

LES RESEAUX DE NEURONES ARTIFICIELS POUR L'AMELIORATION DE LA QUALITE

Rafik BOUASSIDA et Mohamed M. T. LIMAM

Institut Supérieur de Gestion de Tunis
41, Rue de de la liberté, Le Bardo 2000, Tunis, Tunisie

Résumé

Le travail présenté dans cet article s'intéresse à l'utilisation des réseaux de neurones artificiels dans le contrôle de la qualité. Les réseaux de neurones représentent des alternatives aux outils de maîtrise statistique des procédés, plus particulièrement les cartes de contrôle. Des comparaisons basées sur le critère d'erreur de diagnostic ont été effectuées entre des réseaux multicouches et les cartes univariées X-bar, S, R et CUSUM. Ces comparaisons ont été généralisées au cas des cartes bivariées. Les résultats obtenus montrent que le perceptron multicouche est meilleur que les cartes de contrôle univariées dans la détection des variations relatives à une caractéristique donnée. En outre, le perceptron multicouche s'est montré plus performant que les cartes de contrôle bivariées χ^2 puisqu'il détecte aussi bien le dérèglement que la variable qui en est responsable.

Mots clés: Cartes de contrôle, Cartes CUSUM, Maîtrise statistique des procédés, Perceptron multicouche.

1. Introduction

L'importance de la qualité pour assurer la compétitivité, la performance et l'excellence des entreprises n'est plus à démontrer. L'amélioration continue de la qualité en dehors du cycle de production "hors-ligne" est souvent assurée par les plans d'expériences et la régression multiple. Le contrôle des procédés en cours de fabrication est souvent assuré par les outils statistiques de maîtrise des procédés, plus particulièrement les cartes de contrôle. Certes, ces outils ont démontré leur utilité, mais ils présentent certaines insuffisances tant au niveau des conditions d'application qu'au niveau de l'utilisation, comme mentionné par Smith (1994), Yazici et Smith (1992). Les cartes de contrôle sont utilisées pour analyser les paramètres d'un procédé et déterminer son état de contrôle et ce dans le but de différencier les causes communes des causes spéciales (Montgomery (1991)). Ces cartes se basent sur la normalité des données, et sont sensibles aux données manquantes ou bruitées.

Face à ces outils statistiques, se trouve la méthodologie de réseaux de neurones artificiels (RNA). Les RNA ont connu une expansion très rapide. Leur potentiel pour les tâches de reconnaissance de formes dans des domaines comme le traitement de la parole ou de l'image a été mis en évidence depuis une décennie. De même, la capacité de ces réseaux connexionnistes de prendre en compte la nature temporelle de l'information a ouvert de nouveaux champs d'application comme la prévision, la robotique et le contrôle de la qualité. Le principal avantage de l'approche neuronale est son habilité à manipuler des données bruitées et ayant des distributions de probabilité inconnues, ainsi que son aptitude à modéliser simultanément plusieurs stratégies de contrôle telles que les cartes \bar{X} , R, S, CUSUM et χ^2 par un simple réseau. En plus, son intérêt s'explique par l'augmentation des capacités de calcul en volume et en vitesse, la capacité d'adaptation et la possibilité de modéliser des problèmes présentant des relations non linéaires et complexes.

La section 2 définit les cartes de contrôle et présente celles utilisées dans cet article. Dans la section 3 sera présentée l'interprétation de la carte de contrôle Shewhart de la moyenne d'un procédé (\bar{X}) et celle de la dispersion (carte R) par les réseaux de neurones. Cette étude reprend certains des paramètres utilisés par Smith (1994) pour comparer les RNA et la carte (\bar{X} , R). Cependant, vu que l'étude effectuée n'a pris en considération qu'une seule efficacité de carte de contrôle correspondant à un seul effectif d'échantillons, une autre étude similaire tenant compte de trois efficacités différentes de cartes de contrôle sera effectuée. Elle permettra de comparer d'une part les modèles neuronaux aux cartes (\bar{X} , R) et d'autre part ces mêmes modèles aux cartes (\bar{X} , S). La section 4 présente l'utilisation des RNA pour le diagnostic d'un changement de structure (moyenne ou variance) dans les cartes CUSUM de la moyenne et celles de la variance (R et S) et ce pour deux tailles différentes d'échantillons. La section 5 introduira la possibilité d'interprétation de la carte bivariée de la moyenne par les RNA.

2. Les cartes de contrôle

Lorsqu'on est en présence d'un processus de fabrication stable et dans le but de prévenir l'apparition des défauts, on utilise les cartes de contrôle (CC). Les CC constituent à la fois un outil statistique et graphique, qui permet de suivre dans le temps l'évolution des valeurs individuelles d'une variable (caractéristique d'un produit ou paramètre d'un procédé), ou de la moyenne d'un échantillon prélevé régulièrement, et de décider quand intervenir pour modifier ou stopper une évolution non désirée.

L'aspect préventif des cartes de contrôle consiste à anticiper une tendance à la dérive, et à corriger en conséquence le processus à temps avant que des défauts ne soient produits. Evidemment, les CC ne doivent être introduites que si toutes les causes spéciales de variation ont été éliminées, sans quoi toute prédiction du comportement du processus serait impossible.

Le pilotage du processus pour le contrôle en cours de production augmente la rapidité de correction d'un dérèglement et contribue à diminuer la quantité de produits non conformes fabriqués, il est donc un facteur d'amélioration de la qualité et de la productivité.

Les cartes de contrôle de Shewhart se trouvent parmi les CC les plus utilisées. Ces cartes ont été développées par Shewhart en 1931 (Montgomery (1991)) et visent à définir le niveau de la qualité à atteindre et auquel la direction doit se référer, à apporter une aide pour obtenir ce standard, et à fournir un élément de jugement pour décider si ce standard a été atteint.

Pour contrôler un processus, il convient tout d'abord de choisir une ou plusieurs caractéristiques représentant la qualité du produit à contrôler. Pour chacune des caractéristiques retenues, un échantillon doit être constitué périodiquement dans des conditions fixées à l'avance. Les résultats obtenus sur ces échantillons sont résumés par une ou plusieurs valeurs appelées "statistiques d'échantillon", pouvant être par exemple la moyenne, l'écart type ou l'étendue. Ces statistiques peuvent alors être portées sur une carte. Un point sur cette carte représente donc l'état du processus à un moment donné.

Chaque carte comporte une ligne centrale (LC), une limite inférieure de contrôle notée LIC, une limite supérieure de contrôle notée LSC et parfois des limites de surveillance. Les limites de contrôle jouent le rôle d'alarme. En effet, lorsqu'un point se situe à l'intérieur des limites de contrôle, le processus est dit sous contrôle. Par contre, lorsque celui-ci se situe à l'extérieur, alors il est probable que le processus est hors contrôle (figure 1) Dans ce cas, une action corrective doit être entreprise, sinon, les pièces produites ultérieurement auront probablement des défauts.

Les limites de contrôle des cartes de Shewhart sont calculées de telle sorte que la probabilité qu'un point tombe en dehors des limites alors que le processus est stable, reste très faible. La carte de Shewhart habituellement utilisée est la carte dont les limites sont à une distance égale à plus ou moins 3σ de la moyenne de la statistique utilisée, ce qui correspond à une probabilité de 0.0027 (0.00135 sur chaque coté) d'être en dehors des limites lorsque le processus est stable et ce sous

l'hypothèse que la variable observée suit une loi Gaussienne. Les limites de surveillance permettent d'attirer l'attention sur un comportement qui n'est pas encore hors contrôle mais qui nécessite une attention accrue et une surveillance particulière.

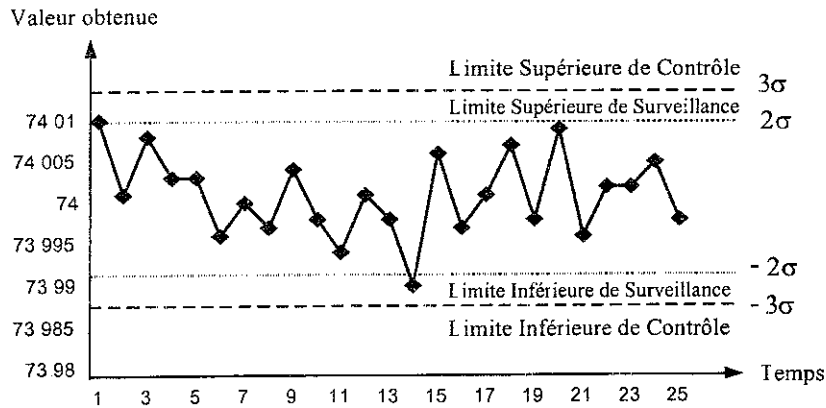


Figure 1: Carte de contrôle typique

Les conclusions que l'on peut tirer de la position d'un point par rapport aux limites de contrôle sont soumises aux incertitudes inhérentes aux fluctuations d'échantillonnage. Deux sortes d'erreur sont alors possibles (dont les risques s'expriment par des probabilités):

* Une erreur dite de risque α consiste à conclure qu'un dérèglement a eu lieu alors qu'aucune cause n'est venue perturber le processus. Ce cas est également désigné sous le nom de "fausse alarme".

* Une autre dite de risque β consiste à conclure que le processus est resté stable alors qu'une cause perturbatrice est venue effectivement le dérégler.

Les cartes de contrôle se différencient non seulement par les méthodes de calcul des limites et les statistiques utilisées, mais aussi par la rapidité avec laquelle un dérèglement est détecté. Cette rapidité se mesure en terme de Période Opérationnelle Moyenne (POM) "Average Run Length" qui est le nombre moyen d'échantillons successifs nécessaires pour s'apercevoir du dérèglement ou autrement dit, le temps moyen d'attente entre le moment où le dérèglement se produit et l'instant où la carte le signale. Cette rapidité dépend évidemment du dérèglement et elle est notée $POM(\delta)$ où δ est le dérèglement maximal toléré de la statistique étudiée, exprimé en nombre d'écart types.

Ainsi une carte de contrôle est dite efficace pour la détection d'un dérèglement δ si $POM(\delta)$ est faible. Cependant, il arrive que la carte signale une sortie des limites sans qu'il ait eu un dérèglement, la $POM(0)$ correspondante est alors le nombre moyen de contrôle avant une fausse alerte qui conduit à des efforts inutiles pour la recherche de causes spéciales inexistantes.

L'efficacité d'une carte de contrôle se mesure avec les deux POM; une bonne carte doit réaliser un compromis entre la période opérationnelle moyenne en cas de dérèglement $POM(\delta)$ et la

POM(0). Souvent, on souhaite détecter un dérèglement δ le plus vite possible (la POM correspondant à δ doit être faible) en ayant le moins de fausses alertes possible (la POM(0) doit être grande). De même, cette efficacité peut être exprimée par la Probabilité d'acceptation (Pa) de conclure à la stabilité d'un processus en fonction du dérèglement de celui-ci; c'est-à-dire la probabilité de se trouver entre les limites de contrôle ou la probabilité de ne pas déceler un dérèglement (risque β).

Les cartes utilisées dans cet article sont dites aux mesures. Elles s'attachent aux variables quantitatives (mesurables), elles réunissent le contrôle du paramètre de centrage et celui de la dispersion (étalement des valeurs de la distribution de la variable autour d'une valeur centrale), et comprennent les cartes sur la moyenne \bar{X} , sur l'étendue R ou sur l'écart type S.

2.2.1 Les cartes de contrôles de l'étendue R

La carte de contrôle de l'étendue R (Range), appelée encore carte W, est utilisée lorsqu'on est en présence de mesures et de valeurs obtenues à partir des observations faites sur un processus. La collecte de ces mesures doit être suffisamment rapide pour permettre la formation de sous-groupes. Les sous-groupes utilisés par ce type de carte sont appelés échantillons. Chaque échantillon est constitué d'un certain nombre de prélèvements individuels qui doivent être l'image la plus représentative possible de la production. La carte R est une carte aux variables utilisée pour traduire la dispersion de l'échantillon prélevé, et contrôler la variabilité d'un processus. Elle est souvent recommandée pour évaluer l'état de contrôle avant qu'une carte X-bar ne soit construite. L'étendue R est calculée à partir de deux observations dans chaque sous-groupe (différence entre la plus grande et la plus petite valeur)

Pour un contrôle par rapport à des valeurs standards connues μ_0, σ_0 , la carte R est définie par: $LSC = D_2\sigma_0$, $LC = d_2\sigma_0$, $LIC = D_1\sigma_0$. Sinon, les limites de contrôle inférieures et supérieures d'une carte R doivent être trouvées à partir de: $\bar{R} \pm 3\hat{\sigma}_R$, où \bar{R} est la moyenne de toutes les étendues des échantillons et $\hat{\sigma}_R$ est l'estimation de l'écart type de R.

2.2.2 Les cartes de contrôles de l'écart type S

La carte de contrôle de Shewhart S est une autre alternative pour le contrôle de la variabilité d'un processus. Comme la carte R, elle se base sur les échantillons formés à partir des observations sur une variable responsable de la qualité d'un produit ou contrôlant un procédé.

Ce type de carte est généralement préféré aux cartes d'étendue parce que contrairement à la carte R où l'étendue est calculée pour chaque échantillon à partir de deux observations seulement, la carte S utilise toutes les observations d'un échantillon pour le calcul de l'écart type S. En outre, la perte d'information engendrée par le calcul de R est négligeable si la taille de l'échantillon est faible (4 ou 5), mais elle est plus importante lorsque la taille de l'échantillon devient grande. De même, plusieurs méthodes statistiques utiles pour l'amélioration de la qualité sont basées sur S ou S² au

lieu de R.

Dans le cas où le contrôle se fait par rapport à des valeurs standards connues μ_0 , σ_0 , la carte S est définie par: $LSC = B_6\sigma_0$, $LC = c_4\sigma_0$, $LIC = B_5\sigma_0$. Si ces valeurs sont inconnues, les limites de contrôle de cette carte sont calculées à partir de: $\bar{S} \pm 3\hat{\sigma}_s$; où \bar{S} est la moyenne de tous les écarts types des échantillons et $\hat{\sigma}_s$ est l'estimation de l'écart type de S.

2.2.3 Les cartes de contrôles \bar{X} (X-bar)

La carte de contrôle de la moyenne X-bar est fréquemment utilisée dans le contrôle statistique de la qualité. Son rôle est d'assurer le suivi d'une caractéristique mesurable représentant le processus à maîtriser et de traduire uniquement la tendance centrale de l'échantillon.

L'utilisation de la carte X-bar pour déterminer si la moyenne du processus est sous contrôle peut logiquement se faire tant que la variabilité du processus est sous contrôle statistique. Pour un contrôle par rapport à des valeurs μ_0 , σ_0 et dans le cas où on a des données groupées, et où on peut disposer à chaque instant du contrôle d'un échantillon ($X_i, i=1, \dots, n$) de n observations d'une loi normale $N(\mu_0, \sigma_0^2)$, la carte X-bar est définie par:

$$LSC = \mu_0 + \frac{3\sigma_0}{\sqrt{n}}, \quad LC = \mu_0, \quad LIC = \mu_0 - \frac{3\sigma_0}{\sqrt{n}}.$$

Lors de la mise en place de la carte, les valeurs μ_0 et σ_0 sont souvent inconnues. La méthode habituelle consiste à remplacer μ_0 et σ_0 par des estimateurs sans biais. Pour cela, on doit prélever un ensemble de k sous-groupes (k échantillons de taille n avec k assez grand, environ 20). La moyenne des moyennes des échantillons notée $\bar{\bar{X}}$ est un estimateur sans biais de μ_0 ce qui n'est pas le cas pour \bar{S} (ni d'ailleurs pour \bar{R}).

Les limites de contrôle d'une carte X-bar sont données à partir de : $\bar{\bar{X}} \pm 3\hat{\sigma}_{\bar{X}}$; $\hat{\sigma}_{\bar{X}}$ est l'estimation de l'écart type de \bar{X} .

Etant donné que chaque carte de contrôle présente certains avantages et détecte certains types de déviations, et dans le but de tirer profit de ces avantages, plusieurs combinaisons de cartes ont été proposées. Les combinaisons les plus usuelles sont (X-bar, R) et (X-bar, S).

Les cartes traditionnelles aux mesures de Shewhart (X-bar, R et S) utilisent à chaque instant t les seules mesures relatives à t, elle détectent uniquement les dérives rapides et les grands écarts de variation des processus. Par conséquent, elles ne sont pas satisfaisantes lorsque les changements dans l'évolution des caractéristiques représentatives d'un processus sont de faible amplitude. Pour la détection de faibles écarts de dérèglement ainsi que les lentes dérives, on utilise des cartes d'applications spécifiques utilisant les observations à l'instant t et l'ensemble ou une partie des observations à des instants antérieurs à t, telles que les cartes à sommes cumulées (CUSUM).

2.2.4 Les cartes de contrôle de la moyenne CUSUM (CUMulative SUM)

D'après Ryan (1989), la carte de contrôle à sommes cumulées CUSUM a été proposée et développée par Page en 1954 et depuis plusieurs modifications ont été faites sur ces cartes.

La carte CUSUM prend en compte les prélèvements antérieurs des échantillons pour évaluer la performance actuelle du processus. La valeur d'un échantillon i portée sur la carte CUSUM est donnée par la formule suivante: $S_i = \sum_{j=1}^i (\text{moy}(x_j) - \mu_0)$ avec

μ_0 : valeur cible; $\text{moy}(x_j)$: moyenne de l'échantillon j .

S_i est la somme cumulée des différences observées entre les valeurs mesurées et la valeur cible. Ainsi, au lieu de placer sur la carte les moyennes successives des échantillons, comme c'est le cas pour les cartes de Shewhart, on suit une somme d'écartés cumulés entre la moyenne présente et une valeur de référence

La version algorithmique de la méthode CUSUM proposée par Lucas en 1982 est basée sur une paire de sommes cumulatives récursives:

$$S_{Hi} = \max \left[0, (z_i - k) + S_{H(i-1)} \right]$$

$$\text{et } S_{Li} = \max \left[0, (-z_i - k) + S_{L(i-1)} \right] \text{ avec } z_i = (\bar{x}_i - \mu_0) / \sigma_{\bar{x}}, S_{H0} = S_{L0} = 0.$$

La première détecte les changements positifs de la moyenne, la deuxième détecte les changements négatifs de celle-ci. Les valeurs de S_{Hi} et S_{Li} sont portées sur une seule carte ou sur deux cartes différentes. Les limites supérieure et inférieure de contrôle sont situées à $\pm h$. Le processus suivi par ce type de carte est considéré hors contrôle si $S_{Hi} > h$ ou $S_{Li} > h$. Le paramètre k est appelé coefficient de sensibilité généralement choisi égal à $\delta/2$ où δ est le dérèglement que l'on souhaite détecter. h est appelé intervalle de décision, en général pris égal à 4 ou à 5. Les valeurs de k et h sont déterminées sur la base des périodes opérationnelles moyennes correspondant aux dérèglages que l'on souhaite détecter et du nombre de fausses alertes

2.2.5 Les cartes de contrôle multivariées de Shewhart

Lorsqu'on est en présence de p ($p > 1$) caractéristiques à contrôler simultanément, la pratique habituelle consiste à utiliser une carte de contrôle unidimensionnelle pour chacune des caractéristiques. Cependant, cette méthode a l'inconvénient de ne pas tenir compte des éventuelles corrélations entre les variables représentant ces caractéristiques. Pour remédier à ces insuffisances, diverses cartes de contrôle multivariées de type Shewhart et autre (CUSUM et EWMA) ont été mises au point pour prendre en considération les relations entre les caractéristiques. L'avantage de la carte multivariée de Shewhart pour le contrôle de la moyenne par rapport à celle univariée a été illustré dans Ryan (1989).

Les caractéristiques d'un procédé sont des variables aléatoires supposées distribuées selon une loi normale multidimensionnelle. L'hypothèse de base d'une carte multivariée est que chaque valeur X est une observation d'un vecteur gaussien $X = (X^1, X^2, \dots, X^p)$ de loi multinormale $N_p(\mu, \Sigma)$ où μ est un vecteur de R^p et Σ une matrice $p \times p$.

Lorsque le contrôle s'effectue par rapport à des valeurs standards μ_0 et Σ_0 , alors l'objectif est de maintenir la moyenne du procédé à μ_0 . Etant donné qu'une carte de contrôle correspond statistiquement à un test, alors contrôler la moyenne revient à effectuer un test multivarié: $\mu = \mu_0$ contre $\mu \neq \mu_0$. La carte de Shewhart correspondant à ce test consiste à suivre en fonction des échantillons une distance de Mahalanobis notée $D^2(\bar{X}_i; \mu_0)$ entre la valeur μ_0 et le vecteur moyenne pour le i ème échantillon \bar{X}_i (\bar{X}_i est une estimation de la moyenne du procédé), sa formule est la suivante:

$$\chi^2 = D^2(\bar{X}_i; \mu_0) = n(\bar{X}_i - \mu_0)' \Sigma_0^{-1} (\bar{X}_i - \mu_0) \text{ avec } n: \text{ taille d'un échantillon.}$$

La distance calculée sera comparée aux limites supérieure et inférieure de contrôle de la carte de Shewhart de la moyenne définies dans Niang (1994) par : $LSC = \chi_p^2 \alpha$; $LIC = 0$. Si la valeur χ^2 d'un échantillon dépasse la limite supérieure de contrôle, alors on peut conclure que la moyenne multivariée du processus n'est plus à μ_0 et l'échantillon correspondant doit être étudié.

Cette carte est encore appelée carte du χ^2 . Elle représente graphiquement en fonction du temps le test précédent répété pour les échantillons successivement prélevés.

3. Interprétation des cartes (X-bar, R) et (X-bar, S) par les RNA

L'approche d'utilisation des RNA dans l'interprétation des cartes (X-bar, R) a été proposée par Smith (1994). Son principe consiste à utiliser un perceptron multicouche, avec la rétropropagation du gradient comme algorithme d'apprentissage, pour chercher et détecter d'une part toute déviation du processus et déterminer d'autre part si la variation se trouve au niveau de la moyenne ou au niveau de la variance.

Le modèle neuronal utilisé est constitué d'une couche d'entrée composée de $n+3$ nœuds et d'une couche de sortie composée d'un nœud unique responsable de l'état de contrôle du processus (n étant la taille d'un échantillon). Les n premiers nœuds de la couche d'entrée constituent les n prélèvements sur la variable contrôlant le processus en question, les trois restants représentent les statistiques des n observations à savoir leur moyenne \bar{X} , leur étendue R et leur écart type S (voir figure 2). Ces statistiques ont été prises en considération par Smith (1994) parce que sans elles, le réseau n'a pas donné des résultats satisfaisants (pour $n=10$). Dans cette architecture, et dans le cas où le nœud de sortie est de type réel, ce dernier aura la valeur 0 1 si le processus est sous contrôle, 0.5 si le processus présente une variation au niveau de la moyenne et 0.9 s'il y a une variation au niveau de la variance. L'architecture utilisée pour le réseau multicouche peut être remplacée par une

autre où la couche de sortie reste inchangée et la couche d'entrée est constituée uniquement des statistiques \bar{X} , R et S.

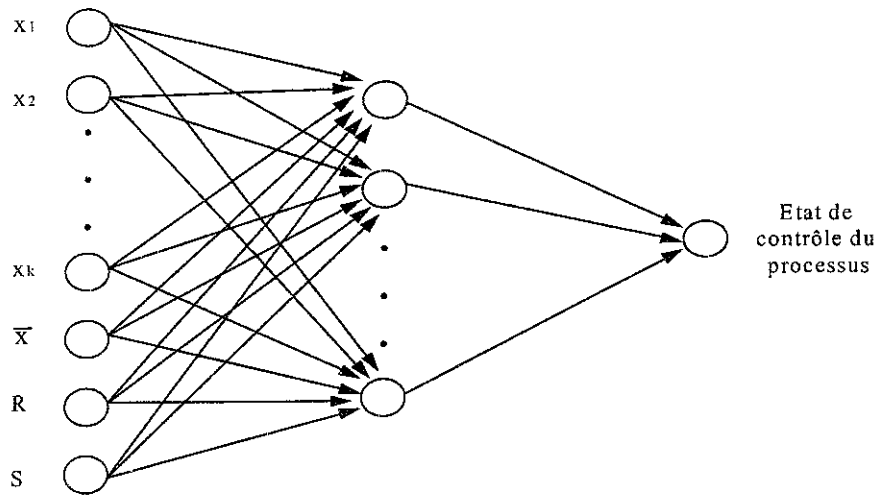


Figure 2: Interprétation des CC par le perceptron multicouche

L'étude comparative effectuée par Smith entre RNA et Carte (\bar{X} , R) pour une taille d'échantillon égale à 10 n'a pris en compte qu'une seule efficacité des cartes \bar{X} et R relative à une probabilité d'acceptation P_a égale à 0.45 pour la carte de la moyenne et une POM (Période Opérationnelle Moyenne) égale à 1.3480 pour la carte R. Il serait intéressant de voir comment les RNA vont se comporter face à des cartes \bar{X} , R et S ayant une meilleure efficacité. Dans cet objectif, trois tailles d'échantillon sont choisies à savoir $n=10$, $n=8$ et $n=5$. Ils sont relatifs à des efficacités respectives de 0.08, de 0.16 et de 0.45 pour la carte \bar{X} . Deux types de variations ont été choisis, les petites et les grandes variations au niveau de la moyenne et de la variance.

Pour les petites variations, la génération des observations relatives à l'apprentissage a été effectuée de la manière suivante:

- Générer M valeurs $N(0, 1)$ représentant l'état de sous contrôle du processus. Avec M : constante choisie en fonction de la taille de l'échantillon.
- Générer M valeurs $N(1.4, 1)$ représentant une variation de la moyenne.
- Générer M valeurs $N(0, 2)$ représentant une variation de la variance.

Pour les grandes variations, on a procédé de la même manière, c'est-à-dire :

- Générer M valeurs $N(0, 1)$ représentant l'état de sous contrôle du processus.
- Générer M valeurs $N(4.2, 1)$ représentant une variation de la moyenne.
- Générer M valeurs $N(0, 3)$ représentant une variation de la variance.

Un ensemble nommé A représentant les petits écarts de variation et un autre B pour les grands écarts de variation sont construits. Chaque élément de ces ensembles est normalisé selon la formule suivante :

$$n_i = x_i / x_{\max}$$

où n_i est la valeur normalisée, x_i est la donnée à normaliser, x_{\max} est la valeur maximale en valeur absolue de tous les exemples de l'ensemble en cours de traitement.

Chaque élément des ensembles A et B est accompagné de l'état de sortie qu'il représente. 75% des exemples d'apprentissage, choisis d'une manière aléatoire, sont utilisés pour l'entraînement du réseau, tandis que les 25% restants sont utilisés pour le test de celui-ci. Dans cet entraînement, la convergence de plusieurs types de réseaux va être testée selon le nombre de couches cachées et le nombre de neurones par couche cachée. Plusieurs réseaux ont été testés, les meilleurs ont été sélectionnés selon le critère de l'erreur RMS et le taux de classification. L'erreur RMS est définie

$$\text{par: RMS} = (\text{SSE} / p_{\max})^2 \quad \text{avec} \quad \text{SSE} = \sum_{p=1}^{p_{\max}} \sum_{q=1}^{q_{\max}} (T(p,q) - O(p,q))^2$$

où p_{\max} est le nombre d'exemples de l'ensemble d'apprentissage, q_{\max} est le nombre de nœuds en sortie, $T(p,q)$ est la sortie q du réseau calculée pour l'exemple p d'apprentissage et $O(p,q)$ est la sortie q désirée du réseau pour l'exemple p d'apprentissage.

Les réseaux retenus vont être comparés avec les cartes de contrôle standard de Shewhart à savoir X-bar R et S. Les exemples d'apprentissage et de test relatifs aux ensembles A et B sont appliqués à ces limites de contrôle et une décision d'un état de hors contrôle est déclarée si un point dépasse les limites de contrôle.

Les comparaisons effectuées entre RNA et cartes de contrôles (X-bar, R) et (X-bar, S), pour différentes efficacités de ces cartes, sont basées sur le pourcentage d'erreur de diagnostic sur l'état de contrôle de procédé. Les résultats, présentés dans le tableau 1, montrent que le modèle de perceptron est plus performant que ces cartes dans la détection des dérèglages survenus au processus et ceci pour les grands et petits écarts de variations.

Tableau 1: Résultats de simulation des RNA et cartes (X-bar, R), (X-bar, S).

n	Ensemble d'App et de Test	Entrée	Nbre de Couches Cachées	Nbre de nœuds par CC	Erreur RMS	Pourcentage d'Erreurs		
						Réseau	Carte (X-bar, R)	Carte (X-bar, S)
10	A	13	2	13	0.03	15.56	21	18.66
	B	3	2	3	0.01	0.44	12.33	11
8	A	11	2	3	0.02	14.67	24.6	22.3
	B	3	2	3	0.01	0.44	10.66	3.33
5	A	3	2	3	0.01	6.66	16.33	15.33
	B	8	2	8	0.00	8	16.66	15.66

Par exemple, pour $n = 10$, le perceptron fait moins d'erreurs que les cartes \bar{X} et R (15.56 % contre 21%) et ce pour les petites variations. Pour les grandes variations, le RNA fait nettement moins d'erreurs (0.44% contre 12.33%) que (\bar{X} , R). De même, le tableau 1 montre que la carte (\bar{X} , S) est plus intéressante que la carte (\bar{X} , R) puisqu'elle fait moins d'erreurs. Ce qui confirme le fait que la carte S est plus sensible à la détection des variations que la carte R.

Notons que dans cette étude, le réseau de neurones effectue les deux types d'erreurs (type α et type β). Par contre, les cartes de contrôle effectuent uniquement les erreurs de type β . Ceci doit être pris en considération lors de l'interprétation des résultats obtenus.

4. Diagnostic des cartes (CUSUM, R) et (CUSUM, S) par les RNA

Etant donnée que les résultats de comparaisons entre RNA et cartes de contrôle de Shewhart pour la détection d'un dérèglement au niveau de la tendance centrale et de la dispersion étaient encourageants, une extension de cette approche au cas des cartes spécialisées dans la détection de petits écarts de variation a été effectuée. Cette approche se base sur les travaux de Smith (1994) relatifs à l'interprétation des cartes \bar{X} et R par les RNA. Le perceptron multicouche proposé dans la section 2 est utilisé pour fournir la connaissance sur l'état de contrôle d'un processus et l'information sur un changement de structure au niveau de la moyenne ou de la variance d'une caractéristique qualité. Ceci va permettre de diminuer le champ d'action et réduire la taille de l'ensemble des causes possibles qui doivent être étudiées.

Le modèle neuronal a pour objectif de voir comment les RNA peuvent substituer la carte (CUSUM, R) ou la carte (CUSUM, S) pour l'étude comportementale d'une caractéristique qualité lors de l'apparition des petites variations. L'apprentissage supervisé utilisé dans cette approche se base sur la rétropropagation du gradient comme algorithme d'entraînement. Son exécution nécessite la constitution d'un ensemble d'apprentissage à partir de données réelles ou simulées. Une fois que le réseau a appris les exemples d'apprentissage, on passe à la phase test au cours de laquelle ce réseau est amené à diagnostiquer et prévoir une sortie adéquate pour toute entrée de l'ensemble test. Dans la phase utilisation, le responsable de la production peut fournir au réseau les données du processus et ce d'une manière en ligne. Dès que la réponse du réseau est obtenue, il doit investiguer les causes qui ont provoqué la variation en se basant sur l'information reçue du réseau et agir en conséquence.

Dans cette étude, le processus considéré et faisant sujet de contrôle est aussi supposé représenté par une caractéristique qualité univariée qui, une fois mesurée à des intervalles de temps réguliers, peut être modélisée par une variable aléatoire x_t normalement distribuée de moyenne et variance constantes. Sans perdre de généralité, on peut supposer que la moyenne du processus μ_0 est égale à 0 et sa variance σ_0^2 est égale à 1. A un instant donné, le processus va présenter une cause spéciale ayant l'un des deux effets suivants:

a - Variation dans la moyenne: $x_t = k + e_t$; où e_t est une variable qui suit la loi $N(0, 1)$.

b - Variation dans la variance: $x_t = ce_t$; où e_t est une variable qui suit la loi $N(0, 1)$

Pour la carte CUSUM, le risque α correspondant aux fausses alarmes est choisi égal à 0.27%, il est calculé lorsque le dérèglement est nul, ce qui correspond à une valeur de POM égale à 370. Pour les cartes R et S, les POM respectives sont assez grandes (POM supérieure à 33.7 pour la carte R et supérieure à 370 pour la carte S) et ce pour les deux valeurs de taille d'échantillon ($n=25$ et $n=16$).

Vu que la carte CUSUM est efficace pour la détection des petits écarts de variation, deux valeurs du dérèglement au niveau de la moyenne vont être prises en considération dans cette étude. La première est égale à σ_0 et correspond à un échantillon de taille $n=25$. La deuxième est prise égale à $1.25\sigma_0$ et elle correspond à un échantillon d'effectif $n=16$. La valeur de POM de la carte CUSUM relative à ces deux tailles d'échantillons est égale à 1.

La variation de l'écart type est prise égale au double de l'écart type initial du processus ($\sigma/\sigma_0 = 2=c$), ce qui correspond à une petite variation de la variance. Les deux dérèglements choisis ($k=1$ et $k=1.25$) correspondent à un coefficient de sensibilité égal à 1 et à un intervalle de décision h égal à 2.517.

La préparation des données d'apprentissage et de test nécessite la création de $3M$ variables :

- M variables $N(0, 1)$ représentant l'état de sous contrôle du processus. M étant une valeur numérique choisie en fonction de la taille de l'échantillon
- M variables $N(1, 1)$ représentant une variation de la moyenne.
- M variables $N(0, 2)$ représentant une variation de la variance.

Pour $n=25$, un ensemble d'entraînement et de test représentant les petits écarts de variation est construit. Toutes les données des exemples sont accompagnées de leur moyenne, leur étendue et leur écart type et sont normalisées selon la méthode présentée dans la section 2.

L'entraînement du réseau va être fait avec 75% des exemples d'apprentissage, choisis d'une manière aléatoire, tandis que le test va s'effectuer avec les 25% des exemples restants. Plusieurs types de réseaux sont testés selon le nombre de couches cachées et le nombre de neurones par couche. La phase de comparaison est entamée après avoir choisi les réseaux ayant donné les meilleurs résultats. Les tests effectués dans cette phase sur ces différents réseaux vont être comparés avec la carte CUSUM et les cartes de contrôle standards de Shewhart à savoir R et S. Les limites de contrôle des cartes R (respectivement S) calculées pour un échantillon de taille 25, une moyenne égale à 0 et un écart type égal à 1 sont 6.058 et 1.420 (respectivement 1.804 et 0.559).

Pour le cas $n=16$, la génération des variables se fait de la même manière sauf que le dérèglement de la moyenne est pris égal à 1.2 l'écart type initial. Les limites de contrôle de la carte R (respectivement S) sont 5.779 et 1.285 (respectivement 1.526 et 0.44).

Vu que la carte CUSUM bilatérale (utilisant S_{Hi} et S_{Li}) de la moyenne fait la somme cumulée des différences entre la valeur mesurée et la valeur de référence μ_0 . Et étant donné que la POM de cette carte est choisie égale à 1. Alors, la comparaison, pour un échantillon donné, va être effectuée à POM égales entre réseau de neurones et carte CUSUM de la manière suivante:

- Calculer pour tous les exemples d'apprentissage, représentant l'état de contrôle du processus, la valeur de z_i , S_{Hi} et S_{Li}
- Choisir un exemple provenant de l'ensemble d'apprentissage et de test, représentant un état quelconque du processus. Calculer son S_{Hi} et S_{Li} et voir si S_{Hi} ou S_{Li} dépasse la valeur de h ou non.
- Conclure enfin sur la validité du diagnostic de la carte (correspond ou non à l'état de contrôle représenté par l'échantillon).

Pour la carte R ou S, on calcule la valeur de R et S de l'exemple et une décision d'un état de hors contrôle est adoptée si un point dépasse les limites de contrôle. Les résultats des tests de comparaison sont présentés dans le tableau 2.

Tableau 2: Résultats de simulation des cartes CUSUM, R et S et du RNA pour $n=25$ et $n=16$

Ensemble d'App et de test du Réseau	Entrée	Nbre de Couches Cachées	Nbre de nœuds par CC	Erreur RMS	Pourcentage d'erreurs		
					Réseau	Carte (CUSUM, R)	Carte (CUSUM, S)
$n=25$	28	2	30	0.00	4	8.33	5.66
$n=25$	28	2	3	0.01	4	8.33	5.66
$n=25$	3	2	3	0.01	0.89	8.33	5.66
$n=16$	19	2	19	0.01	8	13.33	11
$n=16$	19	2	3	0.00	8	13.33	11
$n=16$	3	2	3	0.00	4.44	13.33	11

Le tableau 2 indique que le modèle de perceptron utilisé fait à peu près la moitié des erreurs effectuées par la carte (CUSUM, R), 4% contre 8.33%. Avec un réseau ayant 3 entrées, on peut aller jusqu'à 0.89% contre 8.33%. De même, ce modèle se compare à une carte (CUSUM, S); 4% d'erreur contre 5.66% et peut largement la dépasser si on utilise les caractéristiques statistiques de l'échantillon uniquement. Avec un dérèglement égal à 1.25, le réseau fait moins d'erreurs que la carte (CUSUM, R), 8% contre 13.3%, et peut aller jusqu'à 4.44% contre 13.3%. De même, le perceptron se compare à la carte (CUSUM, S), 8% contre 11%, et la dépasse avec un réseau de 3 entrées; 4.44% contre 11%. L'ajout du cas $n=16$ relatif au dérèglement $\delta=1.25$ montre la diminution de la sensibilité des cartes CUSUM, R et S et l'augmentation de la différence d'erreur entre ces cartes et le RNA puisque la différence d'erreur est passée de 4.77% à 6.56% entre RNA et carte (CUSUM, S)

et de 7.44% à 8.89% entre RNA et carte (CUSUM, R).

5. Interprétation des cartes multivariées par les RNA

Les cartes de contrôle univariées sont utilisées pour le contrôle d'une seule caractéristique qualité d'un produit ou d'un procédé. Lorsqu'on est en présence de p ($p > 1$) caractéristiques à contrôler simultanément, la pratique habituelle consiste à utiliser une carte de contrôle unidimensionnelle pour chacune des caractéristiques. Cependant, cette méthode a l'inconvénient de ne pas tenir compte des éventuelles corrélations entre les variables représentant ces caractéristiques.

Pour remédier à ces inconvénients, on peut avoir recours à des outils statistiques telles que les cartes de contrôle multivariées de Shewhart pour le suivi du comportement global d'un procédé contrôlé par plusieurs paramètres corrélés. Mais, l'utilisation et l'interprétation de ce type de cartes nécessite le concours d'un spécialiste. En plus, si les cartes multivariées sont capables de détecter un dérèglement, elles le sont moins pour la détermination des causes assignables et pour l'identification de ou des variables responsables de l'état de hors contrôle du procédé. Leur nature globale fait qu'elles ne permettent pas en général de savoir sur quelle (s) variable (s) le dérèglement s'est produit.

De ce fait, et étant donné qu'en pratique la plupart des procédés se caractérisent par la dépendance entre les paramètres qui les contrôlent, on propose d'utiliser l'approche neuronale pour contrôler, diagnostiquer, et interpréter l'état d'un processus et remplacer l'utilisation des cartes multivariées de Shewhart. Cette utilisation concerne l'interprétation d'une carte de contrôle bivariée de la moyenne, mais elle peut être étendue pour le cas multivarié. Dans cette approche, le RNA en l'occurrence le perceptron multicouche va être utilisé pour détecter toute déviation globale du procédé et déterminer si la déviation observée est causée par une variation au niveau de la moyenne d'une ou des deux variables à la fois.

Ainsi, dans le cas où on est en présence d'un procédé contrôlé par deux variables corrélées, x_1 et x_2 , le modèle du perceptron sera constitué d'une couche d'entrées composée de $2(n+1)+1$ nœuds et d'une couche de sortie ayant quatre nœuds responsables de l'état de contrôle global du procédé (figure 3). Les $n+1$ premiers nœuds de la couche d'entrée représentent n prélèvements sur la première variable, x_1 , ainsi que leur moyenne. Les $n+1$ nœuds restants constituent les n observations sur la deuxième variable, x_2 , ainsi que leur moyenne. Le nœud restant représente la valeur de la statistique χ^2 définie dans Montgomery (1991) par:

$$\chi^2 = n(\bar{X}_i - \mu_0)' \Sigma_0^{-1} (\bar{X}_i - \mu_0)$$
 avec n : taille d'un échantillon, Σ_0 : matrice variances-covariances et \bar{X}_i : moyenne de l'échantillon i .

Les nœuds binaires de la couche de sortie représentent successivement l'état de contrôle du processus multivarié, la variation au niveau de la moyenne de la variable x_1 , la variation au niveau de la moyenne de x_2 et la variation simultanée au niveau de la moyenne de x_1 et x_2 .

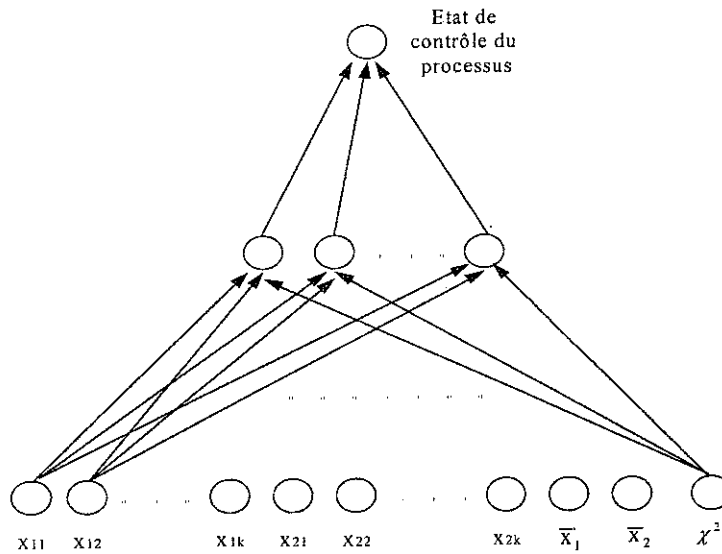


Figure 3 : Interprétation d'une carte bvariée par les RNA

Ce type de données peut être obtenu à partir d'une simulation. En effet, le processus considéré est représenté par deux caractéristiques qualité univariées qui, à cause de leur corrélation, vont être représentées par une caractéristique bidimensionnelle qui, une fois mesurée à des intervalles de temps réguliers peut être modélisée par une variable aléatoire $X_t = \begin{bmatrix} x_{t1} \\ x_{t2} \end{bmatrix}$ de \mathbb{R}^2 de loi normale bvariée de moyenne et de matrice variances-covariances constantes. Sans perdre de généralité, on peut supposer que la moyenne du processus est μ_0 et sa matrice variances-covariances est Σ_0 :

$$\mu_0 = \begin{bmatrix} \mu_{01} \\ \mu_{02} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ et } \Sigma_0 = \begin{bmatrix} \sigma_{01}^2 & \rho\sigma_{01}\sigma_{02} \\ \rho\sigma_{01}\sigma_{02} & \sigma_{02}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

Les deux variables x_{t1} et x_{t2} suivent des lois normales marginales de moyenne $\mu_{01} = 0$ et de variance $\sigma_{01}^2 = 1$ (respectivement de moyenne $\mu_{02} = 0$ et de variance $\sigma_{02}^2 = 1$).

A un instant donné, le processus va présenter une cause spéciale signalée par la variation de sa moyenne et provoquée par la variation de la moyenne de x_{t1} , de x_{t2} ou des deux à la fois. On a alors $X_t = k + e_t$ où e_t est une variable qui suit la loi normale bvariée centrée réduite et $K = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ est l'amplitude de la variation de moyenne en terme d'écart type. Dans le but de simplifier l'étude, nous allons considérer uniquement les variations positives des moyennes ($k_1 > 0$ et $k_2 > 0$). La taille d'un échantillon relative à l'une des deux variables est choisie égale à 4.

Dans cet exemple et pour les deux variables x_{t1} et x_{t2} , les valeurs des dérèglages k_1 et k_2 sont choisies égales à 1 ce qui correspond à des petites variations des moyennes. Pour les grands écarts

de variation au niveau des moyennes, les dérèglages k_1 et k_2 sont choisis égaux à trois fois le dérèglement initial. Avant d'entamer la comparaison entre RNA et carte bivariée, on doit créer le réseau, préparer les exemples relatifs à l'apprentissage et le test, normaliser et fournir ces données au réseau, effectuer son apprentissage puis son test. La préparation des données d'apprentissage et de test pour un perceptron ayant 3 ou 11 nœuds en entrée nécessite tout d'abord la création de plusieurs variables de la manière suivante:

- Générer 400 valeurs relatives à une variable normale standard z_1 de moyenne 0 et de variance 1.
- Générer 400 valeurs relatives à une variable normale standard z_2 de moyenne 0 et de variance 1.

Pour que X_i suit une loi $N(\mu_0, \Sigma_0)$, Bickel et Doksum (1977) proposent la démarche suivante pour générer les variables x_{t1} et x_{t2} à partir de z_1 et z_2 :

$$x_{t1} = \mu_{01} + \sigma_{01}z_1 \quad \text{et} \quad x_{t2} = \mu_{02} + \sigma_{02}(\rho z_1 + z_2\sqrt{1-\rho^2})$$

En fixant les valeurs de μ_{01} , μ_{02} et en appliquant cette méthode sur les variables z_1 et z_2 générées, on aura 400 variables x_{t1} et 400 variables x_{t2} . Vu que la taille d'un échantillon est choisie égale à 4, on aura 100 sous-groupes correspondant à la variable X_i . La génération de ces échantillons pour les différents dérèglages de la moyenne se fait comme suit:

- Pour $\mu_{01} = \mu_{02} = 0$ et $\sigma_{01} = \sigma_{02} = 1$, construire 100 sous-groupes représentant l'état de sous contrôle du processus.
- Pour $\mu_{01} = 1$, $\mu_{02} = 0$ et $\sigma_{01} = \sigma_{02} = 1$, construire 100 sous-groupes représentant l'état de variation au niveau de la moyenne de x_{t1} .
- Pour $\mu_{01} = 0$, $\mu_{02} = 1$ et $\sigma_{01} = \sigma_{02} = 1$, construire 100 sous-groupes représentant l'état de variation au niveau de la moyenne de x_{t2} .
- Pour $\mu_{01} = \mu_{02} = 1$ et $\sigma_{01} = \sigma_{02} = 1$, construire 100 sous-groupes représentant l'état de variation simultanée au niveau de la moyenne de x_{t1} et x_{t2} .

A partir de ces échantillons, on va créer un ensemble nommé A représentant les petits écarts de variation et composé de 400 exemples nécessaires pour l'apprentissage et le test du réseau. Pour les grands écarts de variation, plusieurs autres variables doivent être générées de la même manière, en considérant un dérèglement égal à 3 fois l'écart type initial. Un autre ensemble nommé B représentant ces grands écarts de variation va donc être créé et sera composé de 400 exemples.

Chaque sous-groupe des ensembles A et B doit être accompagné de la moyenne de x_{t1} , celle de x_{t2} et de sa statistique χ^2 . Chaque sous-groupe construit à partir de la génération des variables va être associé à une valeur de sortie qui lui correspond 75% des exemples d'apprentissage, choisis d'une manière aléatoire, sont utilisés pour l'entraînement du réseau, tandis que les 25% restants sont utilisés pour le test de celui-ci. Dans cet entraînement, la convergence de plusieurs types de réseaux va être testée selon le nombre de couches cachées et le nombre de neurones par couche. La comparaison de ces réseaux va être effectuée sur la base de l'erreur de l'apprentissage RMS et le taux de classification.

Plusieurs réseaux ont été retenus, ils vont être comparés avec la carte de contrôle bivariée de la moyenne. Les limites de contrôle de cette carte calculées pour un échantillon de taille égale à 4, suivant une moyenne μ_0 et une matrice variances-covariances Σ_0 sont: LIC = 0 et LSC = $\chi_{p,\alpha}^2$. Plusieurs valeurs de risque α vont être considérées à savoir $\alpha = 1\%$, $\alpha = 5\%$ et $\alpha = 10\%$. Pour chaque valeur de α , les exemples d'apprentissage et de test relatifs aux ensembles A et B sont appliqués aux limites de contrôle correspondantes et une décision d'un état de hors contrôle est déclarée si un point dépasse la limite de contrôle LSC. Les résultats des tests de comparaison entre la carte bivariée et le RNA sont présentés dans le tableau 3

Le tableau 3 montre que, pour les grandes variations au niveau des moyennes (ensemble B), la performance du perceptron multicouche est comparable à celle de la carte bivariée pour un risque α égal à 0.01 et 0.05 (0% contre 0.25% pour $\alpha = 1\%$ et 0% contre 1.25% d'erreurs pour $\alpha = 5\%$). Pour $\alpha = 0.1$, le réseau fait beaucoup moins d'erreurs que la carte χ^2 (0% contre 3.25%). Pour les petits écarts de variation, le réseau de neurones fait moins d'erreurs que la carte pour un risque α égal à 0.01 (28% d'erreur contre 37.25%). Par contre, il fait un peu plus d'erreurs que la carte lorsque les valeurs de α sont prises égales à 0.05 et 0.1 (28% contre 23 % pour $\alpha = 5\%$ et 28% contre 19.75 pour $\alpha = 10\%$)

Les erreurs effectuées par la carte χ^2 sont relatives aux fausses alarmes (risque α) et aux erreurs relatives à la non détection de dérèglement (risque β). Cependant, celles effectuées par le RNA sont composées de l'erreur α , de l'erreur β ainsi que celle relative aux mauvaises classifications bien que l'état de hors contrôle soit détecté. Ceci doit être pris en considération lors de l'interprétation des pourcentages d'erreurs commises par le RNA.

Tableau 3: Résultats de simulation de la carte bivariée et du RNA.

Ensemble d'App et de Test du Réseau	Entrée	Nbre de Couches Cachées	Nbre de nœuds par CC	Erreur RMS	Pourcentage d'erreurs			
					Réseau	Carte Bivariée		
						$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
A	11	2	11	0.02	28	37.25	23	19.75
A	11	2	4	0.15	30.66	37.25	23	19.75
A	3	2	4	0.13	30	37.25	23	19.75
B	11	2	11	0.01	1	0.25	1.25	3.25
B	11	2	4	0.01	0	0.25	1.25	3.25
B	3	2	4	0.02	0	0.25	1.25	3.25

L'étude comparative montre que si la carte bivariée a réussi à détecter les dérèglages dans la moyenne globale du processus, elle ne peut identifier la ou les variables responsables de ces dérèglages. Ceci est possible pour les réseaux de neurones qui indiquent aussi bien la présence d'un dérèglement que la ou les variables qui en sont responsables.

6. Conclusion

Bien que l'utilisation des RNA ait soulevé certaines difficultés au niveau du choix des paramètres d'architecture et d'apprentissage, les comparaisons effectuées entre RNA et cartes (X-bar, R), (X-bar, S), (CUSUM, R), (CUSUM, S) et χ^2 , pour différents niveaux de dérèglement et d'efficacité, et basées sur le pourcentage d'erreur de diagnostic comme critère de comparaison, montrent que les réseaux multicouches sont similaires et même meilleurs que ces cartes dans l'interprétation de l'état de contrôle d'un procédé et de son diagnostic.

Cependant, d'autres comparaisons, basées sur de nouveaux critères ou bien d'autres tailles d'échantillons entre les RNA et autres types de cartes ou de combinaisons de cartes seront nécessaires pour consolider l'utilité et l'efficacité de ces réseaux dans l'interprétation des phénomènes étudiés. De même, l'utilisation d'autres types d'entraînements comme l'apprentissage adaptatif ainsi que l'implantation d'algorithmes plus performants serait plus bénéfique pour l'augmentation de la performance des RNA en ce qui concerne la rapidité de convergence et le taux de reconnaissance et par suite la diminution des erreurs de prévision et de diagnostic.

Les modèles neuronaux étudiés peuvent être dotés d'un tableau de bord capable d'avertir à temps, à partir des résultats fournis par ces modèles, les responsables du suivi de la qualité du produit pour qu'ils puissent agir opportunément et éviter le risque d'avoir des produits non conformes ou de non-qualité. De même, l'utilisation parallèle d'un système interactif d'aide à la décision pourvu d'un système expert serait intéressante pour diagnostiquer les causes des variations détectées par les RNA, orienter l'utilisateur vers les sources d'ennuis, recommander les actions correctives et proposer des solutions adéquates.

7. Bibliographie

- Bickel, J. et Doksum A. (1977), *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day, Inc.
- Montgomery (1991), *Introduction to Statistical Quality Control (Second Edition)*. John Wiley & Sons, Inc.
- Niang, N (1994), *Méthodes multidimensionnelles pour la maîtrise statistique des procédés*. Thèse en statistique obtenue à l'Université de Paris Dauphine.
- Ryan, T P. (1989), *Statistical methods for quality improvement*. John Wiley & Sons
- Smith, A.E. (1994), X-bar and R Control Chart Interpretation Using Neural Computing. *International Journal of Production Research*, Vol, 32, N°2, pp. 309-320.
- Yazici, H. et Smith A.E. (1992), An Intelligent Composite System for Statistical Process Control. *Engineering Applications of Artificial Intelligence*. Vol, 5, N°6, pp. 519-526.