmes de chaque classe,
- taux de classification pour l'ensemble des classes afin de faire ressortir un critère global.

Classe (s)	Taux de classification	Nombre des neurones cachés			
		3	4	5	9
Normale (ω_0)	Mauvaise (%)	28.6	8.6	14.3	14.3
	Non (%)	2.9	20.0	8.6	5.7
Brûleur (ω_1)	Mauvaise (%)	8.6	0	5.7	0
	Non (%)	2.9	0	2.9	0
Echangeur (ω_2)	Mauvaise (%)	0	0	0	0
	Non (%)	0	0	0	0
Relance précoce (ω_{41})	Mauvaise (%)	0	0	0	0
	Non (%)	0	0	5.7	2.9
Relance tardive (ω_{42})	Mauvaise (%)	2.9	0	0	0
	Non (%)	0	5.7	2.9	0
Vanne de régulation (ω_5)	Mauvaise (%)	11.4	0	0	0
	Non (%)	2.9	5.7	8.6	2.9
Ensemble de 6 classes	Mauvaise (%)	8.6	1.4	3.3	2.4
	Non (%)	1.4	5.2	4.8	1.9

Tableau 4.17 Taux de classification de 210 formes (6 classes) de test - RM2

Basé essentiellement sur le taux de mauvaise classification de test, on considère qu'un réseau RM2 avec 4 neurones cachés engendre la meilleure représentation interne du réseau (taux de mauvaise classification de l'ensemble des 6 classes de 1.4 %).

Pour que les probabilités de classification de l'ensemble des 7 classes soient comparables entre les réseaux uniques (RU) et les réseaux multiples (RM), le calcul des probabilités de classification pour les réseaux multiples RM (RM1+RM2) est basé sur les hypothèses concernant la base de test présentées ci-dessus, c'est-à-dire sur une combinaison du RM1 avec le RM2. La structure des réseaux multiples peut donc être présentée comme le montre la Figure 4.14.

Comme avec les réseaux uniques, les probabilités de classification (matrices de confusion) pour les 7 classes à discriminer ont été calculées, sachant que trois réseaux de RM1 (correspondant à 3 bases d'apprentissage BD1, BD2 et BD3) sont respectivement combinés avec le réseau RM2 comportant 4 neurones cachés. Ces trois combinaisons (trois Structures de Réseaux Multiples) sont désormais notées comme suit :

- RMS1 : RM1 (BD1 et 2 neurones cachés) combiné avec RM2 (4 neurones cachés),
- RMS2 : RM1 (BD2 et 2 neurones cachés) combiné avec RM2 (4 neurones cachés),
- RMS3 : RM1 (BD3 et 3 neurones cachés) combiné avec RM2 (4 neurones cachés).

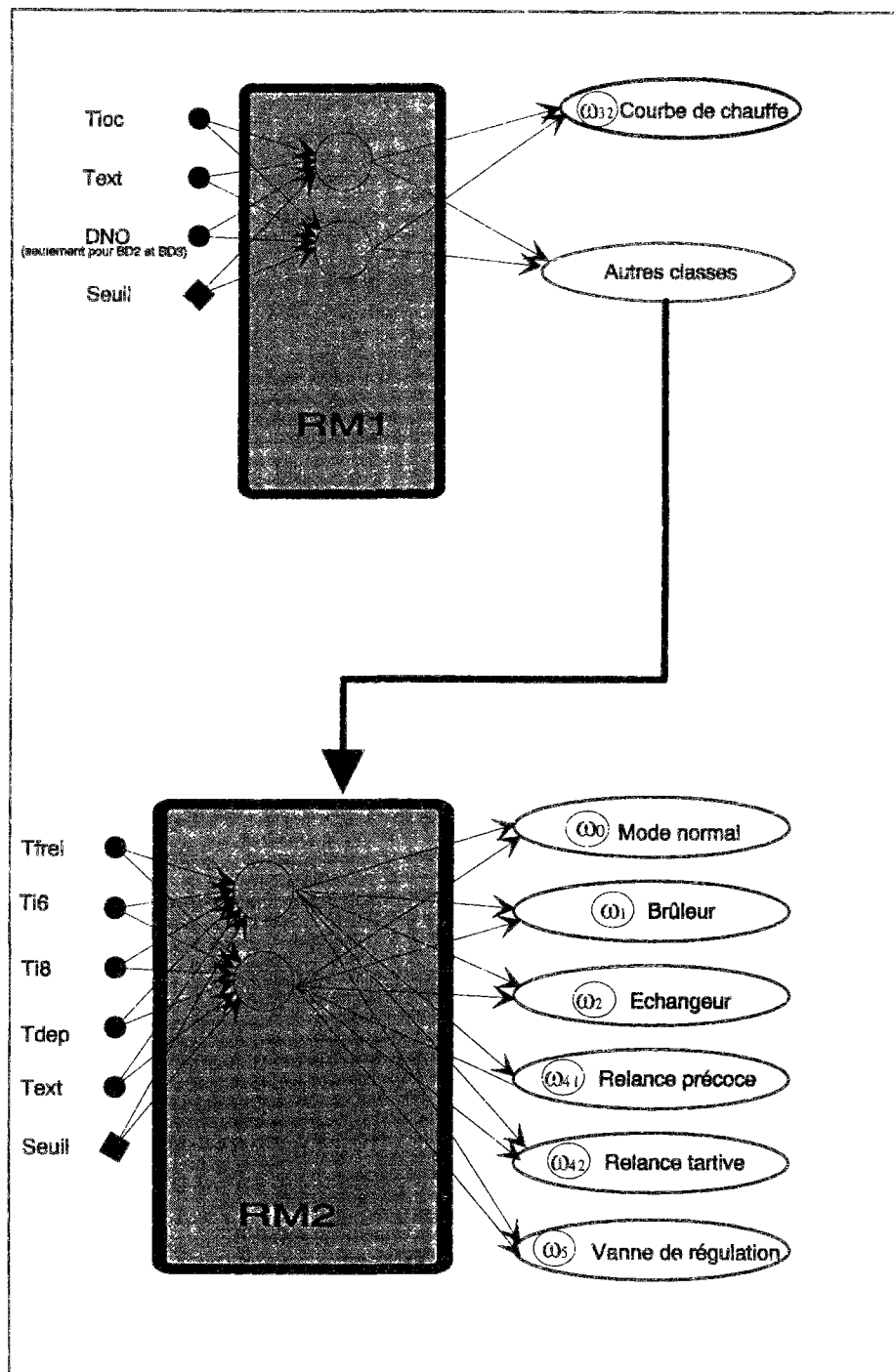


Figure 4.14 Structure de Réseaux Multiples à deux niveaux - RMS

Ces trois matrices de confusion sont présentées dans les tableaux 4.18 - 4.20.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage (*)	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.57	0	0	0.23	0.09	0	0	0.11
Brûleur	0	0.77	0	0.23	0	0	0	0
Echangeur	0	0	0.77	0.23	0	0	0	0
Courbe de chauffe	0.06	0	0	0.91	0.03	0	0	0
Relance précoce	0	0	0	0.17	0.83	0	0	0
Relance tardive	0	0	0	0.31	0	0.63	0	0.06
Vanne de régulation	0	0	0	0.14	0	0	0.80	0.06
PR	0.75							
PFA	0.20							

Tableau 4.18 Probabilités de classification de chaque classe de test (RMS1)

Note :

(*) Les chiffres de cette colonne sont directement issus du Tableau 4.12.

Les nombres des formes des classes testées par le RM2 avec 4 neurones cachés sont :

Classe normale 27 formes = $(1 - 0.23) \times 35$

Classe Brûleur 27 formes = $(1 - 0.23) \times 35$

Classe Echangeur 27 formes = $(1 - 0.23) \times 35$

Classe Courbe de chauffe 3 formes = $(1 - 0.91) \times 35$

Classe Relance précoce 29 formes = $(1 - 0.17) \times 35$

Classe Relance tardive 24 formes = $(1 - 0.31) \times 35$

Classe Vanne de régulation 30 formes = $(1 - 0.14) \times 35$

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage (*)	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.63	0	0	0.11	0.09	0	0	0.17
Brûleur	0	0.91	0	0.09	0	0	0	0
Echangeur	0	0	0.91	0.09	0	0	0	0
Courbe de chauffe	0.09	0	0	0.88	0.03	0	0	0
Relance précoce	0	0	0	0	1.00	0	0	0
Relance tardive	0	0	0	0.26	0	0.68	0	0.06
Vanne de Régulation	0	0	0	0	0	0	0.94	0.06
PR	0.85							
PFA	0.10							

Tableau 4.19 Probabilités de classification de chaque classe de test (RMS2)

Note : (*) Les chiffres de cette colonne sont directement issus du Tableau 4.13.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage (*)	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.71	0	0	0	0.09	0	0	0.20
Brûleur	0	1.00	0	0	0	0	0	0
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0.14	0	0	0.77	0.06	0	0	0.03
Relance précoce	0	0	0	0	1.00	0	0	0
Relance tardive	0	0	0	0.14	0	0.80	0	0.06
Vanne de Régulation	0	0	0	0	0	0	0.94	0.06
PR	0.89							
PFA	0.04							

Tableau 4.20 Probabilités de classification de chaque classe de test (RMS3)

Note : (*) Les chiffres de cette colonne sont directement issus du Tableau 4.14.

4.2.2.3.3 - Conclusion sur les réseaux multiples (RM1+RM2)

Après une phase d'analyse de réseaux multiples, on dispose maintenant de trois réseaux multiples (RMS1, RMS2 et RMS3). Plusieurs remarques méritent d'être mentionnées (voir aussi les trois tableaux 4.18 - 4.20).

(1) Classe normale

Pour la classe normale, le réseau RMS3 engendre un taux de reconnaissance nettement meilleur que le réseau RMS1 (71 % contre 57 %), sachant que la non-détection de la classe normale (20 % avec le réseau RMS3 et 11 % avec le réseau RMS1) n'entraîne pas de fausse alarme ni de conséquence grave ; de plus il n'y a aucune forme de cette classe mal classée dans la classe courbe de chauffe. En revanche, il a 9 % des formes (le même taux qu'avec les réseaux RMS1 et RMS2), mal classés dans la classe relance précoce.

D'autre part, il y a 20 % (avec le réseau RMS3) des formes de la classe normale non classées ce qui est non-négligeable. Pour comprendre le phénomène de non-classification de la classe normale, on présente dans le Tableau 4.21 les vecteurs de sortie réels de tous les exemples (formes) non classés. En effet, on remarque que ces exemples non-classés hésitent fréquemment (en particulier les vendredis) entre la classe normale et la classe relance précoce, peu souvent (peut-être les lundis) entre la classe normale et la classe relance tardive ou la classe brûleur.

Jour de la semaine	Vecteur de sortie réel
Lundi	[+0.8 +1.0 -1.0 -1.0 -1.0 -1.0]
Vendredi	[-0.0 -1.0 -1.0 -0.7 -1.0 -1.0]
Vendredi	[+0.4 -1.0 -1.0 -0.3 -1.0 -1.0]
Lundi	[+0.6 -1.0 -1.0 -1.0 +1.0 -1.0]
Mardi	[+0.8 -1.0 -1.0 +0.5 -1.0 -1.0]
Vendredi	[+0.4 -1.0 -1.0 -0.7 -1.0 -1.0]
Vendredi	[-0.2 -1.0 -1.0 -1.0 -1.0 -1.0]

Tableau 4.21 7 vecteurs de sortie réels des exemples non classés de la classe normale

Le fait que certains exemples de test de la classe normale hésitent entre la classe normale et la classe relance précoce ou tardive est probablement dû à l'utilisation d'un programmeur d'intermittence non-optimal dans notre simulation du système de chauffage. En effet il est possible, avec ce programmeur, que même si le système est régulé normalement, la relance matinale se déclenche parfois précocement ou tardivement. On pourrait donc penser que la non-classification de la classe normale n'est pas due à apprentissage ou choix de réseaux mais plutôt à la base de données simulée, qui a été obtenue en utilisant assez fidèlement un programmeur non-parfait mais correspondant à la réalité. Par conséquent si on espère améliorer la reconnaissance de la classe normale, il faut utiliser un meilleur programmeur d'intermittence et simuler à nouveau une autre base de données.

(2) Classe courbe de chauffe

Pour la classe courbe de chauffe, le réseau RMS3 donne un score de reconnaissance moins élevé que le réseau RMS1 (77 % contre 91 %) ; mais ces formes non détectées (par le RMS3) de la classe courbe de chauffe sont beaucoup plus fréquemment classées dans la classe normale que dans une classe avec défaut, ce qui n'entraîne pas de fausse alarme. Par ailleurs, ces formes non détectées sont très souvent liées à des jours où les apports solaires sont importants (voir aussi la Figure 4.12), c'est-à-dire à des jours où la température intérieure est normale en raison de l'ensoleillement, donc pas d'inconfort pour les usagers du bâtiment. En conclusion, la non-détection de ces formes de la classe courbe de chauffe n'engendre pas une situation gênante pour la maintenance.

(3) Classe relance tardive

Pour la classe relance tardive, le réseau RMS3 présente une performance de classification nettement meilleure que le réseau RMS1 (80 % contre 63 %). Il reste quand même encore 14 % (i.e. 5 formes) de cette classe mal classés (i.e. fausses alarmes) dans la classe courbe de chauffe.

Pour savoir si le deuxième réseau (i.e. RM2) du RMS3 classe correctement ou non ces 5 formes mal classées (par le premier niveau (RM1) du RMS3) de la classe relance tardive, celles-ci sont également testées par le RM2. On a remarqué que ces 5 formes sont bien classées dans la classe relance tardive. Cela nous a conduits à vérifier à nouveau toutes les formes de test classées dans la classe courbe de chauffe par le RM1, et à édicter les règles suivantes (voir la Figure 4.15) :

- une forme classée, par le RM1, dans la classe courbe de chauffe sera également testée par le RM2,
- si elle est classée dans une classe avec défaut, on la rejette car ambiguë au lieu de déclencher une alarme indiquant un défaut de courbe de chauffe,
- en revanche si elle est classée dans la classe normale ou rejetée car ambiguë, on la considère comme un défaut de la classe courbe de chauffe ; et par conséquent une alarme indiquant un défaut de courbe de chauffe se déclenchera.

Avec ces règles, le taux de fausses alarmes sera en principe sensiblement diminué mais le taux de reconnaissance de la classe courbe de chauffe sera également réduit.

Notons que ces règles n'ont pas été appliquées dans les travaux de recherche présentés dans cette thèse, mais ceci nous permettra d'ouvrir, si nécessaire, une autre voie de recherche.

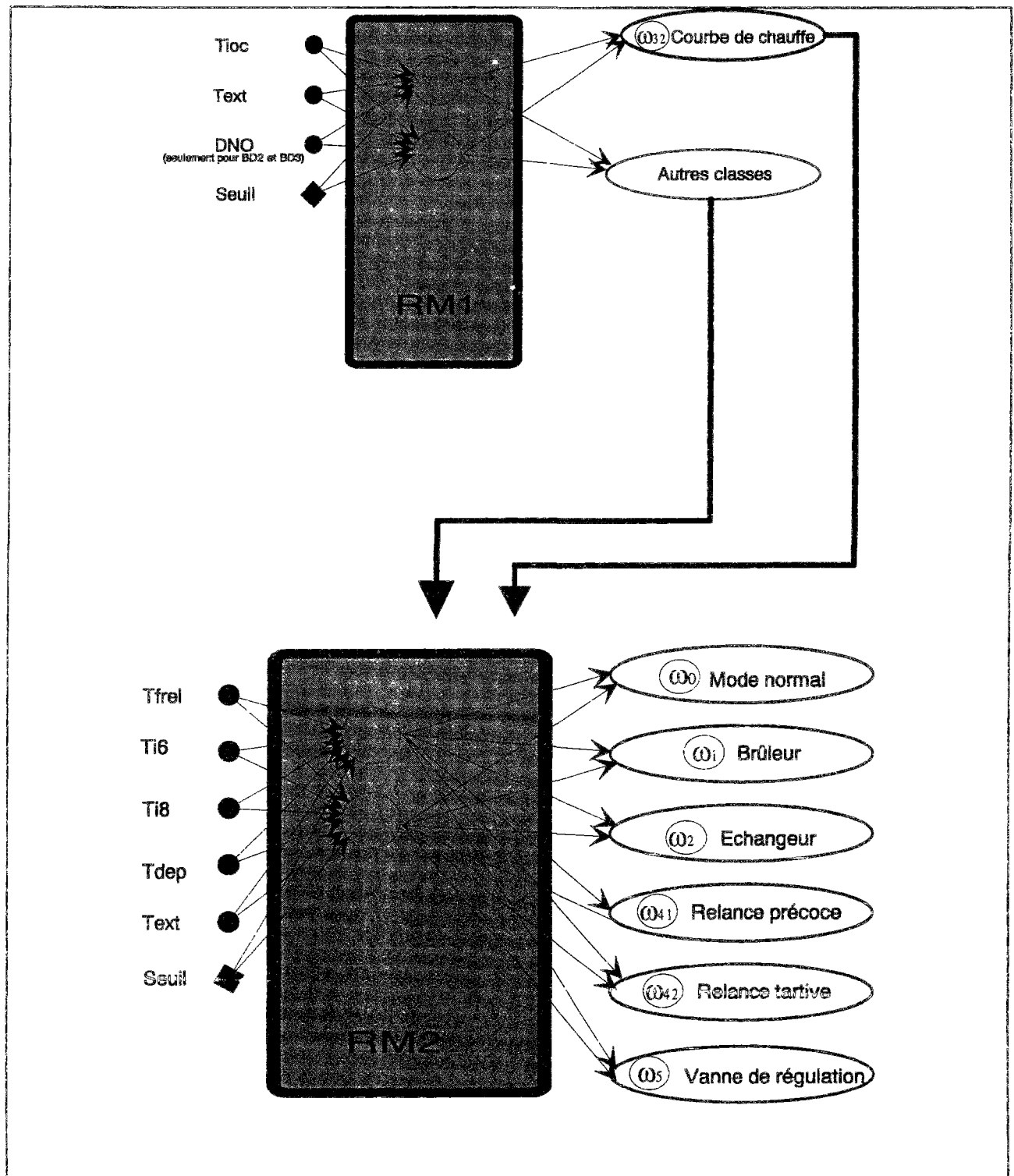


Figure 4.15 Variante de Structure de Réseaux Multiples à deux niveaux

(4) Classes brûleur, échangeur, relance précoce et vanne de régulation

Pour les classes brûleur, échangeur, relance précoce et vanne de régulation, les taux de reconnaissance sont, avec la structure RMS3, respectivement 100 %, 100 %, 100 % et 94 %, donc tout-à-fait satisfaisants. De plus il n'y a aucune fausse alarme due à la présence de ces classes.

(5) Probabilité de Reconnaissance (PR) et Probabilité de Fausse Alarme (PFA)

Les probabilités (PR et PFA) de ces trois structures sont récapitulées dans le Tableau 4.22. Ce tableau montre également que le réseau RMS3 donne une performance globale de classification nettement meilleure par rapport aux RMS1 et RMS2. On retiendra donc, dans la phase généralisation des réseaux vers d'autres systèmes de chauffage non appris, la structure RMS3 en tant que représentant de réseaux multiples, et comparera la performance de celle-ci avec le réseau unique RU2.

Structure de réseau	RMS1	RMS2	RMS3
PR	0.75	0.85	0.89
PFA	0.20	0.10	0.04

Tableau 4.22 Probabilités PR et PFA des trois réseaux multiples

4.2.2.4 - Conclusion

En utilisant une base de données issue d'un système de chauffage simulé, nous avons appris plusieurs types de réseaux neuronaux (3 réseaux uniques RU11, RU12 et RU2 ; 3 réseaux multiples RMS1, RMS2 et RMS3) avec deux algorithmes d'apprentissage : la Rétro-Propagation classique et l'Optimisation Levenberg-Marquardt. Ce dernier algorithme est nettement meilleur tant au niveau du temps d'apprentissage que de la facilité d'utilisation grâce à plusieurs critères d'arrêt (nombre d'itérations, erreur minimum, gradient minimum d'erreur, etc.).

Le choix d'un réseau (unique ou multiple) est basé non seulement sur le résultat d'apprentissage mais aussi sur celui de test. Après une phase d'analyse des résultats (apprentissage et test), nous avons retenu deux structures de réseaux de neurones : RU2 et RMS3. Les probabilités de classification (PR et PFA) sont présentées dans le Tableau 4.23. Ces deux structures seront appliquées, dans le paragraphe 4.3, aux 4 systèmes de chauffage simulés mais non appris par ces deux structures (RU2 et RMS3) afin de tester leur capacité de généralisation.

Structure de réseau	RU2	RMS3
PR	0.95	0.89
PFA	0.01	0.04

Tableau 4.23 PR et PFA des réseaux RU2 et RMS3

En comparant ces deux structures (voir aussi les tableaux 4.10 et 4.20), on remarque :

- que la structure RU2 détecte mieux les classes normale, courbe de chauffe et relance tardive,
- qu'elle engendre moins de fausses alarmes par rapport à la structure RMS3.

Ceci peut s'expliquer peut-être par le fait qu'un réseau unique a plus de chance d'acquérir une connaissance globale qu'un réseau multiple, mais qu'il faut le confirmer avec des données (formes) de généralisation plus importantes. On reviendra sur ce point dans le paragraphe 4.3.

4.3 - TEST DE LA GENERICITE DES RESEAUX DEVELOPPES

INTRODUCTION

L'**idée directrice** est d'apprendre, en utilisant certains défauts d'un système de chauffage spécifique, une (ou plusieurs) structures de réseaux capables de détecter et/ou diagnostiquer d'autres systèmes de chauffage non appris durant la phase d'apprentissage.

On a, dans un premier temps, conçu et simulé quatre systèmes de chauffage pour tester cette capacité de généralisation des réseaux développés. Ces systèmes sont les suivants :

- *système_occ* en augmentant la consigne de température intérieure de la période d'occupation de 19°C à 21°C,
- *système_rad* en augmentant la puissance (surface) des radiateurs dans le bâtiment,
- *système_iner* en augmentant l'inertie du bâtiment (de l'inertie moyenne à l'inertie forte),
- *système_inocc* en augmentant la consigne de température intérieure de la période d'inoccupation de 14°C à 16°C.

Les bases de données utilisées pour tester la généralité des réseaux sont construites par les mêmes jours et les mêmes paramètres que la base de test utilisée pour le développement de ceux-ci (voir paragraphe 4.2), c'est-à-dire 35 jours (exemples ou formes) par classe et par système.

La **normalisation** de paramètre emploie toujours la formule (4.1) mais le choix de $P_{réf}$ et ΔP se fait en fonction du système de chauffage à diagnostiquer, sachant que ce choix influence très sensiblement la discrimination de classes.

Pour rendre cet aspect plus clair, on présente, comme exemple, la normalisation des températures intérieures du système_occ et la normalisation des températures d'eau de départ secondaire du système_rad.

D'abord pour le système_occ.

Notons que dans le cas du système de référence (utilisé pour apprendre les réseaux) :

- $P_{réf} = 19^{\circ}\text{C}$, ce qui représente la consigne de température intérieure de la période d'occupation,
- $\Delta P = 21 - 19 = 2^{\circ}\text{C}$, valeur 21 étant la limite haute (vanne thermostatique) de température intérieure.

Or le système_occ a été obtenu en modifiant la consigne de température intérieure de 19°C à 21°C, les valeurs de $P_{réf}$ et ΔP devraient donc être modifiées comme suit :

- $P_{réf} = 21^{\circ}\text{C}$, ce qui représente la consigne de température intérieure de la période d'occupation,
- $\Delta P = 23 - 21 = 2^{\circ}\text{C}$, valeur 23 étant la limite haute (vanne thermostatique) de température intérieure.

Ensuite pour le système rad.

Comme le dimensionnement des radiateurs a été modifié (coefficient de surdimensionnement α de 1.3 à 2), la courbe de chauffe dans le régulateur doit également être modifiée ; par conséquent, la température d'eau de départ secondaire change (Figure 4.16).

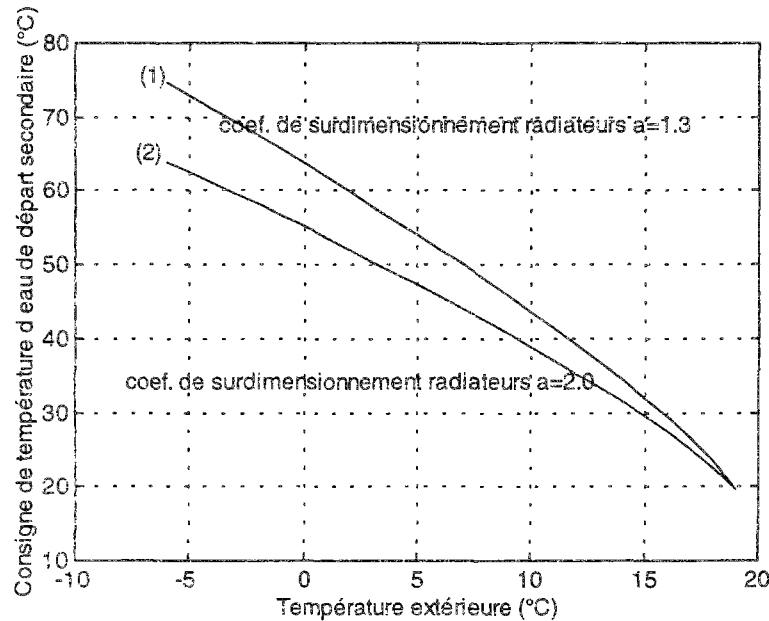


Figure 4.16 Variation de la consigne de température d'eau en fonction du coefficient de surdimensionnement des radiateurs

Les normalisations doivent tenir compte de ce changement, c'est-à-dire :

- pour le système de chauffage régulé par la courbe (1) (Figure 4.16) (i.e. système de référence) :

$$P_{\text{réf}} = T_{\text{cdép}}^{(1)}(T_{\text{ext}} = 5^{\circ}\text{C}), \text{ Température consigne d'eau de départ à la température extérieure de } 5^{\circ}\text{C},$$

$$\Delta P = T_{\text{cdép}}^{(1)}(T_{\text{ext}} = -5^{\circ}\text{C}) - T_{\text{cdép}}^{(1)}(T_{\text{ext}} = 15^{\circ}\text{C})$$

- mais pour le système de chauffage régulé par la courbe (2) (Figure 4.16) (i.e. système dont le coefficient de surdimensionnement des radiateurs a été augmenté) :

$$P_{\text{réf}} = T_{\text{cdép}}^{(2)}(T_{\text{ext}} = 5^{\circ}\text{C}),$$

$$\Delta P = T_{\text{cdép}}^{(2)}(T_{\text{ext}} = -5^{\circ}\text{C}) - T_{\text{cdép}}^{(2)}(T_{\text{ext}} = 15^{\circ}\text{C}).$$

Toutes les modifications des $P_{\text{réf}}$ et ΔP ne sont pas présentées ici. En revanche, le principe de cette approche est édicté ci-dessous :

- pour une classe (normale, brûleur, etc.) donnée, les modifications des $P_{\text{réf}}$ et ΔP d'un paramètre (T_{frel} , T_{loc} , etc.), en fonction de différents systèmes de chauffage, devraient permettre d'obtenir des valeurs numériques du paramètre normalisé pratiquement inchangées quel que soit le système de chauffage. Notons que les valeurs absolues du paramètre de différents systèmes sont généralement différentes.

Ce point est très intéressant car si on arrive à trouver des valeurs des $P_{\text{réf}}$ et ΔP qui suppriment la différence d'un paramètre normalisé de la même classe mais de différents systèmes de chauffage, on pourrait espérer apprendre les réseaux de neurones avec un seul système mais diagnostiquer certains autres systèmes non appris.

4.3.1 - INFLUENCE D'UNE MODIFICATION DE LA CONSIGNE DE TEMPERATURE INTERIEURE DE LA PERIODE D'OCCUPATION

La base de données, qui comporte 245 exemples (35 exemples / classe × 7 classe) et est issue de la simulation du système_occ, sera utilisée pour tester la généralité des deux structures de réseau (RU2 et RMS3).

(1) Test de la structure RU2

Le résultat (ici les probabilités de classification) de test du réseau RU2 est présenté dans le Tableau 4.24.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.66	0	0	0	0.11	0	0	0.23
Brûleur	0.17	0.83	0	0	0	0	0	0
Echangeur	0	0	1.00		0	0	0	0
Courbe de chauffe	0.17	0	0	0.52	0	0	0	0.31
Relance précoce	0.11	0	0	0	0.37	0	0.06	0.46
Relance tardive	0	0	0	0	0	1.00	0	0
Vanne de Régulation	0	0	0	0	0	0	0.94	0.06
PR	0.76							
PFA	0.02							

Tableau 4.24 Probabilités de classification de chaque classe (35 formes)

Commentaires :

- ce tableau montre que les classifications des classes **échangeur**, **relance tardive** et **vanne de régulation** sont toutes très satisfaisantes.
- en revanche les autres classes (**normale**, **brûleur**, **courbe de chauffe**, **relance précoce**) sont tous très mal détectées. Pour mieux comprendre ces mauvaises classifications, on a analysé les vecteurs de sortie réels de ces classes. Cette analyse montre :
 - . que parmi les 8 exemples (i.e. $8/35 = 0.23$) non classés de la classe normale, une grande partie (5 exemples) concerne un jour de lundi et hésite entre la classe normale et la classe relance précoce. Il serait donc possible que la classe normale soit classée dans la classe relance précoce le lundi en particulier mais n'ait aucune chance de se confondre avec les autres classes comportant des défauts.
 - . que parmi les 6 exemples (i.e. $6/35 = 0.17$) de la classe brûleur classés dans la classe normale (i.e. défaut de brûleur non détecté), 5 sont liés aux lundis.
 - . que parmi les 11 exemples (i.e. $11/35 = 0.31$) non classés de la classe courbe de chauffe, une grande partie (7 exemples) hésite entre la classe courbe de chauffe et la classe relance précoce. Cela signifie également qu'un défaut de courbe de chauffe est parfois considéré comme un défaut de relance précoce.

- que parmi les 16 exemples (i.e. 46 %) non classés de la classe relance précoce, une grande partie (12 exemples) hésite entre la classe relance précoce et la classe normale ; c'est-à-dire qu'il est très difficile de détecter la classe relance précoce mais qu'il n'y a, a priori, pas de confusion de cette classe avec les autres classes comportant des défauts.
 - que les taux de reconnaissance de toutes les classes, exceptée la classe relance précoce, sont acceptables à condition que la non-détection d'un défaut pendant quelques jours, pour ce type de système de chauffage, soit acceptable par les équipes de maintenance.
- l'hésitation des exemples (formes) entre la classe normale et la classe relance précoce ou entre la classe courbe de chauffe et la classe relance précoce, ainsi que les problèmes des lundis est probablement liée à utilisation d'un programmeur d'intermittence non-optimal dans notre simulation.

(2) Test de la structure RMS3

Les probabilités de classification de la structure RMS3 sont présentées dans le Tableau 4.25 et le Tableau 4.26.

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.89	0.11	0
Normale	0.26	0.74	0
Brûleur	0.26	0.74	0
Echangeur	0.26	0.74	0
Relance précoce	0	1.00	0
Relance tardive	0.57	0.43	0
Vanne de Régulation	0.06	0.94	0

Tableau 4.25 Probabilités de classification de chaque classe (35 formes) de test du système occ : RM1 avec la BD3 (3 neurones cachés)

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe (*)	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.17	0.06	0	0.26	0	0	0.09	0.42
Brûleur	0	0.29	0	0.26	0	0	0.03	0.42
Echangeur	0	0	0.74	0.26	0	0	0	0
Courbe de chauffe	0.03	0	0	0.89	0.06	0.03	0	0
Relance précoce	0	0.06	0	0	0.66	0	0	0.28
Relance tardive	0	0.23	0	0.57	0	0.17	0.03	0.20
Vanne de Régulation	0	0	0	0.06	0	0	0.94	0
PR	0.55							
PFA	0.30							

Tableau 4.26 Probabilités de classification de chaque classe (35 formes) de test

Commentaires :

- il est impossible de détecter les classes normale, brûleur et relance tardive.
- la classe courbe de chauffe est mieux détectée mais avec des fausses alarmes par rapport au réseau RU2.
- la classe relance précoce est généralement mieux détectée que le réseau RU2.
- globalement la probabilité de reconnaissance de l'ensemble des classes est très faible (PR=0.55), la probabilité de fausse alarme est inacceptable (PFA=0.30) ; la généralisation du réseau RMS3 vers le système_occ serait donc très difficile, voire impossible.
- il serait possible d'améliorer la probabilité de reconnaissance en utilisant un paramètre plus générique (par exemple : Température des fumées - température de départ primaire) que la température des fumées seule (voir la Figure 4.1) pour la discrimination des classes brûleur et échangeur. De même faut-il améliorer encore la normalisation de la température intérieure pour mieux classer la classe courbe de chauffe, la classe relance précoce et la classe relance tardive ? Toutes ces questions devraient être étudiées dans la poursuite des travaux concernant ce sujet.

(3) Conclusion

Avec la structure RU2, la probabilité de fausse alarme est faible ($PFR=0.02$), la probabilité de reconnaissance est médiocre ($PR=0.76$) mais acceptable à condition que la non-détection d'un défaut pendant plusieurs jours n'entraîne pas de grave conséquence.

Les principales difficultés de détection sont dans le cas du réseau RU2, liées au programmeur-régulateur (en particulier le problème de relance précoce). On pourrait espérer améliorer cette situation en utilisant dans notre simulation un programmeur-régulateur plus performant.

Avec la structure RMS3, La PR et la PFA sont toutes inacceptables.

4.3.2 - INFLUENCE D'UN CHANGEMENT DU DIMENSIONNEMENT DES RADIATEURS

La base de données, qui comporte 245 exemples (35 exemples par classe) et est issue de la simulation du système_rad, sera utilisée pour tester la génécité des deux structures de réseau (RU2 et RMS3).

(1) Test de la structure RU2

Les probabilités de classification de test du réseau RU2 sont présentées dans le Tableau 4.25.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.60	0	0	0.11	0.06	0	0	0.23
Brûleur	0	0.94	0	0	0	0	0	0.06
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0	0	0	1.00	0	0	0	0
Relance précoce	0	0	0	0.03	0.77	0	0	0.20
Relance tardive	0.08	0	0	0	0	0.83	0	0.09
Vanne de Régulation	0.03	0	0	0	0	0	0.94	0.03
PR	0.87							
PFA	0.03							

Tableau 4.27 Probabilités de classification de chaque classe (35 formes)

Commentaires :

- on remarque que les classes brûleur, échangeur, courbe de chauffe, vanne de régulation sont très bien détectées.

- en ce qui concerne les classes liées aux classes normale, relance précoce et relance tardive, on a examiné en détail les vecteurs de sortie réels de réseau. Cette analyse montre :

- que les 8 exemples (i.e. $8/35 = 0.23$) non classés de la classe normale hésitent soit entre la classe normale et la classe courbe de chauffe, soit entre la classe normale et la classe relance précoce ou encore entre la classe courbe de chauffe et la classe relance précoce ; mais la possibilité d'être classés dans la classe normale est forte (voir le Tableau 4.28).
- que sur les 7 exemples (i.e. $7/35 = 0.20$) non classés de la classe relance précoce, une grande partie a tendance à se classer dans cette classe et 3 parmi eux sont liés aux lundis (voir le Tableau 4.29) ; et un seul exemple mal classé dans la classe courbe de chauffe est également lié à un lundi. Cela signifie probablement qu'après une longue coupure (week-end) du bâtiment, il est plus difficile de chauffer le bâtiment les lundis que les autres jours de la semaine, par conséquent le défaut de relance précoce ne présente pas un vrai défaut et devient une relance correcte.

Jour de la semaine	Vecteur de sortie réel
Mercredi	[+0.12 -1.00 -1.00 -1.00 -1.00 -1.00 -1.00]
Mardi	[-1.00 -1.00 -1.00 -0.93 -0.95 -1.00 -1.00]
Mercredi	[-0.87 -1.00 -1.00 -1.00 -0.99 -1.00 -1.00]
Vendredi	[-0.94 -1.00 -1.00 -1.00 -0.97 -1.00 -1.00]
Lundi	[+0.68 -1.00 -1.00 +0.87 -1.00 -1.00 -1.00]
Mardi	[-0.99 -1.00 -1.00 -0.40 -1.00 -1.00 -1.00]
Mardi	[-1.00 -1.00 -1.00 -0.96 -0.56 -1.00 -1.00]
Vendredi	[+0.46 -1.00 -1.00 -0.99 -1.00 -1.00 -1.00]

Tableau 4.28 8 vecteurs (8 exemples non classés) de la classe normale

Jour de la semaine	Vecteur de sortie réel
Lundi	[-1.00 -1.00 -1.00 -0.98 +0.20 -0.87 -1.00]
Jeudi	[-1.00 -1.00 -1.00 -1.00 -0.01 -1.00 -1.00]
Lundi	[-1.00 -1.00 -1.00 -0.91 -0.53 -1.00 -1.00]
Mercredi	[-0.99 -1.00 -1.00 -1.00 -0.99 -1.00 -1.00]
Jeudi	[-1.00 -1.00 -1.00 -1.00 -0.99 -0.98 -1.00]
Lundi	[-1.00 -1.00 -1.00 +0.81 +0.91 -0.20 -1.00]
Vendredi	[-1.00 -1.00 -1.00 -1.00 -0.60 +0.13 -1.00]

Tableau 4.29 7 vecteurs (7 exemples non-classés) de la classe relance précoce

que les 3 exemples (i.e. $3/35 = 0.09$) non classés de la classe relance tardive hésitent quant à eux entre elle-même et la classe normale. Cela signifie qu'une partie de la classe relance tardive sera non détectée mais sans confusion de cette classe avec les autres classes comportant des défauts.

- on dirait donc que comme dans le cas du système_occ, les principales difficultés de détection sont liées au programmeur-régulateur (en particulier au problème de relance).

(2) Test de la structure RMS3

Les probabilités de classification de la structure RMS3 sont présentées dans le Tableau 4.30 et le Tableau 4.31.

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.83	0.17	0
Normale	0	1.00	0
Brûleur	0	1.00	0
Echangeur	0	1.00	0
Relance précoce	0	1.00	0
Relance tardive	0.09	0.91	0
Vanne de Régulation	0	1.00	0

Tableau 4.30 Probabilités de classification de chaque classe (35 formes) de test du système_rad RM1 avec la BD3 (3 neurones cachés)

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe (*)	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.68	0	0	0	0.23	0	0	0.09
Brûleur	0	0.94	0	0	0	0	0	0.06
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0.06	0	0	0.83	0.03	0	0	0.09
Relance précoce	0	0	0	0	1.00	0	0	0
Relance tardive	0.03	0	0	0.09	0	0.63	0	0.25
Vanne de Régulation	0	0	0	0	0	0	1.00	0
PR	0.87							
PFA	0.05							

Tableau 4.31 Probabilités de classification de chaque classe (35 formes) de test

Commentaires :

- la probabilité de reconnaissance pour l'ensemble des 7 classes reste la même qu'avec la structure RU2 (i.e. 0.87) ; en revanche cette structure (RMS3) engendre beaucoup plus de fausses alarmes.
- la détection de la **classe normale** reste toujours difficile.
- cette structure détecte moins bien la **classe courbe de chauffe** que la structure RU2.
- elle détecte mieux la classe **relance précoce** mais moins bien la **classe relance tardive** que la structure RU2.

(3) Conclusion

Globalement ces deux structures (RU2 et RMS3) engendrent la même probabilité de reconnaissance ; mais la structure RU2 déclenche moins de fausses alarmes.

Il est difficile de détecter la classe normale quelle que soit la structure employée.

En ce qui concerne la détection des classes avec défauts :

- le réseau RU2 a plus de difficulté à détecter la classe relance précoce (les lundis en particulier) ; cela engendra peut-être des surconsommations d'énergie mais en général pas de répercussion sur le confort des usagers du bâtiment.
- en revanche le réseau RMS3 détecte moins bien la classe relance tardive ; cela pourrait conduire à des conséquences sur le confort.

Globalement il apparaît que le RU2 engendre un meilleur résultat de test que le réseau RMS3 et que les principales difficultés de détection sont liées au régulateur-programmateur (courbe de chauffe ou programme de relance matinale) les lundis en particulier.

4.3.3 - INFLUENCE D'UN CHANGEMENT DE L'INERTIE DU BATIMENT

La base de données, comportant 245 exemples (35 exemples par classe) et issue de la simulation du système_iner, est utilisée pour tester la généricité des deux structures de réseau (RU2 et RMS3).

(1) Test de la structure RU2

Les probabilités de classification de test du réseau RU2 sont présentées dans le Tableau 4.32.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.77	0	0	0.03	0.06	0	0	0.14
Brûleur	0.06	0.86	0	0	0	0.03	0	0.05
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0	0	0	1.00	0	0	0	0
Relance précoce	0.03	0	0	0	0.89	0	0	0.08
Relance tardive	0	0	0	0	0	0.97	0	0.03
Vanne de Régulation	0.06	0	0	0	0	0	0.94	0
PR	0.92							
PFA	0.02							

Tableau 4.32 Probabilités de classification de chaque classe (35 formes)

Commentaires :

- pour la **classe normale**, parmi les cinq formes non classées il y en a trois qui hésitent entre la classe normale et la classe relance précoce. Cela signifie que les fausses alarmes liées à la classe normale indiquent plutôt un défaut de relance précoce que d'autres types de défaut.
- pour les classes avec défauts, la reconnaissance des **classes échangeur, courbe de chauffe, relance tardive et vanne de régulation** est très bonne et, il n'y a aucune fausse alarme en présence de ces trois classes de défaut.
- en revanche la reconnaissance des **classes brûleur et relance précoce** est légèrement moins bonne avec toutefois seulement une forme de test (de la classe brûleur) mal classée dans la classe relance tardive, donc des fausses alarmes minimales. Une analyse plus détaillée de ces deux classes montre :
 - . que pour la classe brûleur, une forme classée dans la classe relance tardive, deux formes classées dans la classe normale et deux formes non classées sont toutes liées aux jours de lundi, c'est-à-dire que le défaut de brûleur pourrait ne pas être détecté ou mal détecté (comme un défaut de relance tardive) en particulier les lundis. En effet pour un bâtiment plus lourd (i.e. inertie plus grande) et d'ailleurs non appris par le réseau testé, la signature d'un défaut de brûleur serait probablement plus difficilement détectable ou ressemblerait à celle d'un défaut de relance tardive, à cause de la complexité de couplage entre le système de chauffage et le bâtiment (changement de l'inertie du bâtiment, utilisation d'un programmeur-régulateur non optimal, etc.).
 - . que pour la classe relance précoce, parmi les trois formes non classées, il y en a deux liées aux jours de lundi ; de même une seule forme classée dans la classe normale est également liée à un lundi. L'explication de ce phénomène serait probablement la même que dans le cas du brûleur.

(2) Test de la structure RMS3

Les probabilités de classification de la structure RMS3 sont présentées dans les tableaux 4.33 et 4.34.

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.85	0.15	0
Normale	0.14	0.86	0
Brûleur	0.14	0.86	0
Echangeur	0.14	0.86	0
Relance précoce	0.03	0.97	0
Relance tardive	0.34	0.66	0
Vanne de Régulation	0.03	0.97	0

Tableau 4.33 Probabilités de classification de chaque classe (35 formes) de test du système iner RM1 avec la BD3 (3 neurones cachés)

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe (*)	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.55	0	0	0.14	0.17	0	0	0.14
Brûleur	0	0.83	0	0.14	0	0	0	0.03
Echangeur	0	0	0.86	0.14	0	0	0	0
Courbe de chauffe	0.09	0	0	0.85	0	0	0	0.06
Relance précoce	0.03	0	0	0.03	0.94	0	0	0
Relance tardive	0	0	0	0.34	0.03	0.60	0	0.03
Vanne de Régulation	0	0	0	0.03	0	0	0.94	0.03
PR	0.80							
PFA	0.15							

Tableau 4.34 Probabilités de classification de chaque classe (35 formes) de test

Commentaires :

- en ce qui concerne la **classe normale**, il est non seulement difficile de la séparer de la classe courbe de chauffe (14 % de mauvaise classification) mais aussi une partie non négligeable (17 %) de cette classe est mal classée dans la classe relance précoce, ce qui entraîne donc de fausses alarmes fréquentes.
- de même il est difficile de séparer les **classes brûleur et échangeur** de la classe courbe de chauffe (14 % de mauvaise classification), ce qui déclenche également des fausses alarmes.
- le réseau RMS3 détecte moins bien la **classe courbe de chauffe** que le réseau RU2 mais il n'y a toujours pas de fausse alarme due au défaut de courbe de chauffe.
- il est très difficile de séparer la **classe relance tardive** de la classe courbe de chauffe (34 % de mauvaise classification de la classe relance tardive), il y a donc des fausses alarmes fréquentes causées par la présence de la classe relance tardive.
- pour l'ensemble des classes à discriminer, la probabilité de reconnaissance ($PR=0.80$) est médiocre et la probabilité de fausse alarme ($PFA=0.15$) est inacceptable.

(3) Conclusion

En comparant les deux structures (RU2 et RMS3) on voit clairement que le réseau RU2 présente une meilleure performance de classification que le réseau RMS3 :

- le RU2 engendre pour l'ensemble de 7 classes un taux de reconnaissance satisfaisant (92 %) et un taux de fausse alarme très faible (2 %),
- le RMS3 engendre pour l'ensemble de 7 classes un taux de reconnaissance médiocre (80 %) et un taux de fausse alarme inacceptable (15 %).

Le test du réseau RU2 montre :

- que la détection des classes brûleur et relance précoce est, semble-t-il, plus difficile les lundis que les autres jours de la semaine,
- que la détection des autres classes avec défauts est très satisfaisante,
- qu'il est toujours relativement plus difficile de détecter la classe normale bien que le taux de reconnaissance soit plus grand que les deux systèmes testés précédemment (i.e. système_occ et système_rad) (voir les tableaux 4.24, 4.27 et 4.32) ; en revanche qu'il n'y a aucune fausse alarme indiquant les défauts autres que les défauts de courbe de chauffe et de relance précoce, si le système est en état de fonctionnement normal (i.e. en présence de la classe normale).

On pourrait conclure que les difficultés de détection liées aux classes normale, brûleur et relance précoce sont probablement liées d'une part au programmeur-régulateur non optimal utilisé dans la simulation et d'autre part au changement de l'inertie du bâtiment.

4.3.4 - INFLUENCE D'UNE MODIFICATION DE LA CONSIGNE DE TEMPERATURE INTERIEURE DE LA PERIODE D'INOCCUPATION

La base de données, comportant 245 exemples (35 exemples par classe) et issue de la simulation du système_inocc, est utilisée pour tester la généralité des deux structures de réseau (RU2 et RMS3).

(1) Test de la structure RU2

Les probabilités de classification de test du réseau RU2 sont présentées dans le Tableau 4.35.

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffage	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.94	0	0	0	0.03	0	0	0.03
Brûleur	0	0.97	0	0	0	0	0	0.03
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0	0	0	1.00	0	0	0	0
Relance précoce	0	0	0	0	0.97	0	0	0.03
Relance tardive	0.03	0	0	0	0	0.94	0	0.03
Vanne de Régulation	0.09	0	0	0	0	0	0.82	0.09
PR	0.95							
PFA	0.004							

Tableau 4.35 Probabilités de classification de chaque classe (35 formes)

Commentaires :

- la classification de l'ensemble des classes exceptée la classe vanne de régulation est très satisfaisante.
- le score relativement bas de la classification de la **classe vanne de régulation** est dû probablement au fait que la température de l'eau de départ du circuit secondaire de la période d'arrêt ou de maintien, utilisée surtout pour la discrimination de la classe vanne de régulation de l'ensemble des autres classes, est fortement liée à la consigne de température intérieure de la période d'inoccupation. Or le système_inocc a été conçu par le changement de la consigne de température intérieure de la période d'inoccupation de 14°C à 16°C, ceci entraîne probablement un accroissement de la difficulté de détection pour la classe vanne de régulation.
- il n'y a aucune confusion de la classe vanne de régulation avec les autres classes comportant des défauts, c'est-à-dire pas de fausse alarme due à la présence de cette classe. Pour cette raison, la non-détection de quelques formes de cette classe ne conduira pas, je pense, à une grave conséquence. La détection du défaut de la vanne de régulation pourrait donc être acceptable.

(2) Test de la structure RMS3

Les probabilités de classification de la structure RMS3 sont présentées dans les tableaux 4.36 et 4.37.

Classe testée	Classées dans la classe		
	Courbe de chauffe	Autres classes	Non classées
Courbe de chauffe	0.74	0.26	0
Normale	0	1.00	0
Brûleur	0	1.00	0
Echangeur	0	1.00	0
Relance précoce	0	1.00	0
Relance tardive	0	1.00	0
Vanne de Régulation	0	1.00	0

Tableau 4.36 Probabilités de classification de chaque classe (35 formes) de test du système inocc - RM1 avec la BD3 (3 neurones cachés)

Classe testée	Classées dans la classe							
	Normale	Brûleur	Echangeur	Courbe de Chauffe	Relance précoce	Relance tardive	Vanne de régulation	Non classées
Normale	0.80	0.03	0	0	0.03	0	0	0.14
Brûleur	0	1.00	0	0	0	0	0	0
Echangeur	0	0	1.00	0	0	0	0	0
Courbe de chauffe	0.14	0	0	0.74	0.03	0	0	0.09
Relance précoce	0	0	0	0	1.00	0	0	0
Relance tardive	0	0	0	0	0	0.86	0	0.14
Vanne de Régulation	0	0	0	0	0	0	0.97	0.03
PR	0.91							
PFA	0.01							

Tableau 4.37 Probabilités de classification de chaque classe (35 formes) de test (RMS3)

Commentaires :

- la détection des **classes normale, courbe de chauffe et relance tardive** est nettement plus difficile par rapport au réseau RU2.
- en revanche la détection de la **classe vanne de régulation** est plus facile qu'avec un réseau RU2. Peut-on dire que le fait que la détection de la classe courbe de chauffe et des autres classes (normale, brûleur, échangeur, relance précoce, relance tardive et vanne de régulation) utilise deux réseaux imbriqués facilite la détection du défaut de vanne de régulation, en particulier quand on applique les réseaux à un système non appris dont la consigne de température intérieure de la période d'inoccupation a été modifiée.

(3) Conclusion

Les résultats de test des réseaux RU2 et RMS3 montrent que la généralisation de réseaux appris par un système de chauffage spécifique simulé vers un autre système simulé dont la consigne de température intérieure de la période d'inoccupation a changé est tout-à-fait possible, restant à savoir si cette généralisation est toujours possible vers un système réel de ce type.

Comme les trois systèmes testés précédemment, la performance du réseau unique (ici RU2) est également meilleure que le réseau multiple (ici RMS3).

4.3.5 - CONCLUSION

On a, jusqu'à maintenant, testé la capacité de généralisation des deux réseaux (RU2 et RMS3) en utilisant cinq systèmes de chauffage. Parmi ces cinq systèmes, un dénommé système spécifique (système_spf) a été utilisé pour l'apprentissage et le test, et les quatre autres (système_occ, système_rad, système_iner et système_inocc) n'ont été utilisés que pour le test.

Les probabilités de reconnaissance et de fausse alarme (PR et PFA) sont récapitulées dans le Tableau 4.38 sachant que chacun des cinq systèmes comporte 245 formes de test (i.e. 7 classes dont chaque classe a 35 formes).

Système		Système_spf	Système_occ	Système_rad	Système_iner	Système_inocc	Moyenne
PR	RU2	0.95	0.76	0.87	0.92	0.95	0.89
	RMS3	0.89	0.55	0.87	0.80	0.91	0.80
PFA	RU2	0.01	0.02	0.03	0.02	0.004	0.02
	RMS3	0.04	0.30	0.05	0.15	0.01	0.11

Tableau 4.38 PR et PFA de test des cinq systèmes avec deux réseaux RU2 et RMS3

On constate que le réseau unique RU2 présente une performance nettement meilleure que le réseau multiple RMS3, quel que soit le système testé excepté le système_rad ; ceci s'explique peut-être par le fait que le réseau unique peut plus facilement acquérir une connaissance globale que le réseau multiple (une confirmation de la conclusion faite à la fin du paragraphe 4.2.2.4).

Avec le réseau RU2, le taux de reconnaissance moyen est de 89 % et le taux de fausse alarme moyen n'est que 2 %, ce qui représente environ deux fausses alarmes (i.e. deux jours pour un diagnostic journalier) tous les cinq mois⁽⁴⁾, donc des fausses alarmes minimales. Il faut signaler que le taux de reconnaissance peut être parfois très différent entre les différents systèmes de chauffage testés.

⁽⁴⁾ 5 mois $\approx$ [100 (jours)] / [5 (jours/semaine) x 4 (semaines/mois)]

En ce qui concerne les taux de reconnaissance de chacune des 7 classes à discriminer, les résultats de test du réseau RU2 montrent :

- que la détection de la **classe normale** du système_spf et du système_inocc est très satisfaisante (respectivement 91 % et 94 % de réussite) ; mais que la détection de cette classe des trois autres systèmes (système_occ, système_rad et système_iner) est médiocre, ceci est probablement dû au programmeur-régulateur utilisé dans la simulation, car la classe normale ne se confond qu'avec la classe relance précoce et/ou la classe courbe de chauffe quel que soit le système testé. On pourrait espérer une meilleure détection de la classe normale en utilisant un programmeur-régulateur plus performant. On résume que **dans le cas où un système est en fonctionnement normal, les fausses alarmes (s'il y en a) indiquent plus probablement des défauts de programmation-régulation que des pannes de fonctionnement d'un équipement**. De plus une grande partie de la classe normale est rejetée par le réseau RU2 car ambiguë, mais n'entraîne pas de fausse alarme, donc pas très gênante pour les équipes de maintenance,
- que la détection de la **classe brûleur est globalement satisfaisante** (taux de reconnaissance moyen des cinq systèmes est de 90 %). Parmi les 175 formes de test de l'ensemble des cinq systèmes (cinq systèmes dont chaque système a 35 formes), il n'y a que deux formes (une du système_spf et une autre du système_iner) classées dans la classe relance tardive, mais aucune confusion avec les autres classes comportant des défauts,
- que le taux de détection de la **classe échangeur est 100 %**, donc pas de difficulté de détection de cette classe ni fausse alarme due à la présence de cette classe,
- que le taux de détection de la **classe courbe de chauffe des système_spf, système_rad, système_iner et système_incc est très satisfaisante (100 %)** ; qu'en revanche la **détection de la classe courbe de chauffe du système_occ est très difficile** (seulement 52 % de réussite) sachant qu'il n'y a toujours pas de fausse alarme due à la présence de cette classe,
- que la détection de la **classe relance précoce des système_spf, système_iner et système_inocc est satisfaisante** (respectivement 91 %, 89 % et 97 % de réussite), et qu'il n'y a pas de fausse alarme en présence de cette classe. En revanche **pour le système_occ et le système_rad, la détection est beaucoup plus difficile** (respectivement 37 % et 77 % de réussite) ; de plus cette classe peut être classée soit dans la classe vanne de régulation (pour le système_occ) soit dans la classe courbe de chauffe (pour le système_rad). Notons que le réseau multiple RMS3 détecte généralement mieux la classe relance précoce que le réseau unique RU2,
- que la détection de la **classe relance tardive est satisfaisante** (taux de reconnaissance moyen des cinq systèmes est de 95 %) sachant que pour le système_rad, le taux de reconnaissance de cette classe (83 %) est nettement inférieur à la valeur moyenne,
- que la détection de la **classe vanne de régulation est satisfaisante** (taux de reconnaissance moyen des cinq systèmes est de 92 %) sachant que pour le système_inocc, le taux de reconnaissance de cette classe (82 %) est nettement inférieur à la valeur moyenne.

En ce qui concerne le programmeur-régulateur, on résume :

- que la non ou mauvaise détection, toutes les classes confondues, est généralement liée au programmeur-régulateur (i.e. courbe de chauffe, relance précoce ou tardive) ; ceci est probablement dû à l'utilisation d'un type de celui-ci non optimal dans la simulation,
- qu'il est plus difficile de détecter certaines classes (normale, brûleur et relance précoce) les lundis que les autres jours de la semaine, une raison pouvant être que le programmeur fonctionne moins bien les lundis (i.e. après une longue coupure de chauffage) que les autres jours de la semaine.

CONCLUSION

Nous avons développé dans ce mémoire une méthode de détection et de diagnostic de défauts pour les systèmes de génie climatique en nous appuyant sur un système de chauffage collectif à eau chaude.

Le développement d'une méthode de détection et de diagnostic impose qu'on dispose d'une base de données représentant le fonctionnement du système comportant différents défauts.

Cette base de données doit être la plus fiable possible pour que la méthode développée soit robuste et utilisable sur un système réel. Pour cette raison, nous avons d'abord appliqué deux approches théoriques de classification de défauts (analyse des dérives de variable et de leurs causes, analyse des modes de défaillance et de leurs effets), ensuite lancé sous forme de questionnaire une enquête auprès d'un nombre important d'exploitants de systèmes de chauffage.

Ce travail nous a permis d'une part de ressortir les défauts les plus importants en fonction de la fréquence d'apparition, de la difficulté de détection et de leur influence sur le confort et les consommations d'énergie, d'autre part de tirer certaines conclusions intéressantes :

- les défauts ayant un impact important sur le confort sont généralement bien détectés soit au début de la vie d'un système de chauffage, soit par l'équipe de maintenance au cours de fonctionnement du système,
- en revanche, les défauts qui ont un impact important uniquement sur les consommations sont aujourd'hui difficiles à détecter,
- de nombreux défauts ayant un impact tant sur le confort que sur les consommations apparaissent au niveau de la programmation-régulation sans qu'il existe de méthode simple pour les détecter.

Ce travail nous a permis également de sélectionner une première série des défauts pour lesquels il paraissait le plus utile de développer des méthodes de détection diagnostic.

D'abord, ces défauts ont été modélisés.

Ensuite nous avons simulé, en utilisant les logiciels Matlab-Simulink, les modes de fonctionnement d'un système de chauffage de référence comportant ou non ces défauts. Ces modes de fonctionnement (ou classes en terme de reconnaissance des formes) sont les suivants :

- classe « normale » correspondant à un mode de fonctionnement normal de référence,
- classe « brûleur » correspondant au mauvais fonctionnement (mauvaise combustion) du brûleur,
- classe « échangeur » correspondant à l'entartrage ou l'encrassement de l'échangeur d'une chaudière,
- classe « courbe de chauffe » correspondant à la courbe de chauffe trop basse,
- classe « relance précoce » correspondant à un défaut du programmeur d'intermittence,
- classe « relance tardive » correspondant à un défaut du programmeur d'intermittence,
- classe « vanne de régulation » correspondant à un défaut de fuite de vanne en position fermée.

Notons que la modélisation-simulation a pris dans un premier temps des défauts francs afin de bien mettre en évidence les différents types de défaut.

Enfin cette modélisation-simulation a permis de simuler quatre autres systèmes de chauffage comportant les mêmes types de défaut.

Disposant d'une base de données simulée, nous avons pu développer une méthode de détection diagnostic de défauts. Après une phase de recherche bibliographique sur les différentes techniques utilisables pour développer une méthode de détection et de diagnostic, nous avons retenu une méthode basée sur la reconnaissance des formes utilisant, comme outil de classification, une méthode connexionniste (réseaux de neurones artificiels). Cette méthode impose deux étapes distinctes : l'apprentissage de réseaux et le test de leurs capacités de généralisation, ce qui impose un ensemble de données pour l'apprentissage et un autre ensemble pour le test.

Ces deux étapes se sont effectuées de façon successive et récursive, c'est-à-dire que si le résultat de test n'était pas satisfaisant, il fallait réapprendre et/ou redéfinir la structure de réseau afin de trouver la meilleure règle de discrimination de différentes classes.

En utilisant seulement des données issues de la simulation du système de référence, l'apprentissage comprend une phase de prétraitement de données et une phase d'établissement d'une règle de discrimination de différentes classes.

Premièrement, puisque les classes ont été déjà définies selon les différents types de défaut, le prétraitement de données a consisté principalement à choisir un espace des formes c'est-à-dire les paramètres pertinents correspondant aux composantes d'un vecteur forme, qui sépare le mieux possible les différentes classes à discriminer. Pour ce faire, nous avons appliqué une approche d'analyse physique suivie par une stratégie de normalisation.

La démarche adoptée pour cette normalisation est de choisir les paramètres de normalisation de telle sorte que le paramètre normalisé soit « indépendant » du système de chauffage appris ; ceci a en effet facilité la généralisation de la règle de discrimination (structure de réseaux apprise) vers d'autres systèmes de chauffage collectif à eau chaude sans nécessité de réapprentissage de celle-ci.

Ce point est très important en terme d'application sur le terrain, car il est souhaitable qu'un outil de détection diagnostic soit appris par un seul système de chauffage (de préférence simulé), mais diagnostique correctement les autres systèmes sans réapprendre cet outil.

Deuxièmement, l'établissement d'une règle de discrimination a consisté à ranger les données issues du prétraitement dans différentes classes en construisant les frontières entre ces classes. Pour ce faire, nous avons utilisé un modèle connexionniste, dénommé réseaux Perceptron Multi-Couches, en appliquant un algorithme d'apprentissage dit Rétro-Propagation (RP) améliorée : Optimisation Levenberg-Marquardt. Notons que cet algorithme a montré dans notre étude une performance nettement meilleure par rapport à une RP classique.

Nous avons appris et testé, en soulignant leurs forces et leurs faiblesses, deux catégories de réseaux (réseaux uniques et réseaux multiples) en utilisant d'une part le système de référence pour l'apprentissage, d'autre part le système de référence et quatre autres systèmes pour le test.

Finalement, nous avons retenu une structure de réseau unique RU2 comportant une couche d'entrée, une couche cachée avec seulement 5 neurones et une couche de sortie, pour la discrimination de l'ensemble des 7 classes. Notons que la taille (i.e. nombre de neurones) de la couche d'entrée est conditionnée par l'espace des formes et celle de la couche de sortie est définie en fonction du nombre de classes à discriminer, mais en revanche le nombre de couches cachées et la taille de chacune de celles-ci sont entièrement indéterminés. Par conséquent, il a fallu essayer (apprendre et tester) un nombre important de structures de réseaux pour sélectionner la meilleure.

Cette approche d'apprentissage-test nous a permis de tirer une première série de conclusions sur la capacité de généralisation du réseau RU2 vers des systèmes de chauffage appris ou non :

- la détection de l'ensemble des 7 modes de fonctionnement sur le système de référence (i.e. utilisé aussi pour l'apprentissage) ou le système dont la consigne de température intérieure de la période d'inoccupation a été modifiée, est très satisfaisante avec un taux de réussite d'environ 95% et un risque de fausse alarme inférieur à 1%,
- quand on augmente la puissance (surface) des radiateurs ou l'inertie du bâtiment, le réseau permet également de détecter correctement ces 7 modes de fonctionnement avec un taux de réussite de l'ordre de 90% et un taux de fausse alarme d'environ 2%,
- en revanche, pour le système dont la consigne de température intérieure de la période d'occupation a été modifiée (par exemple de 19°C à 21°C), le taux de réussite n'est que 76%, mais le risque de fausse alarme est toujours très faible (environ 2%). Notons toutefois que les défauts non détectés sont principalement liés au programmeur-régulateur (courbe de chauffe ou relance de chauffage), et que les défauts de composants tels que brûleur, échangeur et vanne de régulation sont très bien détectés (environ 90% de réussite). Cela explique peut-être le fait qu'un programmeur non optimal (malheureusement ce qui existe dans le marché de régulation actuel !) a une influence non-négligeable sur la détection diagnostic de défauts.

Le système (règle) de décision que nous avons obtenu semble être adéquat pour le diagnostic des défauts portant sur la chaudière (mauvaise combustion, entartrage ou encrassement, etc.), les vannes de régulation (fuite de la boucle primaire vers la boucle secondaire, etc.) et les programmeurs-régulateurs (mauvais réglage de courbe de chauffe, mauvais calculs de relance de chauffage, etc.).

Certes, nous n'avons pu qu'ébaucher la construction d'un prototype de méthode de détection diagnostic pour les systèmes de chauffage collectif à eau chaude, puisque la capacité de généralisation du réseau n'a été testée que par les systèmes de chauffage simulés ; et par ailleurs il ne faut pas oublier que nous n'avons simulé que des défauts francs pour développer ce système de décision. On peut même se demander par là : que se passera-t-il si un défaut se présente dans un système de chauffage avec un degré de sévérité de moindre importance ?

Cependant les résultats de la généralisation des réseaux, en les appliquant aux systèmes de chauffage simulés, sont suffisamment prometteurs pour penser que la généralisation du réseau vers des systèmes réels sur le terrain devrait donner des résultats intéressants. On propose donc de poursuivre ce travail principalement dans les quatre directions suivantes :

- perfectionnement (si nécessaire) de la procédure de normalisation en vue d'améliorer encore la détection des défauts des systèmes non utilisés en phase d'apprentissage,
- validation de l'outil de détection diagnostic ainsi développé sur des systèmes réels,
- implantation de celui-ci dans un système de Gestion Technique du Bâtiment en collaborant avec des industriels ; les contacts pris avec l'industriel leader du marché français prouvent son intérêt pour cette démarche,
- application de la méthodologie développée dans cette thèse à d'autres systèmes de génie climatique.

ANNEXES

Annexe A - Description du système de référence de chauffage collectif à eau chaude	144
ANNEXE A1 - COMPOSANTS DU SYSTEME DE REFERENCE	145
ANNEXE A2 - VARIABLES MESUREES DU SYSTEME DE REFERENCE	149
ANNEXE A3 - DONNEES DE SIMULATION DE SYSTEMES DE CHAUFFAGE	151
 Annexe B - Recensement et classification de défauts	 153
ANNEXE B1 - METHODE D'ANALYSE DES DERIVES DE VARIABLE ET DE LEURS CAUSES	154
ANNEXE B2 - METHODE D'ANALYSE DES MODES DE DEFAILLANCE ET DE LEURS EFFETS	162
ANNEXE B3 - DESCRIPTION DE DEFAUTS TYPQUES (LISTE RESTREINTE)	169
ANNEXE B4 - PRESENTATION DU QUESTIONNAIRE D'ENQUETE	179

Annexe A - Description du système de référence de chauffage collectif à eau chaude

Les sigles que l'on trouve dans les tableaux A1 et A2 signifient :

AEF :	Alimentation Eau Froide
AGN :	Alimentation Gaz Naturel
BP :	Boucle Primaire
BSC :	Boucle Secondaire Chauffage
BSECS :	Boucle Secondaire Eau Chaude Sanitaire
CA :	Circuit Aller
CDCA :	Circuit Distribution Centrale Aller
CDCR :	Circuit Distribution Centrale Retour
CDHE :	Circuit Distribution Horizontale et Emission
CDV :	Circuit Distribution Verticale
CP :	Circuit Primaire
CR :	Circuit Retour
CS :	Circuit Secondaire
DC :	Dispositif de contrôle
EF :	Evacuation des Fumées
ENV :	ENVironnement
PCCC :	Programmation et Conduite des Chaudières en Cascade
SE :	Système d'Expansion
VMC :	Ventilation Mécanique Contrôlée

ANNEXE A1 - COMPOSANTS DU SYSTEME DE REFERENCE

Tableau A1 Liste de composants

Sous process	Position	Composant	Nombre	Numéro
BP	CA	chaudière	2	1
		brûleur	2	2
		échangeur principal	2	3
		rampe gaz interne (filtre et régulateur)	2	4
		contrôleur de débit	2	5
		robinet de vidange	2	6
		purgeur automatique	2	7
		soupape de sûreté	2	8
		thermomètre	2	9
		manomètre	2	10
		vanne d'arrêt	2	11
		bride	4	12
		tube en acier noir		13
		calorifuge		14
		bouteille casse-pression	1	15
	CR	vanne d'arrêt	3	16
		clapet anti-retour	2	17
		bride	10	18
		filtre à tamis	1	19
		accélérateur	2	20
		tube acier		21
		calorifuge		22
	DC	aquastat de sécurité	2	23
		aquastat de réglage	2	24
		sonde de température d'eau chaude	5	25
		capteur de débit d'eau chaude	2	26
		compteur d'énergie	1	27
AGN		coffret type S 2300 (arrêt, détente et comptage)	1	28
		vanne gaz dans le coffret rouge	1	29
		robinet de purge	1	30
		manomètre	2	31
		vanne d'arrêt	2	32
		filtre gaz	2	33
		tube fer noir		34
	DC	capteur de débit gaz	1	35

Sous process	Position	Composant	Nombre	Numéro
AEF		vanne d'arrêt	4	36
		filtre à tamis	1	37
		disconnecteur	1	38
		pot doseur anticorrosion	1	39
		clapet anti-retour	1	40
		soupape de sûreté	1	41
		thermomètre	1	42
		manomètre	1	43
		robinet de vidange	1	44
		(raccordé à l'égout avec écoulement visible)		
		tube fer noir		45
	DC	capteur de débit d'eau froide	1	46
		compteur d'eau	1	47
SE		robinet de purge gaz	1	48
		manomètre	1	49
		soupape de sûreté	1	50
		robinet de vidange	1	51
		vase d'expansion à membrane	1	52
		tube fer noir		53
	DC	capteur de pression d'eau chaude	1	54
EF		conduit de fumée	1	55
		conduit de raccordement	1	56
		anti refoaleur	1	57
		extracteur des fumées	1	58
		réglage de l'extraction	1	59
		pressostat différentiel	1	60
		thermomètre de fumée	1	61
	DC	sonde de température de fumée	1	62
BSC	CDCA	vanne d'arrêt	4	63
		accélérateur double	2	64
		manchon anti-vibratile	4	65
		clapet anti-retour	2	66
		manomètre	2	67
		filtre à tamis	2	68
		vanne 2 voies d'équilibrage de pression	2	69
		robinet à boisseau	2	70
		thermomètre	2	71

Sous process	Position	Composant	Nombre	Numéro
BSC	CDCA	tube acier noir		72
		calorifuge		73
	CDCR	thermomètre	2	74
		vanne d'arrêt	4	75
		clapet anti-retour	1	76
		tube acier noir		77
		calorifuge		78
	CDV	vanne d'arrêt	4	79
		robinet de vidange	2	80
		tube acier noir		81
	CDHE	calorifuge		82
		radiateur	140	83
		purgeur à clef	140	84
		raccord	140	85
		console	140	86
		robinet à boisseau	40	87
		tube cuivre		88
	DC (CDCA)	vanne trois voies	2	89
		servomoteur	2	90
		sonde de température d'eau	3	91
		capteur de débit d'eau	2	92
		capteur de pression différentielle	2	93
	DC (CDCR)	sonde de température d'eau	2	94
		compteur d'énergie	2	95
	DC (CDHE)	tête thermostatique	140	96
		capteur de débit d'eau	140	97
		sonde de température d'air ambiante	140	98
BSECS	CP	accélérateur double	1	99
		clapet anti-retour	1	100
		préparateur d'ECS	1	101
		vanne d'arrêt	2	102
		tube acier noir		103
		calorifuge		104

Sous process	Position	Composant	Nombre	Numéro
BSECS	CS	manomètre	2	105
		thermomètre	2	106
		vanne d'arrêt	2	107
		soupape de sûreté	1	108
		accélérateur double	1	109
		clapet anti-retour	1	110
		égout	1	111
		robinet de puisage	20	112
		tube cuivre		113
	DC	régulateur	1	114
		servomoteur	1	115
		vanne trois voies	1	116
		aquastat de sécurité	1	117
		sonde de température d'eau chaude	4	118
		sonde de température d'eau froide	1	119
		capteur de débit d'eau chaude	3	120
		capteur de débit d'eau froide	1	121
		capteur de pression d'eau chaude	2	122
		vanne thermostatique	1	123
		compteur d'énergie	2	124
		compteur d'eau	1	125
VMC		bouche d'entrée d'air acoustique	220	126
		chicane	120	127
		soupape d'évacuation	20	128
		grille de reprise	10	129
		clapet coupe-feu	10	130
		extracteur	2	131
		gaine spiralee		132
	DC	sonde de température d'air ambiant	2	133
ENV	DC	Sonde de température extérieure	1	134
		sonde détection gaz / Co	1	135
		sonde détection incendie	1	136
PCCC	DC	compensateur en fonction des conditions extérieures	2	137
		programmateur de l'intermittence	2	138
		régulateur de température de départ secondaire	2	139
		limiteur de température de retour primaire	1	140
		"gestionnaire" des chaudières en cascade	1	141
		régulateur des aquastats des chaudières	2	142

ANNEXE A2 - VARIABLES MESUREES DU SYSTEME DE REFERENCE

Tableau A2 Liste des variables mesurées

Sous process	Variable mesurée	Sigle
BP	Température Sortie Chaudière 1	TSC 1
	Température Entrée Chaudière 1	TEC 1
	Débit d'Eau Chaudière 1	DEC 1
	Température Sortie Chaudière 2	TSC 2
	Température Entrée Chaudière 2	TEC 2
	Débit d'Eau Chaudière 2	DEC 2
	Heure de Fonctionnement Chaudière 1	HFC 1
	Heure de Fonctionnement Chaudière 2	HFC 2
	Débit d'eau dans la bouteille casse-pression	DEBCP
	Température Départ Primaire Commune	TDPC
	Température Retour Primaire Commune	TRPC
	Etat Accélérateur Chaudière 1	EAC 1
	Etat Accélérateur Chaudière 2	EAC 2
	Contrôle Accélérateur Chaudière 1	CAC 1
	Contrôle Accélérateur Chaudière 2	CAC 2
	Contrôle Brûleur 1	CB 1
	Contrôle Brûleur 2	CB 2
AGN	Débit Gaz Naturel	DGN
AEI	Débit d'Eau Froide	DEF
	Propriété Electrochimique d'Eau Froide	PEEF
SE	Pression d'Eau Chaude	PEC
EF	Température des Fumées Chaudière 1	TFC 1
	Température des Fumées Chaudière 2	TFC 2
BSC	Température Départ Secondaire Commun	TDSC
	Température Départ Circuit Nord	TDCN
	Température Départ Circuit Sud	TDCS
	Température Retour Circuit Nord	TRCN
	Température Retour Circuit Sud	TRCS
	Débit Circuit Commun Nord	DCCN
	Débit Circuit Commun Sud	DCCS
	Etat Accélérateur 1 Circuit Nord	EACN 1
	Etat Accélérateur 2 Circuit Nord	EACN 2
	Etat Accélérateur 1 Circuit Sud	EACS 1
	Etat Accélérateur 2 Circuit Sud	EACS 2
	Contrôle Accélérateur 1 Circuit Nord	CACN 2
	Contrôle Accélérateur 2 Circuit Nord	
	Contrôle Accélérateur 1 Circuit Sud	
	Contrôle Accélérateur 2 Circuit Sud	

Sous process	Variable mesurée	Sigle
BSC	Pression Différentielle Circuit Nord Pression Différentielle Circuit Sud Contrôle Vanne Trois Voies Nord Contrôle Vanne Trois Voies Sud Position Vanne Trois Voies Nord Position Vanne Trois Voies Sud Débit Radiateur Y Etage X Circuit N Débit Radiateur Y Etage X Circuit S Température Pièce Y Etage X Circuit N Température Pièce Y Etage X Circuit S (Nota : X = 1, 2, ...5 ; Y = 1, 2, ...14)	PDCN PDCS CVTVN CVTVS PVTN PVTVS DRN XY DRS XY TPN XY TPS XY
BSECS	Température Entrée Primaire Température Sortie Primaire Débit Circuit Primaire Température Eau Froide Débit Eau Froide Température Départ Secondaire Température Retour Secondaire Pression Départ Secondaire Pression Retour Secondaire Débit Départ Secondaire Débit Retour Secondaire	TEP TSP DCP TEF ECS DEF ECS TDS TRS PDS PRS DDS DRS
VMC	Température Extraction Nord VMC Température Extraction Sud VMC Etat VMC	TN VMC TS VMC E VMC
ENV	Température EXTérieure Détection Gaz/Co Détection Incendie	TEXT DGC DI

ANNEXE A3 - DONNEES DE SIMULATION DE SYSTEMES DE CHAUFFAGE

Les simulations se sont effectuées sous l'environnement SIMULINK. leurs résultats (variables) ont été stockés dans l'espace de travail (workspace en anglais) MATLAB sous forme *fichier.mat*.

Pour chaque classe simulée, les fichiers, comportant 54 variables au total, sont les suivants :

(1) Horloge de simulation (fichier : temps.mat)

1 Variable : temps de simulation

(2) Météo (fichier : text.mat)

1 Variable : temperature exterieure

(3) Conduite en séquence des chaudières (fichier : cond.mat)

6 Variables : temperature consigne aquastat 1
temperature consigne aquastat 2
M/A (Marche/Arrêt) bruleur 1
M/A bruleur 2
M/A pompe 1
M/A pompe 2

(4) Programmation d'intermittence et régulation (fichier : regul.mat)

7 Variables : heure d'une journée (0h - 24h)
jour de la semaine (0 samedi, 1 dimanche, 2 lundi, 3 autres jours)
consigne de chauffage intermittent (0 arrêt, 1 normal, 2 relance, 3 réduit)
position requise de Vanne-3-Voies nord
position requise de Vanne-3-Voies sud
température consigne de Vanne-3-Voies nord
température consigne de Vanne-3-Voies sud

(5) Boucle primaire (génération de chaleur) (fichier : looprim.mat)

16 Variables : température des fumées chaudière 1 (après filtrage pass-bas)
température de départ chaudière 1
température de retour chaudière 1
débit chaudière 1
consommation de gaz accumulée chaudière 1
consommation de gaz instantanée chaudière 1
température des fumées chaudière 2
température de départ chaudière 2
température de retour chaudière 2
débit chaudière 2
consommation de gaz accumulée chaudière 2
consommation de gaz instantanée chaudière 2
température de départ primaire commun
température de retour primaire commun
débit de la bouteille-casse-pression
température des fumées chaudière 1 (avant filtrage pass-bas)

(6) Boucle secondaire - Circuit de distribution nord (fichier : loopsecN.mat)

7 Variables : débit avant Vanne-3-Voies nord
température avant Vanne-3-Voies nord
température de départ secondaire après Vanne-3-Voies nord
position réelle de Vanne-3-Voies nord
température résultante de zone nord
température de retour secondaire nord avant la bouteille casse pression
débit de retour secondaire nord avant la bouteille casse pression

(7) Boucle secondaire - Circuit de distribution sud (fichier : loopsecS.mat)

7 Variables : débit avant Vanne-3-Voies sud
température avant Vanne-3-Voies sud
température de départ secondaire après Vanne-3-Voies sud
position réelle Vanne-3-Voies sud
température résultante de zone sud
température de retour secondaire sud avant la bouteille casse pression
débit de retour secondaire sud avant la bouteille casse pression

(8) Structure de bâtiment (fichier : mur.mat)

3 Variables : température des murs entre zone nord et l'extérieure
température des murs entre zone nord et zone sud
température des murs entre zone sud et l'extérieure

(9) Apports énergétiques (fichier : app.mat)

6 Variables : apports internes zone nord
apports internes zone sud
apports solaires zone nord
apports solaires zone sud
puissance totale zone nord (radiateur+apports internes+apports solaires)
puissance totale zone sud (radiateur+apports internes+apports solaires)

L'objectif de stocker un nombre si important de variables est double :

- d'une part, le fonctionnement du système (chauffage intermittent, régulation en fonction de température extérieure, couplage bâtiment-système, etc.) étant complexe, les comportements dynamiques de ces variables sont très utiles pour vérifier et justifier la crédibilité du simulateur développé,

- d'autre part, l'ensemble des variables constitue d'une base de données représentant le fonctionnement du système comportant différents défauts. Cette base de données devrait nous permettre de ressortir des variables portant les informations de défauts les plus significatives et les plus pertinentes, qui seront utilisées pour développer des méthode de détection et de diagnostic.

Annexe B - Recensement et classification de défauts

ANNEXE B1 - METHODE D'ANALYSE DES DERIVES DE VARIABLE ET DE LEURS CAUSES

TSC1 (ou TSC2) : Température Sortie Chaudière 1

trop haute

- dysfonctionnement du thermostat de sécurité de chaudière
- mauvais réglage du brûleur (voir CB1)
- débit d'eau de la chaudière 1 trop bas (voir DEC1)

trop basse

brûleur de sécurité :

- alarme électrique,
- alarme de débit d'eau de chaudière,
- alarme de pression de gaz naturel,
- alarme de sécurité de gaz naturel.

brûleur/chaudière sous-dimensionnés

mauvais fonctionnement du brûleur

rendement de chaudière trop bas

mauvais réglage du thermostat-limiteur de chaudière

déperdition thermique par les tuyaux de distribution trop importante (dégradation de l'isolation thermique)

mauvais réglage du brûleur (voir CB1)

débit d'eau de la chaudière 1 trop haut (voir DEC1)

TEC1 trop basse (voir TEC1)

fluctuation

brûleur sur-dimensionné

mauvais réglage du brûleur (voir CB1)

fluctuation du débit d'eau de la chaudière 1 (voir DEC1)

TEC1 (ou TEC2) : Température d'Entrée de Chaudière 1

trop haute

TRPC trop haute (voir TRPC)

débit d'eau de la chaudière 1 trop bas (voir DEC1)

trop basse

déperdition thermique par les tuyaux de distribution trop importante (dégradation de l'isolation thermique)

TRPC trop basse (voir TRPC)

débit d'eau de la chaudière 1 trop haut (voir DEC1)

fluctuation

fluctuation de la TRPC (voir TRPC)

fluctuation du débit d'eau de la chaudière 1 (voir DEC1)

DEC1 (DEC2) : Débit d'Eau de la Chaudière 1*nul*

pompe de recyclage de la chaudière 1 ne marche pas

mauvais fonctionnement de la pompe de recyclage de la chaudière 1

vanne d'arrêt à l'entrée de la chaudière 1 fermée

mauvais fonctionnement du clapet anti-retour à l'entrée de la chaudière 1

trop bas

mauvais fonctionnement de la pompe de recyclage de la chaudière 1

bouchage de tuyaux par la boue

mauvais équilibrage entre les circuits de chaudière 1 et chaudière 2

vanne d'arrêt à l'entrée de la chaudière 1 non-complètement ouverte

bouchage de l'échangeur principal de la chaudière 1

pompe de recyclage de la chaudière 1 sous-dimensionnée

air dans les tuyaux

trop haut

pompe de recyclage de la chaudière 1 sur-dimensionnée

mauvais équilibrage entre les circuits de chaudière 1 et chaudière 2

fluctuation

fluctuation de la position de la vanne-trois-voies

boucle primaire non-indépendante de la boucle secondaire

TFC1 (ou TFC2) : Température des Fumées de Chaudière 1*trop haute*

encrassement de l'échangeur principal de la chaudière 1

excès d'air trop important au niveau de combustion

trop basse

condensation de la vapeur d'eau des fumées dans la cheminée : dégradation de la cheminée

HFC1 (ou HFC2) : Heures de Fonctionnement de Chaudière 1*trop haute*

rendement de la chaudière 1 trop bas

HFC1 et HFC2 simultanément moyennes

mauvais fonctionnement du GS (Gestionnaire en Séquence des chaudières)

TDPC : Température de Départ de la boucle Primaire Commune*trop basse*

perte thermique par les tuyaux trop importante (dégradation de l'isolation thermique)

TSC1 et/ou TSC2 trop basses (voir TSC1)

DEC1 et/ou DEC2 trop hauts (voir DEC1)

trop haute

TSC1 et/ou TSC2 trop hautes (voir TSC1)

DEC1 et/ou DEC2 trop bas (voir DEC1)

fluctuation

fluctuation de TSC1 et/ou TSC2, DEC1 et/ou DEC2 (voir TSC1, DEC1)

TRPC : Température de Retour de la boucle Primaire Commune*trop basse*

bypass (bouteille casse-pression) mal-dimensionné

déperdition thermique par les tuyaux trop importante (dégradation de l'isolation thermique)

TDCN (et/ou TDCS) trop basse (voir TDCN)

TRCN (et/ou TRCS) trop basse (voir TRCN)

DEBCP trop bas (voir DEBCP)

TDPC trop basse (voir TDPC)

trop haute

TDPC trop haute (voir TDPC)

TRCN (et/ou TRCS) trop haute (voir TRCN)

fluctuation

fluctuation de TDPC (voir TDPC)

fluctuation de TRCN (et/ou TRCS) (voir TRCN)

DEBCP : Débit d'Eau de la Bouteille Casse-Pression*nul*

bouchage complet des tuyaux

bouchage complet de la bouteille casse-pression par des boues

mauvais fonctionnement des pompes de recyclage des chaudières 1 et 2

mauvais réglage des pompes de recyclage des chaudières 1 et 2 par GS.

DEC1 et DEC2 nuls (voir DEC1)

trop bas

bouchage des tuyaux
 bouchage de la bouteille casse-pression par des boues
 bouteille casse-pression sous-dimensionnée
 DEC1 et/ou DEC2 trop bas (voir DEC1)

trop haut

bouteille casse-pression sur-dimensionnée
 DEC1 et/ou DEC2 trop hauts (voir DEC1)

négatif

débit des boucles secondaires supérieur à celui de la boucle primaire

EAC1 (ou EAC2) : Etat de l'Accélérateur (pompe) de la Chaudière 1*différent de CAC1*

mauvais fonctionnement de la pompe

CAC1 (ou CAC2) : Contrôle de l'accélérateur de la Chaudière 1

pompe marche au moment où elle doit s'arrêter ou pompe s'arrête au moment où elle doit marcher

consigne du thermostat de chaudière mal réglée par CC1(Contrôleur de la température de Chaudière1)
 mauvais réglage de la séquence entre pompe de recyclage et brûleur

CB1 (ou CB2) : Contrôle du brûleur 1*problèmes quelconques*

CC1 ne fonctionne pas
 mauvaise commande de CC1 par GS

fluctuation

mauvais réglage de CC1
 mauvais réglage de la séquence entre pompe de recyclage et brûleur
 fluctuation de DEC1 (voir DEC1)

TDSC : Température de départ des boucles secondaires communes*trop basse*

dysfonctionnement de GS
 débit d'eau dans la boucle primaire trop bas ($DEC1+DEC2 < DCCN+DCCS$: DEBCP négatif)
 TSC1 et/ou TSC2 trop basse (voir TSC1)

trop haute

dysfonctionnement de GS
TSC1 et/ou TSC2 trop hautes (voir TSC1)

fluctuation

fluctuation des signals de sortie de GS
fluctuation des TSC1 et/ou TSC2 (voir TSC1)

TDCN (ou TDCS): Température de Départ du Circuit Nord*trop basse*

fuite du bypass vers la voie directe de la vanne-trois-voies du circuit nord
dysfonctionnement du RTD (Régulateur de la Température de Départ du circuit nord)
déperdition thermique par les tuyaux trop importante (dégradation de l'isolation thermique)
TDSC trop basse (voir TDSC)
DCCN trop bas (voir DCCN)

trop haute

fuite de la voie directe vers le bypass de la vanne-trois-voies du circuit nord
dysfonctionnement du RTD (Régulateur de la Température de Départ du circuit nord)
TDSC trop haute (voir TDSC)

fluctuation

RTD n'est pas adapté à la vanne-trois-voies
mauvaise autorité de la vanne-trois-voies
hystérésis mécanique de la vanne-trois-voies très importante
embouage de la vanne-trois-voies
TDSC trop basse (voir TDSC)

TRCN (ou TRCS) : Température de Retour du Circuit Nord*trop basse*

dysfonctionnement de la pompe du circuit secondaire
dysfonctionnement de vannes thermostatiques dans des pièces (ouverture complète de vanne par exemple : TPNXY trop haute)
déperdition thermique trop importante de la zone concernée du bâtiment (ouverture inadéquate de fenêtre, vent très fort, ...)
déperdition thermique par les tuyaux trop importante (dégradation de l'isolation thermique)
TDCN trop basse (voir TDCN)
TPNXY trop basse : émission importante de chaleur par radiateur (voir TPNXY)
DCCN trop bas (voir DCCN)

trop haute

apports (chaleur) gratuits très important dans la zone concernée (solaire, humain, par machines, ...)

vannes thermostatiques fermées dans certaines pièces (dysfonctionnement ou action humaine)

embouage de certains radiateurs

vannes-deux-voies dans certains circuits d'étage fermées par accident

PDCN trop basse (voir PDCN)

TDCN trop haute (voir TDCN)

TPNXY trop haute (voir TPNXY)

fluctuation

fluctuation de TDCN (voir TDCN)

TDCN trop haute (donc vanne thermostatique de pièce presque fermée cause un pompage de cette vanne)

DCCN (ou DCCS) : Débit d'eau du Circuit Commun Nord*trop haut*

mauvais équilibrage hydraulique entre circuit nord et circuit sud

trop bas

mauvais équilibrage hydraulique entre circuit nord et circuit sud

dysfonctionnement de la pompe du circuit nord

dysfonctionnement de la vanne d'équilibrage de pression du circuit nord

présence de boues dans des tuyaux ou des radiateurs

dysfonctionnement du PI (Programmateurs d'intermittence)

DEC1+DEC2 trop bas (voir DEC1)

fluctuation

cavitation de la pompe du circuit nord

fluctuation de la PVTVN (voir PVTVN)

EACN1 (ou EACN2 ou EACS1 ou EACS2) : Etat Accélérateur 1 du Circuit Nord*différent de l'état normal*

dysfonctionnement du PI

dysfonctionnement de la pompe

PDCN (ou PDCS) : Pression Différentielle du Circuit Nord*trop haute*

dysfonctionnement de la vanne-deux-voies d'équilibrage de pression (du fait par exemple d'un embouage)

trop basse

fuite de la vanne-deux-voies d'équilibrage de pression

PVTVN (ou PVTVS) : Position de la Vanne-Trois-Voies du circuit Nord*trop fermée*

- dysfonctionnement du RTD (Régulateur de la Température de Départ) du circuit nord
- dysfonctionnement du PI
- dysfonctionnement du LTR (Limiteur de la Température de Retour)
- blocage de la vanne-trois-voies

trop ouverte

- dysfonctionnement du RTD du circuit (ou du PI, ou du LTR)
- blocage de la vanne-trois-voies

DRNXY (ou DRSXY) : Débit du Radiateur de la pièce Y de l'étage X du circuit Nord*trop bas*

- vanne thermostatique du radiateur fermée anormalement
- consigne (température) de la vanne thermostatique trop basse
- mauvais équilibrage hydraulique entre circuit nord et circuit sud

trop haut

- ouverture de fenêtres dans certaines pièces
- ouverture complète anormale de la vanne thermostatique
- consigne (température) de la vanne thermostatique trop haute

TPNXY (ouTPSXY) : Température de la Pièce Y de l'étage X du circuit Nord*trop haute*

- apports gratuits (chaleur) trop importants
- consigne (température) de la vanne thermostatique trop haute
- blocage en position ouverture de la vanne thermostatique
- radiateur de la pièce sur-dimensionné
- fuite de la vanne thermostatique en position fermée
- TDCN trop haute (voir TDCN)

trop basse

- apports gratuits (chaleur) trop bas
- consigne (température) de la vanne thermostatique trop basse
- blocage en position fermée de la vanne thermostatique
- radiateur de la pièce sous-dimensionné
- durée de la relance matinale de chauffage trop courte
- arrêt optimal de chauffage trop tôt
- embouage du radiateur
- ouverture inadéquate de fenêtres
- TDCN trop basse (voir TDCN)
- DRNXY trop bas (voir DRNXY)

fluctuation

- fluctuation des apports gratuits
- autorité insuffisante de la vanne thermostatique

PEEF : Propriétés Electrochimiques de l'Eau Froide*différentes des états normaux*

- dysfonctionnement du pot doseur anticorrosion
- mauvaise étanchéité de pompes
- mauvaise étanchéité de la vase d'expansion
- production des boues

DEF : Débit d'Eau Froide*trop haute*

- fuite d'eau dans le réseau de chauffage
- corrosion du système
- production des boues

PEC : Pression d'Eau Chaude du réseau de chauffage*trop haute*

- dysfonctionnement de la vase d'expansion

ANNEXE B2 - METHODE D'ANALYSE DES MODES DE DEFAILLANCE ET DE LEURS EFFETS

Brûleur et chaudière

manque de gaz

alarme : pression du gaz naturel trop basse

alarme

pression du gaz naturel trop haute

alarme

arrêt automatique du brûleur pour des raisons de sécurité

alarme

surchauffe d'eau

alarme

débit d'eau trop bas dans la chaudière

DEC1 (ou DEC2) trop bas

débit d'air comburant trop haut

taux de O₂ trop haut dans le gaz de combustion

diminution du rendement de chaudière

débit d'air comburant trop bas

augmentation du taux de CO dans le gaz de combustion

changement de couleur du gaz de combustion

diminution du rendement de chaudière

combustion incomplète

diminution du rendement de chaudière

encrassement de l'échangeur principal de chaudière

TFC1 (ou TFC2) trop haute

diminution du rendement de chaudière

TSC1 (ou TSC2) trop basse

dégradation de l'isolation thermique de chaudière

diminution du rendement de chaudière

corrosion de chaudière

fuite d'eau vers le coté fumées ou vers la chaufferie

Vanne-deux-voies

bloquée en position fermée

DEC1 (et/ou DEC2) nul

DCCN (et/ou DCCS) nul

fuite en position fermée

DEC1 (et/ou DEC2) non nul (généralement négatif) quand la vanne est fermée

mauvais rendement de la boucle primaire quand une seule chaudière fonctionne

TSC1 est pratiquement égale à TSC2, et $TSC1 < TEC1$ (et/ou $TSC2 < TEC2$)

encrassement

DEC1 (et/ou DEC2) trop bas

consommation énergétique de la pompe trop importante

Vanne-trois-voies

voie bypass fermée

TDCN (et/ou TDCS) et TRCN (et/ou TRCS) trop hautes

voie directe fermée

TDCN (et/ou TDCS) et TRCN (et/ou TRCS) trop basses

fuite du bypass

TDCN (et/ou TDCS) trop basses pendant la période de relance matinale par exemple

fuite de la voie directe

TDCN (et/ou TDCS) trop hautes pendant la période d'arrêt par exemple

mauvaise autorité hydraulique

fluctuation de TDCN (et/ou TDCS)

fluctuation de TPNXY (et/ou TPSXY)

Pot doseur anticorrosion (système du traitement d'eau d'alimentation de chaudière)

dysfonctionnement

corrosion de tuyaux et de radiateurs

fuite d'eau froide dans la chaufferie

production des boues

pompe arrêtée par des boues

variation de PEEF

DEF trop haut

Pompe

défaillance complète

DEC1 et/ou DEC2 nuls

DCCN et/ou DCCS nuls

DEC1 négatif (si la pompe de recyclage de la chaudière 1 est en défaillance), ...

DEBCP négatif (défaillance de la pompe de recyclage de la chaudière 1 et/ou de la pompe de recyclage de la chaudière 2)

encrassement

DEC1 ou DEC2 ou DCCN ou DCCS trop bas

cavitation

DEC1 ou DEC2 ou DCCN ou DCCS trop bas

bruit

mauvaise étanchéité de pompe

air dans des tuyaux

variation des propriétés électrochimiques d'eau des réseaux de chauffage

tuyau

gel

Radiateur

mauvaise autorité de la vanne thermostatique

instabilité de la position de cette vanne

instabilité de TPNXY ou TPSXY

TPNXY ou TPSXY trop haute ou trop basse

air dans radiateur

bruit

TPNXY ou TPSXY trop basse

bloqué par des boues

TPNXY ou TPSXY trop basse

DEF trop haut

Vase d'expansion

défaillance complète

PEC (Pression d'Eau Chaude) trop haute

mauvaise étanchéité à l'air

variation de PEEF

DEF trop haut

Ventilation Mécanique Controlée (VMC)

arrêt total

- débit d'air neuf nul
- dégradation de la qualité de l'air intérieur
- consommation d'énergie trop basse
- TRCN et/ou TRCS trop hautes

encrassement de la VMC

- diminution du débit d'air neuf
- dégradation de la qualité de l'air intérieur
- consommation d'énergie trop basse
- TRCN et/ou TRCS trop hautes

réglage incorrect de l'horloge de VMC (VMC fonctionne quand le bâtiment est non-occupé)

- baisse rapide de TPNXY et/ou TPSXY pendant la période d'arrêt ou d'entretien de chauffage
- durée de la relance matinale de chauffage trop longue
- TPNXY et/ou TPSXY trop basses au début de la période d'occupation
- consommation d'énergie trop haute

réglage incorrect de l'horloge de VMC (VMC s'arrête quand le bâtiment est occupé)

- dégradation de la qualité de l'air intérieur
- TPNXY et/ou TPSXY trop haute pendant la période d'occupation
- consommation d'énergie trop basse

Bâtiment et occupant

ouverture inadéquate de fenêtre

- baisse rapide de TPNXY ou TPSXY
- ouverture complète de vannes thermostatiques

variation sensible des apports gratuits entre différentes zones

- différences sensibles de TPNXY ou TPSXY entre différentes pièces

variation de la durée d'occupation

- consommation d'énergie trop haute ou trop basse
- variation de la valeur moyenne de TPNXY ou TPSXY

dégradation de l'isolation thermique du bâtiment

augmentation de consommation d'énergie

isolation thermique ne correspond pas au calcul

TRCN et/ou TRCS trop basses

usage incorrect de thermostats

TPNXY ou TPSXY trop hautes ou trop basses

variation de TPNXY ou TPSXY avec des apports gratuits

Sondes de température

biais de sondes

la mesure ne correspond pas à la valeur réelle

variation brutale d'une variable mesurée

sondes de mesure de températures d'eau pas toujours irriguées

fausses valeurs mesurées

sonde de la vanne thermostatique influencée par le radiateur

instabilité de TPNXY ou TPSXY

sondes de mesure de la température intérieure de référence mal positionnées

variation de TPNXY ou TPSXY avec chaleur solaire, apports gratuits, ...

températures intérieures de référence utilisées par le PI (Programmateur Intermittence) pas représentatives

période de relance matinale trop courte

courbe de chauffe pas bien adaptée au bâtiment

TPNXY ou TPSXY trop basses dans certaines pièces

Courbe de chauffe

courbe de chauffe mal réglée : TDCN ou TDCS trop hautes

vannes thermostatiques de radiateurs pratiquement fermées

fluctuationb de TPNXY ou TPSXY

courbe de chauffe mal réglée : TDCN ou TDCS trop basses

TPNXY ou TPSXY trop basses

Programmateurs d'intermittence

commutation trop tôt entre relance matinale et régime normal de chauffage

TPNXY ou TPSXY trop basses au début de la matinée (surtout de lundi matin)

commutation trop précoce entre régime réduit et arrêt (ou régime réduit) de chauffage

TPNXY ou TPSXY trop basse à la fin de l'après-midi

commutation trop tardive entre régime réduit (ou arrêt) et relance matinale de chauffage

TPNXY ou TPSXY trop basse au début de la matinée

commutation trop précoce entre régime réduit (ou arrêt) et relance matinale de chauffage

relance matinale s'arrête avant le début d'occupation

fausses informations de la boucle secondaire transmises au Gestionnaire de chaudière en Séquence (GS)

TDSC trop haute ou trop basse

Régulateur de température de départ

pas adapté à la vanne-trois-voies

fluctuation de TDCN et/ou TDCS

ne fonctionne pas

vanne-trois-voies reste ouverte ou fermée ou en position initiale

TDCN et/ou TDCS différentes de leur valeurs normales

Limiteur de température de retour

mauvaise loi de régulation

TRPC trop basse

voie directe de la vanne-trois-voies reste fermée pendant une durée très longue

TDCN et TDCS trop basses

Gestionnaire de chaudière en séquence

mauvais choix du nombre des chaudières utilisées

durées de fonctionnement de deux chaudières incorrectes

charge moyen du fonctionnement de chaudière trop basse

rendement de chaudière trop bas

Contrôleur de température de chaudière

mauvaise séquence entre le fonctionnement de pompe et celui de brûleur

variation rapide du nombre des chaudières en fonctionnement

marche-arrêt de brûleur trop fréquent : encrassement de l'échangeur, baisse du rendement de chaudière

mauvais réglage

marche-arrêt de brûleur trop fréquent : encrassement de l'échangeur, baisse du rendement de chaudière

fluctuation de TSC1 ou TSC2

TFC1 ou TFC2 trop haute

dysfonctionnement catastrophique : brûleur fonctionne sans arrêt

chaudière en état de sécurité (surchauffe)

dysfonctionnement catastrophique : brûleur ne fonctionne pas

TSC1 et/ou TSC2 trop basses

TDCN et/ou TDCS trop basses

Alimentation de gaz

pas de gaz

alarme de brûleur

vanne d'arrêt de gaz fermée

alarme de brûleur

encrassement du filtre d'alimentation de gaz

alarme du brûleur

TSC1 et/ou TSC2 trop basses

Environnement

température extérieure inférieure à sa valeur pris en compte au niveau de conception

TPNXY et/ou TPSXY trop basses

ANNEXE B3 - DESCRIPTION DE DEFAUTS TYPIQUES (LISTE RESTREINTE)

Composant n°2 : brûleur

Le brûleur est considéré comme un ensemble. Quand un défaut quelconque se présente dans le brûleur, un agent de maintenance doit intervenir pour chercher les causes de ce défaut. Il doit être capable de déterminer par lui-même quelle partie du brûleur est en dysfonctionnement. L'objectif de la détection de défaut est simplement de surveiller si le brûleur fonctionne normalement ou non.

Défaut 1 :	Arrêt anormal
Dérive de variable :	Alarmes (alarme de débit, alarme de sécurité de gaz, alarme de pression de gaz, alarme de surchauffe d'eau, alarme de circuits cassés)
Symptôme :	Température à la sortie de chaudière trop basse, baisse de température intérieure
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire :	Méthodes de détection de défaut existent déjà. Elles sont fournies par fabricant de chaudière
Défaut 2 :	Mauvaise combustion
Dérive de variable :	Dérive de la température des fumées
Symptôme :	Augmentation de consommation de l'énergie, variation de la couleur des, pollution de l'environnement
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	Possibilité de détection par l'O ₂ contenue dans les fumées ou par la dérive de la température des fumées

Composant n°3 : échangeur principal de chaudière.

L'échangeur est considéré comme un ensemble.

Défaut 1 :	Fuite
Dérive de variable :	Alarme du débit d'eau de chaudière
Symptôme :	Eau dans la chaufferie
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire :	Alarme de manque d'eau existe déjà
Défaut 2 :	Encrassement de l'échangeur principal
Dérive de variable :	Température à la sortie de chaudière trop basse
Symptôme :	Mauvaise efficacité de l'échangeur, augmentation de consommation de l'énergie
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Composant n°20, 64 : accélérateur.

Défaut 1 :	Défaillance complète
Dérive de variable :	Débit d'eau nul dans des circuits
Symptôme :	Arrêt anormal du brûleur, pas de besoin de chauffage dans le bâtiment
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	La détection du dysfonctionnement de moteur de pompe (par une alarme interne) existe déjà. Certains défauts mécaniques comme la casse de l'axe de pompe, etc. ne peuvent pas détectés par cette alarme interne. Ils pourraient être détectés par un débitmètre ou un manomètre différentiel.
Défaut 2 :	Encrassement
Dérive de variable :	Débit d'eau trop bas dans des circuits
Symptôme :	Encrassement de pompe de chaudière (boucle primaire) : températures de départ des circuits nord et sud trop basses. Encrassement de pompe de la boucle secondaire : Température intérieure du bâtiment trop basse.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	
Défaut 3 :	Cavitation
Dérive de variable :	Débit d'eau dans des circuits trop bas
Symptôme :	Bruit. Défaut de pompe de chaudière : températures de départ des circuits nord et sud trop basses. Défaut de pompe de la boucle secondaire : température intérieure du bâtiment trop basse
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	
Défaut 4 :	Pompe de chaudière sous-dimensionnée
Dérive de variable :	
Symptôme :	Températures de départ des circuits nord et sud trop basses dans certains cas
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	Ce défaut se présente principalement quand la voie directe de la vanne-trois-voies (n°89) est presque fermée.

Composant n°28 - 35 : alimentation du gaz naturel

Défaut 1 :	Arrêt anormal
Dérive de variable :	Débit de gaz naturel nul
Symptôme :	Arrêt du brûleur
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire	
Défaut 2 :	Fuite de gaz naturel
Dérive de variable :	Débit de gaz trop bas, pression de gaz trop basse.
Symptôme :	Arrêt du brûleur
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire	

Composant n°89 : vanne-trois-voies (et servo-moteur n°90)

Défaut 1 :	Mauvaise autorité hydraulique
Dérive de variable :	Fluctuation de la température de départ de circuit secondaire.
Symptôme :	Température intérieure du bâtiment non stable. Endommagement de vanne.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	
Défaut 2 :	Fuite de la voie bypass ou course de limiteur mal positionnée
Dérive de variable :	Température de départ de circuit secondaire trop basse quand la consigne de température trop haute (par exemple pendant la période de relance matinale de chauffage ou la température extérieure trop basse).
Symptôme :	Température intérieure ne peut pas atteindre à sa consigne au début de l'occupation du bâtiment, ou température intérieure trop basse quand la température extérieure trop basse.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	
Défaut 3 :	Fuite de la voie directe ou course de limiteur mal positionnée
Dérive de variable :	Température de départ de circuit secondaire trop haute quand la consigne de température trop basse (par exemple pendant la période d'arrêt)
Symptôme :	Température intérieure trop haute pendant la période d'inoccupation.
Nécessité de développer des méthodes de détection de défaut :	OUI.
Commentaire :	

Défaut 4 :	Fermeture de la voie bypass
Dérive de variable :	Température de départ du circuit secondaire trop haute.
Symptôme :	Surchauffe du bâtiment. Augmentation de consommation de l'énergie. Ouverture de fenêtre par les occupants ...
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	

Défaut 5 :	Fermeture de la voie directe
Dérive de variable :	Température de départ du circuit secondaire trop basse.
Symptôme :	Souschauffe du bâtiment
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	

Défaut 6 :	Défauts mécanique et électrique
Dérive de variable :	Température de départ de circuit secondaire ne correspond pas à sa consigne
Symptôme :	Fluctuation de la température intérieure
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	

Composant n° 101 : préparateur d'ECS

Défaut 1 :	Entartrage
Dérive de variable :	Débit d'ECS trop bas ou température d'ECS trop basse.
Symptôme :	Plainte d'usager contre l'insuffisance de la température ou du débit d'ECS. Augmentation de consommation de l'énergie
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Composant n°83 : radiateur

Défaut 1 :	Mauvaise autorité de la vanne thermostatique
Dérive de variable :	Instabilité de la température intérieure
Symptôme :	Plainte d'occupant
Nécessité de développer des méthodes de détection de défaut :	à discuter
Commentaire	

Défaut 2 :	Vanne thermostatique dans la pièce de régulation
Dérive de variable :	Température intérieure ne correspond pas à sa consigne (trop haute ou trop basse)
Symptôme :	Plainte d'occupant. Les apports gratuits ne sont pas récupérés
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire	

Défaut 3 :	Air dans radiateur
Dérive de variable :	Température intérieure trop basse
Symptôme :	Bruit. Souschauffe du bâtiment
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire	

Défaut 4 :	panne de la vanne thermostatique
Dérive de variable :	Température intérieure trop haute ou trop basse
Symptôme :	Mauvaise chauffe du bâtiment. Les apports gratuits ne peuvent pas être récupérés
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	

Composant n°38 : Disconnecteur

Défaut 1 :	Défaillance complète
Dérive de variable :	Débit négatif possible de l'eau d'alimentation de chaudière (eau froide)
Symptôme :	Eau dans les circuits de chauffage s'écoule vers les réseaux de distribution d'eau froide, ce qui entraîne une pollution de l'eau froide
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire	

Composant n° 48 - 54 : système d'expansion.

Défaut 1 :	Fuite du gaz de remplissage
Dérive de variable :	Pression de l'eau dans les circuits de chauffage trop haute ou trop basse
Symptôme :	
Nécessité de développer des méthodes de détection de défaut :	à discuter
Commentaire	

Composant n°55 : conduit des fumées

Défaut 1 :	Corrosion
Dérive de variable :	
Symptôme :	Non-étanchéité du conduit des fumées
Nécessité de développer des méthodes de détection de défaut :	à discuter
Commentaire	

Composant : bâtiment et occupant

Défaut 1 :	Ouverture de fenêtre par occupant
Dérive de variable :	Baisse rapide de température intérieure
Symptôme :	Gaspillage de l'énergie. Ouverture de la vanne thermostatique.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Défaut 2 :	Apports gratuits trop importants dans la pièce de conduite
dérive de variable :	Variation rapide de la température de la pièce de conduite
Symptôme :	Température intérieure trop basse dans certaines pièces
Nécessité de développer de méthodes de détection de défaut :	OUI
Commentaire :	

Composant : ensemble de l'installation de chauffage et de production d'ECS.

Défaut 1 :	Mauvais équilibrage hydraulique entre circuit nord et circuit sud.
Dérive de variable :	Température différente entre différentes zones
Symptôme :	Surchauffe et/ou souschauffe dans certaines pièces du bâtiment. Gaspillage de l'énergie
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	Beaucoup de méthodes existent déjà pour détecter ce défaut. La GTB pourrait être utilisée pour les mettre en oeuvre.

Défaut 2 :	Fuite
Dérive de variable :	Pression et/ou débit de l'eau de chauffage trop bas
Symptôme :	Eau dans le bâtiment. Besoin permanent de l'eau froide (alimentation de chaudière). Corrosion de l'installation.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Défaut 3 :	Embouage
Dérive de variable :	Température intérieure trop basse.
Symptôme :	Souschauffe du bâtiment.
Nécessité de développer des méthodes de détection de défaut :	à discuter
Commentaire :	

Composant : sondes de température.

Défaut 1 :	Dérive de mesure
Dérive de variable :	Valeur mesurée ne correspond pas à la valeur réelle.
Symptôme :	Fausses alarmes. Inconfort. Augmentation de consommation de l'énergie. Régulateur ne fonctionne pas correctement.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	
Défaut 2 :	Sondes immergées pas toujours irriguées
Dérive de variable :	Valeurs mesurées fausses
Symptôme :	Fausses alarmes. Inconfort. Augmentation de consommation de l'énergie. Régulateur ne fonctionne pas correctement.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	
Défaut 3 :	Sonde de la vanne thermostatique influencée par radiateur
Dérive de variable :	Fluctuation de la température intérieure.
Symptôme :	Inconfort des occupants
Nécessité de développer des méthodes de détection de défaut :	à discuter
Commentaire :	
Défaut 4 :	Sonde de température dans la pièce de conduite mal positionnée (apports gratuits trop importants, zone avec peu de débit d'air, pièce non-représentative).
Dérive de variable :	
symptôme :	Durée de relance matinale trop courte, courbe de chauffe pas adaptée au bâtiment, températures intérieures dans certaines pièces trop basses au début de matinée.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	
Défaut 5 :	Sonde de la température d'eau trop près du point de mélange.
Dérive de variable :	
Symptôme :	Mauvaise régulation de la température de départ secondaire.
Nécessité de développer des méthodes de détection de défaut :	NON
Commentaire :	
Défaut 6 :	Perturbation électrique sur le fil de sonde
Dérive de variable :	Instabilité de la valeur mesurée
Symptôme :	Mauvaise régulation
Nécessité de développer des méthodes de détection de défaut :	à discuter
Commentaire :	

Composant n°137 : Régulateur en fonction de la température extérieure.

Défaut 1 :	Courbe de chauffe mal réglée
Dérive de variable :	Température intérieure corrélée avec la température extérieure.
Symptôme :	Inconfort des occupants, vanne thermostatique presque fermée.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	Courbe de chauffe auto-adaptative est proposée pour résoudre ce problème.

Composant : Horloge.

Défaut 1 :	Durée d'occupation mal déterminée par occupant.
Dérive de variable :	Température intérieure ne correspond pas au besoin des occupants.
Symptôme :	Plainte des occupants et/ou gaspillage de l'énergie.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Défaut 2 :	Horloge mal réglé
Dérive de variable :	Température intérieure ne correspond pas au besoin des occupants.
Symptôme :	Plainte des occupants et/ou gaspillage de l'énergie.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Composant n°138 : programmeur d'intermittence (PI).

Défaut 1 :	Commutation trop tard entre régime réduit (ou arrêt) et relance matinale de chauffage
Dérive de variable :	Température intérieure trop basse au début de matinée.
Symptôme :	Plainte des occupants surtout le lundi matin quand il fait très froid
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	PI est proposé pour résoudre le problème.

Défaut 2 :	Commutation trop tôt entre régime réduit (ou arrêt) et relance matinale
Dérive de variable :	
Symptôme :	Relance matinale termine avant l'occupation du bâtiment. gaspillage de l'énergie.
Nécessité de développer de méthodes de détection de défaut :	OUI
Commentaire :	PI est proposé pour résoudre le problème.

Défaut 3 : **Commutation trop tôt entre relance matinale et régime normal de chauffage**

Dérive de variable :	Température intérieure trop basse surtout le lundi matin.
Symptôme :	Plainte des occupants.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Défaut 4 : **Commutation trop tôt entre régime normal et arrêt (ou régime réduit).**

Dérive de variable :	Température intérieure trop basse à la fin d'après-midi.
Symptôme :	Plainte des occupants
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Défaut 5 : **Fausse information de la boucle secondaire transmise au gestionnaire de chaudière en séquence.**

Dérive de variable :	Température à la sortie de chaudière trop haute ou trop basse.
Symptôme :	Température de départ secondaire trop basse. Gaspillage de l'énergie. Inconfort des occupants.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Composant n°139 : Régulateur de la température de départ secondaire.

Défaut 1 : **Pas adapté à la vanne-trois-voies**

Dérive de variable :	Fluctuation de la température de départ secondaire
Symptôme :	Endommagement de la vanne. Fluctuation de la température intérieure.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Défaut 2 : **Ne fonctionne pas**

Dérive de variable :	Température de départ secondaire différente de sa valeur normale.
Symptôme :	Vanne-trois-voies reste ouverte ou fermée ou en position initiale. Inconfort des occupants. Gaspillage de l'énergie.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

Composant n°141 : "gestionnaire" des chaudière en séquence.

Défaut 1 :	Mauvaise choix du nombre des chaudières utilisées
dérive de variable :	Durée de fonctionnement de chaudière pas correcte
Symptôme :	Mauvais rendement de chaudière. Charge moyen du fonctionnement de chaudière trop basse. Marche-arrêt de brûleur trop fréquent.
Nécessité de développer des méthodes de détection de défaut :	OUI
Commentaire :	

ANNEXE B4 - PRESENTATION DU QUESTIONNAIRE D'ENQUETE

MODE D'EMPLOI DU QUESTIONNAIRE D'ENQUETE

Ce questionnaire d'enquête doit nous permettre de recenser les défauts les plus fréquents des installations de chauffage à eau chaude.

Le questionnaire d'enquête comporte une liste de composants d'une installation de chauffage à eau chaude. Pour chacun de ces composants est donnée une liste de défauts possibles.

L'enquête doit permettre pour chacun de ces défauts de déterminer :

- sa fréquence d'apparition (colonnes C1 ou C2 du questionnaire),
- si il est facile aux équipes d'exploitation de détecter ce défaut (colonne C3 du questionnaire d'enquête),
- son incidence sur le confort des utilisateurs (colonne C4 du questionnaire d'enquête),
- son incidence sur les consommations d'énergie (colonne C5 du questionnaire d'enquête).

Pour ce faire nous vous demandons pour chaque défaut d'indiquer un chiffre (0, 1, 2, 3) dans les colonnes C1 ou C2, C3, C4, C5 du questionnaire d'enquête.

Les colonnes C1 et C2 correspondent à la fréquence des défauts. La colonne C2 est utilisée pour les défauts liés à la conception de l'installation, la colonne C1 pour les autres défauts.

Seule une des deux colonnes C1 et C2 est à remplir pour chaque défaut.

Dans la colonne 1 mettez le chiffre :

- 0 si la durée moyenne s'écoulant entre deux apparitions de ce défaut est supérieure à 15 ans,
- 1 si cette durée est comprise entre 7 et 15 ans,
- 2 si cette durée est comprise entre 2 et 7 ans,
- 3 si cette durée est inférieure à 2 ans

Dans la colonne 2 mettez le chiffre :

- 0 si ce défaut de conception est rare c'est à dire si on le trouve sur moins de 1% des installations,
- 1 si ce défaut de conception est peu fréquent c'est à dire si on le trouve sur 1 à 5% des installations,
- 2 si ce défaut de conception est fréquent c'est à dire si on le trouve sur 5 à 20% des installations,
- 3 si ce défaut de conception est très fréquent c'est à dire si on le trouve sur plus de 20% des installations.

La colonne C3 correspond à la difficulté qu'a le personnel de maintenance à détecter le défaut.

Mettez le chiffre :

- 0 si il existe généralement un dispositif automatique permettant de détecter le défaut et de le signaler au personnel de maintenance,
- 1 si le défaut est généralement détecté lors d'une visite de routine du personnel de maintenance,
- 2 si la détection du défaut nécessite une analyse détaillée du fonctionnement de l'installation par le personnel de maintenance,
- 3 si le défaut peut ne pas être détecté lors d'une analyse détaillée du fonctionnement de l'installation par le personnel de maintenance.

La colonne C4 correspond au confort des occupants et à la disponibilité de l'eau chaude sanitaire.

Mettez dans cette colonne le chiffre :

- 0 si l'occupant ne s'aperçoit pas des conséquences du défaut,
- 1 si l'occupant s'aperçoit des conséquences du défaut plus d'une journée après qu'il ait eu lieu,
- 2 si l'occupant s'aperçoit des conséquences du défaut dans la journée où il a lieu,
- 3 si l'occupant s'aperçoit des conséquences du défaut en moins de dix minutes.

La colonne C5 correspond à l'impact du défaut sur la consommation d'énergie.

Mettez dans cette colonne le chiffre :

- 0 si le défaut n'entraîne pas d'augmentation de la consommation d'énergie,
- 1 si la présence du défaut conduit à une consommation correcte mais non optimale,
- 2 si la présence du défaut entraîne un gaspillage d'énergie,
- 3 si la présence du défaut entraîne un gaspillage d'énergie important.

Tableau récapitulatif décrivant les caractéristiques pour évaluer l'importance d'un défaut de composant	
Colonne C1 : <i>Fréquence de défaut</i>	
0	Durée moyenne entre deux défauts supérieure à 15 ans
1	Durée moyenne comprise entre 7 et 15 ans
2	Durée moyenne comprise entre 2 et 7 ans
3	Durée moyenne inférieure à 2 ans
Colonne C2 : <i>Taux de présence de défaut (seulement pour les défauts de conception)</i>	
0	Défauts de conception rares (moins de 1% des installations)
1	Défauts de conception peu fréquents (moins de 5% des installations)
2	Défauts de conception fréquents (de 5 à 20% des installations)
3	Défauts de conception très fréquents (plus de 20% des installations)
Colonne C3 : <i>Détection de défauts</i>	
0	Il existe un système automatique de détection de défaut
1	Le défaut est détecté lors d'une visite de routine du personnel de maintenance
2	La détection de défaut nécessite une analyse détaillée du fonctionnement de l'installation par le personnel de maintenance
3	Le défaut peut ne pas être détecté lors d'une analyse détaillée du fonctionnement de l'installation par le personnel de maintenance
Colonne C4 : <i>Confort des occupants et disponibilité de l'eau chaude sanitaire</i>	
0	L'occupant ne s'aperçoit pas des conséquences
1	L'occupant pourrait s'apercevoir des conséquences plus d'une journée après l'apparition du défaut
2	L'occupant s'apercevra des conséquences dans moins d'une journée
3	L'occupant s'apercevra des conséquences dans moins de dix minutes
Colonne C5 : <i>Consommation d'énergie</i>	
0	La consommation d'énergie n'est pas augmentée
1	L'utilisation de l'énergie n'est pas optimisée
2	Gaspillage de l'argent de l'occupant
3	Gaspillage important de l'argent de l'occupant

QUESTIONNAIRE D'ENQUETE SUR LES DEFAUTS DE FONCTIONNEMENT DES SYSTEMES DE CHAUFFAGE A EAU CHAUDE

Nom de la personne
ayant répondu au questionnaire :

Société :

Adresse :

.....

.....

N° de téléphone :

.....

Fonction :

Nota : Ces coordonnées seront utilisées pour vous envoyer les résultats de l'enquête.

En cas de difficulté pour remplir ce questionnaire contacter :

Monsieur LI
CSTB - Service GEC
BP 02
77421 Marne la vallée Cedex 2
Téléphone : (1) 64 68 83 13

Le questionnaire est à renvoyer à l'adresse ci dessus.

COMPOSANT	DEFAUT	Effet d'un défaut pour une caractéristique choisie				
		C1	C2	C3	C4	C5
Brûleur	* Arrêt de sécurité					
	* Mauvaise combustion (par exemple : combustion incomplète, excès d'air trop important, défaut d'air, ect...)					
	* Panne électrique					
	* Panne mécanique					
	* Sousdimensionnement (C)**					
	* Surdimensionnement (C)					
	* Autres					
Chaudière	* Fuite d'eau vers les circuits des fumées					
	* Dégradation de l'isolation thermique					
	* Corrosion					
	* Encrassement (circuit des fumées)					
	* Entartrage (circuit d'eau)					
	* Autres					
Soupape de sûreté	* Fuite					
	* Embouage					
	* Entartrage					
	* Autres					
Pompe	* Panne électrique					
	* Panne mécanique					
	* Encrassement (entartrage, embouage, ...)					
	* Cavitation					
	* Sousdimensionnement(C)					
	* Surdimensionnement(C)					
	* Autres					

** (C) : Défaut de conception

COMPOSANT	DEFAUT	Effet d'un défaut pour une caractéristique choisie				
		C1	C2	C3	C4	C5
Alimentation de gaz	* Arrêt anormal					
	* Fuite de gaz					
	* Perturbation importante de la pression de gaz					
	* Autres					
Manchon vibratile (à l'entrée et à la sortie de pompe)	* Vibration					
	* Eclatement					
	* Autres					
Vanne 3 voies, vanne 2 voies et Servo-moteur	* Fuite de la voie directe vers le bypass					
	* Fuite du bypass vers la voie directe					
	* Vanne non étanche en position fermée					
	* Sifflement					
	* Embouage					
	* Entartrage					
	* Panne électrique					
	* Panne mécanique					
	* Limiteur de course mal positionné					
	* Mauvaise autorité hydraulique (C)					
	* Autres					
Préparation d'ECS	* Fuite d'eau sur le circuit d'ECS					
	* Fuite d'eau entre circuit primaire et circuit secondaire					
	* Entartrage					
	* Corrosion					
	* Sousdimensionnement (C)					
	* Surdimensionnement (C)					
	* Autres					

COMPOSANT	DEFAUT	Effet d'un défaut pour une caractéristique choisie				
		C1	C2	C3	C4	C5
Radiateur	* Air dans radiateur					
	* Embouage					
	* Entartrage					
	* Sousdimensionnement général (C)					
	* Surdimensionnement général (C)					
	* Sousdimensionnement et surdimensionnement dans certaines pièces ou dans des zones particulières (C)					
	* Autres					
Vanne thermostatique	* Panne de la vanne thermostatique					
	* Mauvaise autorité de la vanne thermostatique(C)					
	* Autres					
Disconnecteur et circuit d'eau froide	* Défaillance complète					
	* Fuite vers les réseaux de l'eau de ville					
	* Fuite d'eau sur le circuit					
	* Autres					
Système d'expansion	* Fuite ou manque du gaz de remplissage					
	* Gel					
	* Autres					
Conduit des fumées	* Corrosion					
	* Feu dans le conduit					
	* Autres					
Bâtiment	* Dégradation de l'isolation thermique du bâtiment					
	* Isolation thermique ne correspond pas aux calculs (C)					
	* Autres					
Occupant	* Ouverture "inadéquate" de fenêtre					
	* Usage incorrect de thermostat					
	* Autres					

COMPOSANT	DEFAUT	Effet d'un défaut pour une caractéristique choisie				
		C1	C2	C3	C4	C5
L'ensemble de l'installation	* Mauvais équilibre hydraulique entre les circuits (C)					
	* Déséquilibre de chauffage entre différentes zones (C)					
	* Dégradation de l'isolation thermique de l'installation					
	* Isolation thermique ne correspond pas aux calculs (C)					
	* Fuite légère d'eau dans la chaufferie					
	* Fuite importante d'eau dans la chaufferie					
	* Mauvais assemblage entre tuyaux et joints (C)					
	* Air dans des tuyaux					
	* Embouage					
	* Entartrage					
	* Corrosion générale					
	* Gel de tuyaux					
	* Autres					
Sondes de température	* Sonde de la vanne thermostatique influencée par le radiateur ou d'autres sources de chaleur (rayonnement solaire, ...) (C)					
	* Sondes de température extérieure mal positionnées (C)					
	* Sondes de la température d'eau (circuits de départ) trop près du point de mélange (C)					
	* Sonde de mesure de la température intérieure placée dans une pièce non représentative (C)					
	* Panne électrique des sondes					
	* Panne mécanique des sondes					
	* Perturbation électrique sur le fil des sondes					
	* Autres					
Régulateur en fonction de la température extérieure	* Courbe de chauffe mal réglée					
	* Pas adapté à la vanne-3-voies (C)					
	* Panne électrique					
	* Panne mécanique					
	* Autres					

COMPOSANT	DEFAUT	Effet d'un défaut pour une caractéristique choisie				
		C1	C2	C3	C4	C5
<i>Limiteur de température de retour</i>	* Mauvaise loi de régulation					
	* Consigne de température de retour mal réglée					
	* Panne électrique					
	* Panne mécanique					
	* Autres					
<i>Programmateurs d'intermittence</i>	* Mauvais calcul de l'heure de relance par l'optimiseur					
	* Mauvais calcul de l'heure d'arrêt par l'optimiseur					
	* Mauvais réglage					
	* Panne électrique					
	* Panne mécanique					
	* Autres					
<i>Horloge</i>	* Horloge pas à l'heure					
	* Des horaires d'occupation non cohérents avec l'occupation réelle					
	* Panne électrique					
	* Panne mécanique					
	* Autres					
<i>"Gestionnaire" des chaudières en cascade</i>	* Mauvais choix du nombre des chaudières utilisées					
	* Mauvais réglage					
	* Panne électrique					
	* Panne mécanique					
	* Autres					
<i>Régulateurs des aquastats des chaudières</i>	* Mauvaise séquence d'arrêt entre pompe et brûleur					
	* Mauvais réglage					
	* Panne électrique					
	* Panne mécanique					
	* Autres					

BIBLIOGRAPHIE

- [**Anderson, 1989**] Anderson, D. & al. *A quasi real time expert system for commercial building HVAC diagnostics*. ASHRAE Trans., V.95, Pt.2, 1989.
- [**Azcarraga, 1991**] Azcarraga, A. & A. Giacometti. 1991. *A prototype-bases incremental network model for classification tasks*. Actes Neuro-Nîmes 91, pp 121-134.
- [**Barschdorff, 1991**] Barschdorff, D & A. Both. 1991. *Adaptive Nearest Neighbour Neural Network for Pattern Recognition*. Actes Neuro-Nîmes 91, pp 241-251.
- [**Brunet, 1990**] Brunet, J., D.Jaume, M.Labarrère, A.Rault & M.Vergé. 1990. *Détection et diagnostic de pannes*. Editions Hermès.
- [**Caccavelli, 1991**] Caccavelli, D., E. Hutter, S. Nibel & P. Picard. *Gas Fired Heating Plant Management*. Proc. Building Simulation'91-IBPSA, pp169-174, August 1991, Nice-France.
- [**Carrier, 1992**] Carrier Corporation. 1992. *Product specification, Model#48DJ*, Syracuse, NY, USA.
- [**Chow, 1957**] Chow, C.K. 1957. *An optimum character recognition system using decision functions*. R.E. Trans. on Information Theory, Vol. 13, pp 21-27.
- [**Culp, 1989**] Culp, C.H. *Expert system in preventive maintenance and diagnostics*. ASHRAE Journal, p24-27, August, 1989.
- [**Denker, 1987**] Denker, J., D. Schwartz, B. Wittner, S. Solla, R. Howard, L. Jackel & J. Hopfield. 1987. *Large automatic learning, rule extraction and generalization*. Comptis Systèmes 1, pp. 877-922.
- [**Demuth, 1994**] Demuth, H. & M. Beale. 1994. *Neural Network Toolbox - For use with MATLAB*. The MATH WORKS Inc..
- [**Devijver, 1982**] Devijver, P.A. & J. Kittler. 1982. *Pattern recognition, a statistical approach*. Prentice-Hall, Englewood Cliffs.
- [**Diday, 1980**] Diday, E. & al. 1980. *Optimisation en classification automatique*. INRIA.
- [**Dubuisson, 1990 a**] Dubuisson, B. 1990. *Diagnostic et reconnaissance des formes*. Editions Hermès.
- [**Dubuisson, 1990 b**] Dubuisson, B. 1990. *Decision with reject option*. Proceedings Eusipco.
- [**Duda, 1973**] Duda, R.O. & P.E. Hart. 1973. *Pattern classification and scene analysis*. Wiley-Interscience Publication.

- [**Fabron, 1995**] Fabron, C. 1995. *Régulation de systèmes thermiques en régime dynamique - Application au chauffage intermittent des bâtiments*. Thèse de doctorat de l'Ecole des Mines de Paris.
- [**François, 1990**] François, C., E.Hutter & J.C.Visier. *Etude sur les formalismes de modèles simplifiés de bâtiment*. CSTB-GEC n°89-4971, Mars 1990.
- [**Frank, 1987**] Frank, P.M. 1987. « *Fault diagnosis in dynamic systems via state estimation - a survey* », S. Tzafestas (eds.), « *System fault diagnostics, reliability, and related knowledge - based approaches* », Vol. 1, 35-98. D.Reidel Publishing Co.
- [**Fukunaga, 1972**] Fukunaga, K. 1989. *Introduction to statistical pattern recognition*. Academic Press.
- [**Hebb, 1949**] Hebb, D.O. 1949. *The Organization of Behavior*. NY, Wiley-Editions.
- [**Hertz, 1991**] Hertz, J., A. Krogh & R.G. Palmer. 1991. *Introduction to the theory of neural computation*. Addison-Wesley Publishing Company.
- [**Hyvärinen, 1993**] Hyvärinen, J. & al. 1993. *Building Optimisation and Fault Diagnosis (BOFD) system concept*. Collective document - IEA Annex 24, VTT, Finland.
- [**Isermann, 1983**] Isermann, R. 1983. *Process fault detection based on modelling and estimation methods - a survey*. Automatica, Vol. 20, 387-404.
- [**Jacobs, 1988**] Jacobs, R.A. 1988. *Increased rates of convergence through learning rate adaptation*. Neural networks, 1, pp 295-307.
- [**Laret, 1980**] Laret, L. 1980. *Contribution au développement de modèles mathématiques du comportement thermique transitoire de structures d'habitation*. Thèse de doctorat de l'Université de Liège, Belgique.
- [**Le Cun, 1985**] Le Cun, Y. 1985. *Une procédure d'apprentissage pour réseau à seuil asymétrique*. Proc. of COGNITIVA 85, Cesta-Afcet, pp 599-604.
- [**Le Cun, 1987**] Le Cun, Y. 1987. *Modèles connexionnistes de l'apprentissage*. Thèse de Doctorat de l'Université de Paris 6.
- [**Le Cun, 1989**] Le Cun, Y. 1989. *Generalization and Network Design Strategies*. R. Pfeifer, Z. Schreter, F. Fogelman-Soulié & L. Steels (Eds.), Elsevier Science Publishers B.V., North-Holland.
- [**Li, 1995a**] Li, X., H. Vaézi-Néjad & J.C. Visier. *Application de réseaux neuronaux à la détection des défauts d'installations de génie climatique*. Proc. 2ème conférence internationale sur l'automatisation industrielle, Vol.1, pp179-184, Juin 1995, Nancy, France.
- [**Li, 1995b**] Li, X., H. Vaézi-Néjad & J.C. Visier. *Development of a fault diagnosis method for heating system by using neural networks*. Proc. Tsinghua-HVAC-'95, September 1995, Beijing, China.
- [**Li, 1995c**] Li, X. & J.C. Visier. *Méthodes automatiques de détection des défauts et dérives des installations de chauffage - Les défauts de fonctionnement des installations de chauffage collectif*. Cahiers du CSTB, n° 365, Décembre 1995, France.
- [**Li, 1996**] Li, X., H. Vaézi-Néjad & J.C. Visier. *Development of a fault diagnosis method for heating systems using neural networks*. ASHRAE Trans., Winter meeting, February 1996, Atlanta, USA.

[Masson, 1992] Masson, M.H. *Contribution à l'élaboration d'une méthode de décision avec rejet par réseaux de neurones - Application à la surveillance de système*. Thèse de Doctorat de l'Université de Technologie de Compiègne.

[Matlab,1992] Matlab. 1992. *Matlab - High-Performance Numeric Computation and Visualisation Software - User's Guide and Reference Guide*. The MATH WORKS Inc.

[Mc Culloch, 1943] Mac Culloch, W.S. & W. Pitts. 1943. *A logical calculus of the ideas immanent in nervous activity*. Bulletin of Math. Biophysics, Vol. 5, pp 115-143.

[Minsky, 1969] Minsky, M. & S. Papert. 1969. *Perceptrons : an introduction to computational geometry*. MIT Press, Cambridge, Massachusetts.

[Morant,1982] Morant, M.A., J.M.DOLS & J.Hannay. 1982. *Water Boiler model*. Technical Form, IEA - Annex 10, Laboratoire de Physique du Bâtiment - Université de Liège, Belgique.

[Ozog, 1985] Ozog, H. 1985. *Hazard Identification Analysis and Control*. Chemical engineering, Feb. 18, pp 161-170.

[Parker, 1985] Parker, D.B. 1985. *Learning logic*. Tech Report TR-47, Center for Computational Research in Economics and Management Science, MIT, Cambridge, Massachusetts.

[Paternelle, 1987] Paternelle, S. & P. Caravelle. 1987. *Learning networks of neurons with boolean logic*. Europhysics Letters 4, pp 503-508.

[Pau,1975] Pau, L.F. 1975. *Diagnostic des pannes dans les systèmes*. Cepadues, Toulouse.

[Pau,1981] Pau, L.F. 1981. *Failure diagnosis and performance monitoring*. Marcel Dekker Inc., New York and Basel.

[Ragot, 1990] Ragot, J, M.Darouach, D.Maquin & G.Bloch. 1990. *Validation de données et diagnostic*. Editions Hermès.

[Reid, 1977] Reid, R.C., J.M. Prausnitz & T.K. Sherwood. 1977. *The Properties of Gases and Liquids*. Graw-Hill Book Company.

[Rosenblatt, 1958] Rosenblatt, F. 1962. *Principales of Neurodynamics : Perceptrons and the theory of brain mechanisms*. Spartan books, Washington D C.

[Rumelhart, 1986] Rumelhart, D., G. Hinton & R. Williams. 1986. *Learning internal representations by error propagation in Parallel distributed processing : Exploring the micro-structure of cognition*. MIT Press, Cambridge, Massachusetts.

[Sacadura, 1978] Sacadura, J.F. (Coordonnateur). 1978. *Initiation aux transferts thermiques*. Editions Technique et Documentation, Paris.

[Simon, 1984] Simon, J.C. 1984. *La reconnaissance de formes par algorithmes*. Collection Etudes et Recherches en informatique, Masson.

[Simulink] Simulink. 1992. *Simulink - A Program for Simulating Dynamic Systems*. The MATH WORKS Inc.

[Spriet, 1982] Spriet, J.A. & G.C. Vansteenkiste. 1982. *Computer-aided modelling and simulation*. Academic Press.

[Tou, 1974] Tou, J.T. & R.C. Gonzales. 1974. *Pattern recognition principles*. Addison-Wesley Publication, Reading, Massachusetts.

[Vaézi-Néjad, 1990] Vaézi-Néjad, H. 1990. *Détection de défauts d'instruments de mesure*. Thèse de Doctorat de l'Université de Nancy 1.

[Villemeur, 1988] Villemeur, A. 1988. *Sûreté de fonctionnement des systèmes industriels*. Eyrolles.

[Visier, 1994] Visier, J.C., H. Vaézi-Néjad & M. Jandon. *Results of the application of a quality assessment procedure to different building energy management systems*. Proc. The 3th IEEE Conference on Control Applications, Vol.3, pp 1867-1872, August 1994, University of Strathclyde, Glasgow-Scotland-UK.

[Wagner,1988] Wagner, J. & R.Shoureshi. *Observer designs for diagnostics of nonlinear processes and systems*. Proc. ASME Winter meeting, November 1988, Chicago-USA.

[Widrow, 1960] Widrow, B & M.E. Hoff. 1960. *Adaptive Switching Circuits*. Ire Wescon. Conv Record, Part 4, pp 96-104.