

AI.1  
9  
910

Université de Montréal

Les réseaux de neurones et la prévision des taux de change quotidiens

par  
Slim Chamakhi

Centre de Documentation  
Dép. de sciences économiques  
Université de Montréal  
C. P. 6128, Succ. "A"  
Montréal, Qué., Canada, H3C 2S1

Département de sciences économiques  
Faculté des arts et des sciences

Rapport de recherche présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de M.Sc.  
en Sciences Économiques  
option Économétrie

novembre 2002

© Slim Chamakhi, 2002

Université de Montréal  
Faculté des études supérieures

Ce rapport de recherche intitulé :  
Les réseaux de neurones et la prévision des taux de change quotidiens

présenté par  
Slim Chamakhi

a été évalué par un jury composé des personnes suivantes :

Benoit Perron  
directeur de recherche

Silvia Gonçalves  
deuxième lecteur

## Table des Matières

<i>Sommaire</i>	iv
<i>Liste des Graphiques</i>	v
<i>Liste des Tableaux</i>	vi
<b>Section I. – Introduction</b>	1
<b>Section II. – Les Réseaux de Neurones</b>	4
Définition	4
Estimation	7
Sélection et Validation	9
Évaluation	11
<b>Section III. – Les Études Antérieures</b>	16
Kuan & Liu (1995)	17
Franses & Van Homelen (1998)	17
Franses & Van Griensven (1998)	18
Gençay (1999)	19
Gradojevic & Yang (2000)	20
<b>Section IV. – Les Déterminants des Taux de Change</b>	23
Les Rendements Passés	23
L'Analyse Technique	24
Les Différentiels de Taux d'intérêt à Court Terme	25
<b>Section V. – Les Données</b>	28
Les Rendements Quotidiens	29
Les Changements Quotidiens des Différentiels	30
<b>Section VI. – Les Résultats Empiriques</b>	31
Le Dollar Canadien	31
Le Franc Français	32
Le Mark Allemand	33
La Livre Anglaise	34
<b>Section VII. – Conclusion</b>	37
<i>Bibliographie</i>	vii
<i>Annexe</i>	ix

## Sommaire

Les réseaux de neurones sont employés pour la prévision des rendements quotidiens de cinq séries de taux de change par rapport au dollar américain : le dollar canadien, le franc français, le Mark allemand et la Livre anglaise. Trois types de variables sont considérés comme inputs : les réalisations passées des rendements, des variables empruntées à l'analyse technique et des différentiels de taux d'intérêt à court terme. La performance prédictive des modèles est évaluée selon deux critères. Le premier mesure la capacité du modèle à prévoir le signe du rendement futur alors que le second mesure la précision des prévisions. Pour éviter les problèmes de sur-ajustement, une technique de validation est également employée.

Les résultats sont sensibles à la série étudiée et à la variable introduite comme input du modèle. Les réseaux de neurones sont plus performants que le modèle de référence : la marche aléatoire.

**Mots-clefs :** Réseaux de neurones, validation, performance prédictive

## Liste des Graphiques

<b>Graphique 1.</b> – Représentation du Réseau de Neurones (3,2)	5
<b>Graphique 2.</b> – Analyse Technique – Dollars Canadiens/Dollar Américain	ix
<b>Graphique 3.</b> – Taux de Change Dollars Canadiens/Dollar Américain	x
<b>Graphique 4.</b> – Taux de Change Franc Français/Dollar Américain	xi
<b>Graphique 5.</b> – Taux de Change Mark Allemand/Dollar Américain	xii
<b>Graphique 6.</b> – Taux de Change Livre Anglaise/Dollar Américain	xiii
<b>Graphique 7.</b> – Différentiel Taux Libor Dollar Canadien – Dollar Américain	xiv
<b>Graphique 8.</b> – Différentiel Taux Libor Franc Français – Dollar Américain	xv
<b>Graphique 9.</b> – Différentiel Taux Libor Mark Allemand – Dollar Américain	xvi
<b>Graphique 10.</b> – Différentiel Taux Libor Livre Anglaise – Dollar Américain	xvii

### Liste des Tableaux

<b>Tableau I.</b>	– Statistiques – Rendements Quotidiens de Taux de Change	xviii
<b>Tableau II.</b>	– Statistiques – Changements Quotidiens des Différentiels des Taux LIBOR 3 mois	xix
<b>Tableau III.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change CAD/USD Rendements Passés	xx
<b>Tableau IV.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change CAD/USD Signaux de vente-achat, $n_1 = 1, n_2 = 50$	xxi
<b>Tableau V.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change CAD/USD Signaux de vente-achat, $n_1 = 1, n_2 = 200$	xxii
<b>Tableau VI.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change CAD/USD Changements quotidiens des différentiels de taux d'intérêt	xxiii
<b>Tableau VII.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change FF/USD Rendements Passés	xxiv
<b>Tableau VIII.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change FF/USD Signaux de vente-achat, $n_1 = 1, n_2 = 50$	xxv
<b>Tableau IX.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change FF/USD Signaux de vente-achat, $n_1 = 1, n_2 = 200$	xxvi
<b>Tableau X.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change FF/USD Changements quotidiens des différentiels de taux d'intérêt	xxvii
<b>Tableau XI.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change DM/USD Rendements Passés	xxviii
<b>Tableau XII.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change DM/USD Signaux de vente-achat, $n_1 = 1, n_2 = 50$	xxix
<b>Tableau XIII.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change DM/USD Signaux de vente-achat, $n_1 = 1, n_2 = 200$	xxx
<b>Tableau XIV.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change DM/USD Changements quotidiens des différentiels de taux d'intérêt	xxxi
<b>Tableau XV.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change BP/USD Rendements Passés	xxxii
<b>Tableau XVI.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change BP/USD Signaux de vente-achat, $n_1 = 1, n_2 = 50$	xxxiii
<b>Tableau XVII.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change BP/USD Signaux de vente-achat, $n_1 = 1, n_2 = 200$	xxxiv
<b>Tableau XVIII.</b>	– Performance du Réseau ( $p, q$ ) – Taux de Change BP/USD Changements quotidiens des différentiels de taux d'intérêt	xxxv

## I. Introduction :

Le développement de stratégies de placements profitables sur les marchés financiers requiert des prévisions sur les mouvements futurs des prix. Ainsi, il est primordial de pouvoir compter sur un modèle de prévision fiable afin d'ajuster les portefeuilles selon le comportement attendu des marchés. Ce rapport tentera de vérifier s'il est possible de construire un modèle de prévision de taux de change quotidiens à l'aide d'une méthode non-paramétrique, les réseaux de neurones.

L'utilisation des méthodes non-paramétriques dans la modélisation de processus dynamiques non-linéaires est de plus en plus répandue, tant en économie financière que dans d'autres domaines tels que la médecine ou la physique. Elles sont préférées aux méthodes paramétriques car elles ont l'avantage de demander un minimum de restrictions sur la forme supposée de la relation entre les variables du modèle et peuvent aider à découvrir des relations cachées par des hypothèses plus restrictives.

Il existe deux classes de méthodes non-paramétriques. Les méthodes locales (le noyau et le "*Nearest Neighbor*") utilisent les observations dans un voisinage du point d'intérêt alors que les méthodes globales (approximation en séries et les réseaux de neurones) exploitent l'ensemble des observations disponibles pour obtenir un estimateur de l'espérance conditionnelle au point d'intérêt. Cybenko (1989) et Hornik *et al* (1989) ont démontré que les réseaux de neurones pouvaient bien approximer une vaste classe de fonctions non-linéaires, les modèles plus classiques de séries chronologiques n'ayant pas cette propriété. Les réseaux de neurones ont souvent été utilisés avec succès dans l'analyse empirique de variables économiques ou financières à haute fréquence, nous pouvons nommer Kuan et White (1994) ou Samson et White (1995) à titre d'exemple. Le caractère non-linéaire des séries étudiées et la flexibilité des réseaux à capter ces non-linéarités justifient leur emploi pour bâtir des modèles de prévision des rendements quotidiens de taux de change.

Une description détaillée des réseaux est donnée dans la section II du rapport.

La flexibilité des réseaux de neurones – conjuguée à la faible performance des modèles paramétriques – a amené un nombre croissant de chercheurs à exploiter ces méthodes pour essayer de prévoir les comportements futurs des variables financières. Kuan et Liu (1995), Franses et Van Homelen (1998), Franses et Van Griensven (1998), Gençay (1999) et Gradojevic et Yang (2000) étudient tous des variantes des réseaux de neurones pour bâtir des modèles de prévision des rendements des taux de changes quotidiens. Nous présentons leurs études à la section III de ce rapport.

Les modèles étudiés dans ces articles se distinguent les uns des autres par le type de variable introduit pour décrire le mouvement des taux de change. Nous définissons à la section IV les différents modèles de taux de change considérés dans ce rapport. Trois types de variables sont introduits à tour de rôle comme inputs de nos réseaux, les réalisations passées des rendements quotidiens de taux de change, des variables empruntées à l'analyse technique (i.e. moyennes mobiles) et des différentiels de taux d'intérêt nominaux.

Certains auteurs, comme Hsieh (1989), Chinn (1991) ou Brooks (1996) par exemple, ont démontré que les séries de rendements de taux de change suivaient des processus dynamiques non-linéaires. Il semble alors justifié d'inclure les réalisations passées de la variable dépendante comme inputs du réseau pour expliquer les variations des taux de change quotidiens. Les outils d'analyse technique sont largement utilisés par les analystes dans leur décision d'achat ou de vente d'actifs financiers. L'étude des mouvements récents des prix permet de détecter des changements de tendance et d'ajuster en conséquence les positions prises sur le marché. La décision d'achat ou de vente est formée en comparant le niveau actuel des prix à une moyenne mobile des prix passés. La différence entre ces deux séries de prix est utilisée comme input dans la deuxième version du modèle. Enfin, dans un système de change flottant, l'observation des taux d'intérêt sur le marché monétaire international est indispensable pour essayer de comprendre (et de prévoir) l'évolution des taux de change, car leurs variations se reflètent sur celles des cours des devises. Le



différentiel de taux d'intérêt à court terme est donc introduit comme input dans la troisième version du réseau de neurones étudiée. Nous voulons déterminer ainsi quel modèle de taux de change permet de produire les prévisions les plus précises.

Les résultats empiriques de notre étude sont présentés à la section VI. Nous y rapportons les valeurs des différentes mesures utilisées pour évaluer la performance de nos réseaux de neurones dans la prévision des rendements quotidiens de quatre séries de taux de change par rapport au dollar américain : le dollar canadien, le franc français, le Mark allemand et la Livre anglaise. Nous utilisons les critères de sélection de Akaike et Schwarz comme outil de validation. Nous voulons choisir le réseau (parmi les différents couples  $(p,q)$  considérés) qui décrit le mieux la dynamique existant entre l'output et les inputs à l'intérieur de l'échantillon d'estimation tout en gardant un certain degré de généralisation pour produire les meilleures prévisions.

Avant de procéder à l'analyse des résultats, nous présentons à la section V des statistiques descriptives sur les séries étudiées. La section VII conclut le rapport.

## II. Les Réseaux de Neurones :

La structure à couches multiples des réseaux de neurones et leur capacité à traiter l'information n'est pas sans rappeler le fonctionnement du cerveau humain. La popularité grandissante de ce type de modèles ces dernières années a donné naissance à une nomenclature qui leur est propre. Les notions de couches, de nœuds cachés, d'inputs, d'output et d'entraînement sont des éléments de modélisation qui sont traités dans cette section. Nous présentons aussi les méthodes utilisées pour estimer les paramètres des modèles de réseaux de neurones de même que les modifications apportées pour améliorer les propriétés numériques de l'algorithme d'apprentissage et obtenir un meilleur estimateur de l'espérance conditionnelle. La dernière partie de la section définit les différentes statistiques utilisées pour évaluer la performance de nos modèles à prévoir les rendements futurs des taux de change.

### 1) Définition :

Le réseau de neurones est une fonction de régression non-linéaire caractérisant la relation entre une variable dépendante  $r_t$  (cible) et un vecteur de variables indépendantes  $x_t$  (inputs). Plutôt que de définir une relation spécifique, le réseau est construit de manière à combiner plusieurs fonctions non-linéaires simples dans une structure à couches multiples.

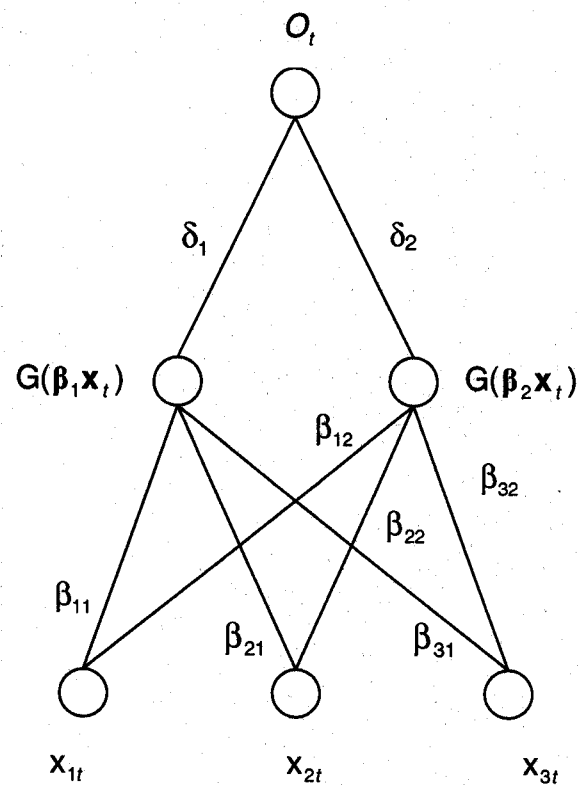
Le réseau (3,2) présenté à la page suivante est composé de  $p = 3$  inputs, de  $q = 2$  nœuds cachés et d'1 output  $O_t$ . Chaque élément du vecteur d'inputs  $x_t$  :

$$(1) \quad x_t = (x_{1t}, x_{2t}, x_{3t})'$$

se voit attribuer un poids  $\beta_j$  :

$$(2) \quad \beta_1 = (\beta_{11}, \beta_{21}, \beta_{31}) \quad \& \quad \beta_2 = (\beta_{12}, \beta_{22}, \beta_{32})$$

Graphique 1 : Représentation du Réseau (3,2)



avant d'être transformé par une fonction d'activation  $G(z)$  :

$$(3) \quad G(z) = \frac{1}{1 + e^{-z}}$$

et constituer les 2 nœuds de la couche cachée. L'utilisation de la forme logistique n'est pas obligatoire. Une quelconque fonction bornée peut servir de fonction d'activation (tangente ou hyperbolique par exemple).

Les 2 nœuds cachés sont à leur tour reliés à l'output  $O_t$  par les paramètres  $\delta_j$  :

$$(4) \quad O_t = \delta_1 G(\beta_1 x_t) + \delta_2 G(\beta_2 x_t)$$

L'output  $O_t$  est une somme pondérée de fonctions non-linéaires simples, elles-mêmes construites à partir de combinaisons linéaires des variables explicatives utilisées comme inputs,  $O_t$  représente notre estimateur de la variable d'intérêt  $r_t$ . Le réseau de neurones à composante linéaire est largement utilisé dans les travaux empiriques récents (Samson et White, 1995; Franses et Van Homelen, 1998; Franses et Van Griensven, 1998) et constitue une extension du réseau (3,2) présenté au graphique 1. Le vecteur d'inputs  $\mathbf{x}_t$  est introduit linéairement dans le modèle (4).

$$(5) \quad O_t = \alpha \mathbf{x}_t + \sum_{j=1}^q \delta_j G(\beta_j \mathbf{x}_t)$$

$\alpha \mathbf{x}_t$  exprime la partie linéaire de la relation entre le vecteur d'inputs  $\mathbf{x}_t$  et la variable cible  $r_t$ , alors que les composantes logistiques  $G(\beta_j \mathbf{x}_t)$  mesurent le degré de non-linéarité. Une constante est aussi ajoutée comme variable explicative. Le vecteur d'inputs comprend donc  $p + 1$  éléments:

$$(6) \quad \mathbf{x}_t = (1, x_{t1}, \dots, x_{pt})'$$

Pendant le processus d'entraînement, les poids  $\alpha_t$ ,  $\delta_j$ ,  $\beta_{ij}$  reliant les différentes couches du réseau (des inputs à l'output) sont ajustés pour réduire l'écart  $\varepsilon_t$  entre l'output du réseau  $O_t$  et la variable cible  $r_t$ . Les paramètres du modèle donné en (5) sont estimés par moindres carrés non-linéaires. Le processus itératif d'optimisation est arrêté lorsque la somme des écarts  $\varepsilon_t$  au carré entre l'output du réseau  $O_t$  et la variable cible  $r_t$  atteint sa valeur minimum.

2) Estimation :

Soit  $Q(\theta)$  la somme au carré des écarts  $\varepsilon_t$  entre l'output  $O_t$  et la variable dépendante  $r_t$  :

$$(7) \quad Q(\theta) = \sum_{t=p+1}^T \varepsilon_t^2 = \sum_{t=p+1}^T (r_t - O_t)^2$$

et  $\theta = (\alpha, \delta, \beta)$  le vecteur de dimension  $1 \times ((p+1)(q+1) + q)$  qui regroupe tous les paramètres du modèle, avec :

$$(8) \quad \begin{array}{ll} \alpha = (\alpha_0, \alpha_1, \dots, \alpha_p) & 1 \times (p+1) \\ \delta = (\delta_1, \dots, \delta_q) & 1 \times q \\ \beta = (\beta_1, \dots, \beta_q) & 1 \times q(p+1) \\ \text{et } \beta_j = (\beta_{0j}, \beta_{1j}, \dots, \beta_{pj}) & 1 \times (p+1) \end{array}$$

L'entraînement du réseau de neurones se résume à estimer les paramètres du modèle (5) par la minimisation de la somme des erreurs au carré donnée par  $Q(\theta)$ . L'exercice d'optimisation se fait par moindres-carrés non-linéaires (l'algorithme de Broyden-Fletcher-Goldfarb-Shanno par exemple, BFGS par la suite). Le processus itératif d'entraînement est arrêté lorsque l'algorithme a convergé. La somme des écarts au carré  $Q(\theta)$  possède plusieurs minimums locaux et la convergence de l'algorithme ne garantit pas l'atteinte du minimum global. Pour améliorer les propriétés numériques de l'algorithme de BFGS, certaines modifications sont apportées au modèle (voir Franses et Van Dijk, 2000, pages 220 à 222) :

a) Normalisation des variables du réseau :

$$(9) \quad z_i^* = \frac{z_i - \bar{z}_i}{\sigma(z_i)}$$

avec  $\bar{z}_i$  et  $\sigma(z_i)$  les estimateurs de la moyenne et de l'écart-type de  $z_i$  respectivement. La variable transformée  $z_i^*$  est centrée réduite avec une moyenne nulle et un écart-type égal à 1. Toutes les variables du modèle sont transformées, les inputs  $x_i$  et la variable cible  $r_i$ . Les variables sont donc symétriquement distribuées autour de leur moyenne, 0.

b) Imposition de pénalités à la somme des écarts au carré  $Q(\theta)$  :

La fonction-objectif définie en (7) est augmentée d'un terme de pénalité :

$$(10) \quad Q(\theta)^a = Q(\theta) + r_\alpha \sum_{i=1}^p \alpha_i^2 + r_\delta \sum_{j=1}^q \delta_j^2 + r_\beta \sum_{i=1}^p \sum_{j=1}^q \beta_{ij}^2$$

L'imposition de pénalités empêche les éléments individuels de  $\theta$  de devenir excessivement élevés pendant le processus de minimisation. Les pénalités sont définies a priori. Pour les réseaux considérés dans ce rapport, les valeurs des pénalités ont été fixées à  $r_\alpha = 0.01$  et  $r_\delta = r_\beta = 0.0001$  de la même manière que Franses et Van Dijk (2000). Il est nécessaire de normaliser les variables du réseau pour que les pénalités imposées à  $Q(\theta)$  soient efficaces.

c) Valeurs initiales :

Le choix des valeurs initiales  $\hat{\theta}^{(0)}$  est critique pour l'algorithme d'estimation. Le bon choix peut accélérer le temps d'entraînement et faciliter la convergence de l'algorithme de BFGS vers la solution optimale. Pour obtenir un meilleur point de

départ, nous avons généré dix vecteurs  $\hat{\theta}_\ell^{(0)}$ ,  $\ell=1,\dots,10$ , à partir de la loi Normale centrée réduite et celui qui produit la plus petite erreur quadratique moyenne pour l'échantillon d'estimation est choisi comme point de départ pour le processus de minimisation. L'algorithme du Simplexe est d'abord employé pour explorer de larges parties de l'espace des paramètres. Le vecteur  $\hat{\theta}_s^{(0)}$  ainsi obtenu est utilisé comme vecteur initial pour l'algorithme de BFGS.

Le réseau de neurones défini en (5) est entraîné pour différents couples  $(p,q)$ . Il est nécessaire maintenant de déterminer le couple d'inputs  $p$  et de nœuds cachés  $q$  qui permet d'obtenir l'estimateur de  $\theta$  qui produit les meilleurs résultats. Nous développons à cette fin une technique de sélection qui nous permettra de choisir le modèle  $(p,q)$  que nous utiliserons pour l'exercice de prévision. Les critères de Akaike et Schwarz sont calculés pour un échantillon de validation.

### 3) Sélection et Validation :

Avant de procéder à l'exercice de prévision, il faut déterminer la combinaison d'inputs  $p$  et de nœuds  $q$  qui livre le réseau optimal. Une stratégie courante est d'estimer les paramètres de réseaux  $(p,q)$  pour un nombre donné de valeurs de  $p$  et de  $q$  et utiliser des critères de sélection pour choisir le modèle le plus efficace. Les critères les plus courants sont le critère de Akaike :

$$(11) \quad AIC = n \ln(s^2) + 2k$$

ou le critère de Schwarz :

$$(12) \quad BIC = n \ln(s^2) + k \ln(n)$$

avec  $k = (p + 1)(q + 1) + q$  le nombre total de paramètres,  $n$  la taille de l'échantillon considéré et  $s^2$  l'estimateur de la variance résiduelle.

Ces fonctions de sélection pénalisent pour l'addition d'un trop grand nombre de variables dans le modèle et permettent de choisir simultanément le nombre d'inputs  $p$  et le de nœuds cachés  $q$ . Le modèle  $(p,q)$  qui produit la valeur minimum pour ces fonctions est sélectionné. Il faut noter ici que Schwarz (*BIC*) a tendance à sélectionner des modèles plus conservateurs (lire moins complexes) que Akaike (*AIC*) car il pénalise davantage l'inclusion de variables additionnelles. L'entraînement des réseaux de neurones se fait dans le but d'obtenir un degré maximal de généralisation. Nous procédons à une étape de validation pour éviter le sur-ajustement des données et s'assurer de sélectionner des réseaux  $(p,q)$  capables de généraliser les relations apprises à de nouveaux exemples (à l'extérieur de l'échantillon d'estimation).

Nous proposons d'utiliser les critères de sélection de Akaike et de Schwarz définis ci-haut comme outil de validation.

À cette fin, nous divisons l'échantillon disponible en trois groupes :

- un échantillon d'estimation (50% de l'échantillon total);
- un échantillon de validation (25% de l'échantillon total);
- un échantillon de prévision (25% de l'échantillon total);

Certains auteurs attribuent 60% de l'échantillon total pour l'entraînement des réseaux, 30% pour la validation et les 10% restants pour la prévision. Nous préférons utiliser des échantillons de tailles comparables pour les étapes de validation et de prévision. Nous réservons quatre années de données quotidiennes pour l'estimation et deux années de données quotidiennes pour les exercices de validation et de prévision respectivement pour un total de huit années de données quotidiennes.

Les données du premier échantillon (seulement) sont utilisées pour minimiser la somme des erreurs au carré présentée en (7) par l'algorithme de BFGS et obtenir un estimé du vecteur de paramètres  $\theta$ . L'exercice est répété pour toutes les combinaisons  $(p,q)$  avec  $p \in \{1, 2, 3, 4, 5\}$  et  $q \in \{1, 2, 3\}$ . Nous avons un total de quinze modèles et quinze vecteurs de paramètres à estimer.

Les critères de sélection de Akaike et Schwarz sont calculés pour l'échantillon de validation en utilisant les vecteurs de paramètres  $\hat{\theta}$  estimé à l'étape précédente



(pour les quinze modèles considérés). Les modèles  $(p,q)$  qui minimisent ces critères sont sélectionnés pour l'exercice de prévision.

Les vecteurs de paramètres  $\hat{\theta}$  obtenus à la première étape (estimation) et qui minimisent les critères de sélection calculés à la deuxième étape (validation) sont utilisés pour générer des prévisions à un jour en utilisant les données du troisième échantillon (prévision). Nous supposons ici que l'information pour les périodes  $t-1, t-2, \dots, t-p$  est disponible pour les prévisions à la période  $t$ , il en est de même pour les prévisions au temps  $t+1$ , et ainsi de suite pour les prévisions des périodes suivantes.

À titre de comparaison, nous calculons les critères *AIC* et *BIC* pour l'échantillon d'estimation. Ces critères sont employés dans d'autres contextes pour sélectionner le meilleur modèle parmi un ensemble de modèles concurrents sans recourir à une étape de validation au préalable. Nous voulons vérifier si l'utilisation des critères d'information au moment de l'estimation permet de choisir un modèle plus performant que les réseaux qui ont été sélectionnés à l'étape de validation.

La performance hors-échantillon des réseaux  $(p,q)$  est évaluée en utilisant deux critères : le premier est basé sur la prévision de signe (prévoir la direction des changements futurs) et l'autre sur la dispersion des prévisions par rapport aux rendements observés.

#### 4) Évaluation :

Les mesures telles que le taux de succès, la statistique de Pesaran et Timmermann (1992), l'erreur quadratique moyenne et la statistique de Diebold et Mariano (1995) sont utilisées pour mesurer la performance prédictive des réseaux  $(p,q)$  ainsi obtenus.

a) Le Taux de Succès :

Le taux de succès est souvent utilisé par les chercheurs intéressés à prévoir le sens des variations futures des prix d'actifs financiers. Anticiper le sens futur du mouvement des taux de change est essentiel pour ajuster les positions prises sur le marché. Cette mesure consiste à comptabiliser le nombre de fois que le modèle  $(p,q)$  étudié prévoit correctement le signe du rendement.

Soit  $r_t$  le rendement observé au temps  $t$  et soit  $\phi(x_t, \hat{\theta})$  notre estimateur du rendement  $r_t$ , une fonction  $\phi$  des inputs  $x_t$ , du vecteur des estimateurs  $\hat{\theta}$ , du nombre  $p$  de retards utilisés pour les inputs et du nombre  $q$  de nœuds cachés du réseau.

Le taux de succès  $SR$  (Success Ratio) est défini comme :

$$(13) \quad SR = \frac{1}{m} \sum_{t=1}^m \mathbf{I}_t [ r_t \cdot \phi(x_t, \hat{\theta}) > 0 ]$$

et  $m$  représente l'horizon de prévision. Le test suppose que  $r_t$  et  $\phi(x_t, \hat{\theta})$  sont distribuées symétriquement par rapport à 0. Nous nous assurons de respecter cette condition en normalisant les variables, tel que défini en (9).

Chaque fois que le réseau  $(p,q)$  produit le bon (mauvais) signe, la fonction  $\mathbf{I}_t[.]$  prend la valeur un (zéro).  $SR$  correspond donc à la proportion de prévisions de signes exactes sur le nombre total de prévisions  $m$  calculées.

b) La Statistique de Pesaran et Timmermann (1992) :

Pesaran et Timmermann (1992) ont développé une statistique assez simple pour vérifier si la valeur de  $SR$  calculée en (13) diffère significativement de la valeur de  $SR$  qui aurait été obtenue si les processus  $r_t$  et  $\phi(x_t, \hat{\theta})$  étaient indépendants :

$$(14) \quad PT = \frac{(SR - SRI)}{\sqrt{\text{var}(SR) - \text{var}(SRI)}}$$

$SR$  est le taux de succès défini en (13) et  $SRI$  le taux de succès s'il y a indépendance entre  $r_t$  et  $\phi(\mathbf{x}_t, \hat{\theta})$ ,  $\text{var}(SR)$  et  $\text{var}(SRI)$  sont les variances respectives de  $SR$  et  $SRI$ . La dérivation de  $SRI$  et des variances respectives est présentée dans l'article de Pesaran et Timmermann (1992) de même que dans le volume de Franses et Van Dijk (2000) aux pages 43 et 44.

Pesaran et Timmerann démontrent que sous l'hypothèse nulle d'indépendance des deux processus, la statistique  $PT$  suivait asymptotiquement une distribution Normale centrée réduite.

Si la valeur de  $PT$  est positive et significative (à un seuil de 5 ou 10% par exemple), on peut conclure que le taux de succès ( $SR$ ) est significatif et le modèle  $(p, q)$  produit des prévisions de signes fiables. Inversement, si la statistique  $PT$  est négative et significative, les prévisions de signes ne sont pas fiables.

### c) L'Erreur Quadratique Moyenne :

Une mesure alternative utilisée pour évaluer la performance des réseaux de neurones est l'erreur de prévision quadratique moyenne (Mean Square Prediction Error):

$$(15) \quad MSPE = \frac{1}{m} \sum_{t=1}^m (r_t - \phi(\mathbf{x}_t, \hat{\theta}))^2$$

Le modèle qui produit l'erreur quadratique moyenne la plus faible est considéré comme meilleur. Une manière plus explicite de présenter cette mesure consiste à prendre le rapport entre le  $MSPE$  du modèle  $(p, q)$  étudié et le  $MSPE$  d'un modèle de référence, souvent la Marche Aléatoire, Gençay (1999).

Un ratio inférieur à 1 indique que le réseau  $(p, q)$  produit une erreur quadratique moyenne inférieure à l'erreur quadratique moyenne obtenue à partir de la Marche Aléatoire.

d) La Statistique de Diebold et Mariano (1995) :

La statistique de Diebold et Mariano (1995) permet de comparer des fonctions d'erreurs de prévisions de deux modèles concurrents : les réseaux de neurones  $(p, q)$  et un modèle de référence, la Marche Aléatoire.

Soit  $g_{1t} = (r_t - \phi(\mathbf{x}_t, \hat{\theta}))^2$  la fonction d'erreur de prévision quadratique (ou toute autre fonction de l'erreur de prévision, par exemple l'erreur absolue) du réseau  $(p, q)$  et  $g_{2t}$  la même fonction d'erreur de prévision de la Marche Aléatoire, la fonction du différentiel des erreurs de prévision est donnée par:

$$(16) \quad d_t = g_{1t}(\cdot) - g_{2t}(\cdot) \quad t = 1, \dots, m$$

Nous vérifions alors s'il existe une différence significative entre les deux fonctions d'erreurs de prévision  $g_{1t}$  et  $g_{2t}$  en testant si la moyenne du différentiel  $d_t$  est significativement différente de zéro. Sous l'hypothèse nulle, le différentiel moyen est égal à zéro,  $E[d_t] = 0$ , ce qui signifie qu'il n'y a pas de différences dans la précision des prévisions obtenues à partir des deux modèles concurrents.

Soit  $\mu(d_t)$  un estimateur de l'espérance de la fonction du différentiel  $d_t$  et  $\sigma(d_t)$  un estimateur de l'écart-type de cette même fonction, la statistique de Diebold et Mariano s'écrit alors :

$$(17) \quad DM = \frac{\mu(d_t)}{\sigma(d_t)}$$

Diebold et Mariano montrent que sous l'hypothèse nulle, cette statistique suit asymptotiquement une loi de distribution Normale centrée réduite.

Diebold et Mariano utilisent un estimateur de la variance asymptotique du différentiel moyen comme dénominateur pour tenir compte de la dépendance des prévisions. En effet, pour un horizon de prévision de  $k$  périodes, les erreurs des prévisions seront au maximum  $k - 1$  dépendantes (Diebold et Mariano,

1995). Sous l'hypothèse d'indépendance, nous utilisons l'écart-type  $\sigma(d_t)$ . En effet, nous ne générons que des prévisions à une période, avec  $k = 1$  jour et les erreurs de prévisions sont indépendantes.

Une statistique  $DM$  significativement différente de 0 (aux seuils de 5 ou 10% par exemple) implique l'existence d'une différence dans la précision des prévisions des deux modèles comparés. Si  $DM$  est significativement positive, le réseau  $(p,q)$  est plus performant que le modèle de référence (ici la Marche Aléatoire) alors qu'une valeur de  $DM$  significative et négative implique que le réseau  $(p,q)$  est moins performant que la Marche Aléatoire. Si l'hypothèse nulle  $E[d_t] = 0$  n'est pas rejetée, le réseau de neurones  $(p,q)$  est comparable à la Marche Aléatoire en terme de précision des prévisions, il n'est ni plus ni moins performant que le modèle de référence auquel nous le comparons au niveau du test utilisé.

Les éléments de modélisation des réseaux de neurones ont été présentés dans cette section. Nous avons aussi discuté des difficultés liées à l'emploi de techniques non-linéaires d'optimisation et avons proposé différentes méthodes pour améliorer les propriétés numériques de notre algorithme. Nous avons également élaboré une technique de sélection pour choisir les meilleurs modèles parmi tous ceux qui ont été estimés et avons décrit les mesures nécessaires pour évaluer la performance prédictive de ces modèles.

Trois versions différentes de réseaux de neurones sont étudiées dans ce rapport. Nous voulons déterminer s'il est possible d'améliorer la performance de ces réseaux quand différentes variables explicatives sont employées comme inputs. Avant de discuter des différents modèles de taux de change que nous étudions dans ce rapport, nous résumons à la section suivante quelques articles qui ont été consultés pour notre recherche. Les méthodologies et résultats de ces recherches ont largement influencé notre approche empirique.

### **III. Les Études Antérieures :**

Les limites des techniques linéaires dans la prévision des mouvements des données financières ont amené un nombre important de chercheurs à exploiter des méthodes alternatives. Parmi les techniques non-linéaires, les modèles non-paramétriques s'avèrent plus populaires car ils demandent un minimum d'hypothèses et produisent des résultats assez intéressants comme en témoignent les publications récentes sur le sujet.

Les réseaux de neurones, par leur flexibilité et simplicité, deviennent un outil de choix dans la prévision des variables financières. Les auteurs des différentes recherches présentées plus loin exploitent tous des variantes de ces réseaux pour bâtir des modèles de prévisions des rendements des taux de change quotidiens.

Les premiers chercheurs utilisent les réalisations passées de la variable d'intérêt comme inputs de leurs modèles avec plus ou moins de succès. En effet, les réseaux entraînés permettent d'obtenir des prévisions de signes intéressantes et arrivent dans certains cas à faire mieux que la Marche Aléatoire (sous l'hypothèse nulle de l'efficacité des marchés, la marche aléatoire est le modèle de référence car les prix reflètent toute l'information disponible sur le marché et il n'est donc pas possible d'en prévoir les mouvements futurs).

Le deuxième groupe d'auteurs concentre ses efforts dans la recherche de nouvelles relations expliquant les mouvements à court terme des taux de change. Pour prévoir le sens des variations futures, des fonctions des taux passés (comparaison entre des moyennes mobiles de longueurs différentes) sont utilisées pour certains modèles alors que d'autres exploitent des variables reflétant le volume actuel des transactions des devises sur les marchés des changes. Les réseaux qui en découlent sont plus performants. Ils prévoient correctement la direction des rendements futurs dans des proportions plus élevées et produisent des erreurs quadratiques moyennes inférieures ou égales aux erreurs quadratiques moyennes de la Marche Aléatoire.

1) Kuan et Liu (1995) :

Les rendements passés sont utilisés comme inputs des réseaux de neurones pour générer des prévisions des rendements quotidiens de cinq séries de taux de change par rapport au dollar américain : la Livre anglaise, le dollar canadien, le Mark allemand, le Yen et le franc suisse. Les auteurs procèdent à un exercice de validation pour sélectionner les trois modèles  $(p,q)$  les plus performants (les estimateurs qui produisent l'erreur quadratique moyenne la plus faible sont retenus pour l'exercice de prévision) par taux de change étudié. Pour chacun des réseaux ainsi choisis, ils génèrent des prévisions à un jour pour trois échantillons de prévision de tailles différentes qu'ils comparent aux valeurs observées. Pour évaluer les réseaux, Kuan et Liu calculent les taux de succès (*SR*) et la statistique de Pesaran et Timmermann (*PT*). Ils présentent également les écarts quadratiques de prévision moyens (Mean Square Prediction Error *MSPE*) du réseau  $(p,q)$  et les comparent aux écarts obtenus par la Marche Aléatoire.

Certains réseaux produisent des prévisions de signes significatives alors que d'autres livrent des *MSPE* inférieures aux *MSPE* de la Marche Aléatoire. Le réseau  $(p,q)$  sélectionné pour un taux de change donné génère des résultats différents selon l'horizon de prévision considéré et un modèle  $(p,q)$  donné se comporte différemment pour chacun des taux de change pour lequel il a été sélectionné.

2) Franses et Van Homelen (1998) :

Franses et Van Homelen tentent aussi de construire des modèles de prévision des taux de change en utilisant les retards des rendements quotidiens comme inputs de leurs réseaux de neurones. Les paramètres ne sont estimés qu'une fois et sont utilisés pour calculer des prévisions à un jour pour quatre séries de taux de change quotidiens par rapport au Florin : le dollar américain, le dollar canadien, la Livre anglaise et le Yen.

Une technique de validation est utilisée pour minimiser l'impact du choix de l'échantillon d'estimation sur la performance prédictive des réseaux considérés. Les mesures rapportées par les auteurs sont le taux de succès (*SR*) et la statistique de Pesaran et Timmermann (*PT*). Ils calculent également les critères de sélection de Akaike (*AIC*) et Schwarz (*BIC*) pour l'échantillon d'estimation afin de vérifier s'ils facilitent le choix du réseau ( $p, q$ ) qui produit les prévisions les plus fiables.

Pour le dollar américain, les valeurs de *PT* qui sont significatives sont toutes négatives et les *SR* sont rarement supérieurs à 50%. Pour le dollar canadien, aucune des statistiques de *PT* obtenues n'est significative bien que les réseaux considérés donnent des taux de succès plus élevés. La conclusion est sensiblement la même pour la Livre anglaise. Pour le Yen, les auteurs rapportent quelques cas où *PT* est significativement positive avec des taux de succès supérieurs à 50%.

Enfin, les critères de Akaike et Schwarz ne permettent pas de sélectionner les modèles les plus performants.

### 3) Franses et Van Griensven (1998) :

Franses et Van Griensven (1998) se tournent vers l'analyse technique pour tenter de bâtir de meilleurs modèles de prévision des rendements de cinq séries de taux de change quotidiens par rapport à la Livre anglaise : le Mark allemand, le dollar américain, le Florin, le Yen et le franc suisse. La différence entre des moyennes mobiles de court terme et des moyennes mobiles de long terme est introduite comme variable explicative de leur modèle. Une différence positive indique un mouvement à la hausse du cours de la devise étudiée (rendement prévu positif) et une différence négative un mouvement à la baisse (rendement prévu négatif). Ils fixent la longueur ( $n_1$ ) de la moyenne mobile de court terme à 1 jour et la longueur ( $n_2$ ) des moyennes mobiles de long terme à 50 et 150 jours. Les prévisions à un jour sont obtenues pour deux échantillons, 1994 et 1995 alors que les paramètres ne sont estimés qu'une fois en utilisant les observations



de 1981 à 1993. Pour évaluer la performance des réseaux  $(p, q, n_1, n_2)$ , Franses et Van Griensven utilisent le taux de succès ( $SR$ ) et la statistique de Pesaran et Timmermann ( $PT$ ).

Les modèles qui utilisent une moyenne mobile de long terme de longueur  $n_2 = 150$  produisent clairement les meilleurs résultats pour les échantillons de prévision. En général, les auteurs obtiennent des valeurs de  $PT$  significatives pour au moins deux séries peu importe le nombre de nœuds  $q$  considéré : le Yen et le franc suisse pour 1994, le Mark allemand et le Florin pour 1995. Quand  $n_2$  est fixé à 50 jours, les réseaux sont moins performants. Pour 1994, le franc suisse est la seule devise pour laquelle la statistique  $PT$  est significative. Pour 1995, les auteurs obtiennent une statistique significative pour le dollar américain pour les cinq valeurs de  $q$  utilisées.

#### 4) Gençay (1999) :

L'article de Gençay se veut une étude comparative des différents modèles (paramétriques et non-paramétriques) qui sont couramment utilisés dans la prévision des taux de change. Six modèles sont estimés pour tenter de prévoir les rendements futurs de cinq séries de taux de change quotidiens par rapport au dollar américain : la Livre anglaise, le Mark allemand, le franc français, le Yen et le franc suisse. Les modèles comparés sont la Marche Aléatoire, le GARCH(1,1), le "Nearest Neighbor", le réseau de neurones avec rendements passés et les réseaux de neurones avec différences entre moyennes mobiles de court terme de longueur  $n_1 = 1$  et moyennes mobiles de long terme de longueur  $n_2 = 50$  et 200 jours.

L'auteur procède à un exercice de validation croisée pour déterminer la structure optimale des réseaux de neurones (celle qui donne les erreurs quadratiques moyennes minimales). Une fois cette combinaison optimale déterminée et les paramètres du modèle estimés, des prévisions à un jour sont générées et comparées aux valeurs observées. Les mesures de performance telles le ratio de l'erreur quadratique moyenne de prévision ( $MSPE$ ) du modèle concerné

relativement à l'erreur quadratique de la Marche Aléatoire, le taux de succès (*SR*) et la statistique de Diebold et Mariano (*DM*) sont rapportées.

Les méthodes non-paramétriques, le "*Nearest Neighbor*" et les réseaux de neurones sont plus fiables que la Marche Aléatoire ou le GARCH(1,1), en termes de prévisions de signe (*SR* plus élevés) et prévisions ponctuelles (moyennes des erreurs au carré inférieures). Parmi ces modèles non-paramétriques, la méthode locale, le "*Nearest Neighbor*" produit des prévisions plus précises que les réseaux de neurones. Les réseaux de neurones qui utilisent les règles d'analyse technique produisent des prévisions plus précises que le réseau qui utilise les retards des rendements comme inputs.

##### 5) Gradojevic et Yang (2000) :

Gradojevic et Yang introduisent de nouvelles variables dans leur réseau de neurones pour essayer d'obtenir des modèles de prévisions de taux de change plus précis. Les agents économiques qui transigent sur le marché des changes (les banques, les fonds mutuels, les fonds de pension et autres) analysent tout l'ensemble d'information disponible, incluant les données fondamentales, avant de décider de la position à prendre sur le marché. Les cambistes observent cette demande de devises extraite des commandes placées et fixent les prix des devises en conséquence. Une quantité importante de commandes pour une devise (i.e une forte demande) devrait se traduire par une appréciation du cours de la devise et par conséquent un rendement positif. Les auteurs utilisent donc cette variable de flux (de même que des variables macro-économiques) comme input des réseaux de neurones pour tenter de prévoir les mouvements futurs du taux de change quotidien Canada/États-Unis.

Pour s'assurer de ne pas surentraîner leurs modèles, Gradojevic et Yang utilisent la technique du "*Early Stopping*" qui consiste à arrêter l'entraînement des réseaux dès que la fonction d'erreur de l'échantillon de validation commence à augmenter. Ils font appel à deux critères pour mesurer la performance des réseaux : la racine carrée de l'erreur quadratique moyenne (*RMSE*) et le taux de

succès ( $SR$ ). Les prévisions des différents réseaux de neurones sont comparées aux prévisions générées à partir d'une Marche Aléatoire et deux modèles linéaires où les inputs sont utilisés comme régresseurs. Les taux de succès obtenus par les réseaux sont toujours supérieurs aux taux de succès des modèles concurrents et les  $RMSE$  constamment inférieures ou égales aux  $RMSE$  des modèles concurrents.

Les résultats des différentes études présentées ci-haut viennent confirmer que les réseaux de neurones sont susceptibles de mettre à profit les non-linéarités présentes dans les processus des rendements de taux de change pour obtenir des prévisions plus précises (i.e. des taux de succès plus élevés et des erreurs quadratiques moyennes inférieures). L'introduction de nouvelles variables explicatives permet d'améliorer la performance prédictive de ces techniques comme l'indiquent les conclusions de Franses et Van Griensven (1998) ou Gençay (1999). Les résultats obtenus restent cependant sensibles à la série de taux de change étudiée et à la structure considérée pour le réseau. Un réseau donné peut produire des résultats intéressants pour une série de taux de change mais pas pour l'autre. Le nombre de nœuds cachés  $q$  ne semble pas avoir beaucoup d'impact sur les résultats. Pour un échantillon de prévision et une variable explicative donnés, des valeurs significatives de  $PT$  peuvent être obtenues pour les mêmes séries quelle que soit la valeur de  $q$  considérée. Ces résultats établissent l'importance de développer des techniques de validation pour bâtir des réseaux appropriés à la variable étudiée. Il est nécessaire de mettre en place une stratégie de sélection qui facilitera le choix de la structure du réseau à utiliser pour une série de taux de change. Kuan et Liu (1995) et Gençay (1999) par exemple, utilisent des fonction d'erreurs quadratiques moyennes pour sélectionner la combinaison  $(p,q)$  optimale. Les réseaux  $(p,q)$  qui produisent les erreurs quadratiques moyennes les plus petites pour l'échantillon d'estimation sont retenus pour l'étape de prévision.

Nous proposons une technique de sélection basée sur des critères d'information tels le critère de Akaike et de Schwarz appliquée sur l'échantillon de validation et

voulons vérifier s'il s'agit là d'une méthode fiable pour sélectionner les réseaux  $(p,q)$  qui produisent des prévisions plus précises des rendements futurs des taux de change quotidiens. Samson et White (1995) appliquent le critère de Schwarz à l'échantillon d'estimation et n'obtiennent pas de résultats concluants.

Nous présentons à la section suivante les trois versions différentes des réseaux de neurones que nous entraînons et validons dans le but de bâtir ces modèles de prévision. Nous discutons des données utilisées pour évaluer nos trois modèles à la section V et nos résultats empiriques sont traités à la section VI.

#### IV. Les déterminants des Taux de Change :

Soit  $P_t$ ,  $t = 0, \dots, T$  la série de taux de change quotidiens observés, exprimés en unités de devises étrangères pour un dollar américain.

La série des rendements est alors définie comme :

$$(17) \quad r_t = \ln(P_t) - \ln(P_{t-1}) \quad t = 1, \dots, T$$

Trois versions différentes du réseau de neurones sont étudiées dans ce rapport. Le but est de trouver l'estimateur de la moyenne conditionnelle qui génère les prévisions les plus fiables.

##### 1) Les rendements passés :

Dans un contexte dynamique, il est commun d'utiliser des modèles auto-régressifs pour expliquer le comportement de la variable dépendante. Il semble alors justifié d'inclure les réalisations passées des rendements comme inputs du réseau de neurones :

$$(18) \quad r_t = \phi(r_{t-p}, \theta) + \varepsilon_t \quad t = p+1, \dots, T$$

$\varepsilon_t$  est indépendamment distribué de moyenne 0 et de variance  $\sigma_t^2$ ,

$\phi(\cdot)$  est la fonction qui relie  $r_t$  aux inputs – le réseau de neurones avec composante linéaire décrit en (5) –  $p$  représente l'ordre des retards,  $q$  le nombre de nœuds de la couche cachée et  $\theta$  le vecteur des paramètres du modèle.

Les premières recherches sur les réseaux de neurones tentent d'exploiter cet aspect dynamique qui caractérise la relation entre la variable d'intérêt et ses réalisations passées pour bâtir des modèles de prévision des taux de change plus ou moins performants : Kuan et White (1994), Kuan et Liu (1995), Franses et Van Homelen (1998).

## 2) L'analyse technique :

Les règles techniques sont largement utilisées par les analystes dans leurs décisions d'achat et de ventes de différents actifs financiers. Ces règles permettent de déterminer des changements de tendance dans l'évolution des prix d'un actif en étudiant leurs comportements passés.

Soit  $m_t(n)$  la moyenne mobile d'ordre  $n$  :

$$(19) \quad m_t(n) = \frac{1}{n} \sum_{i=0}^{n-1} P_{t-i}$$

La règle consiste à comparer une moyenne mobile de court terme à une moyenne mobile de long terme. Un signal d'achat (de vente) est donné quand la différence entre la moyenne mobile de court terme et la moyenne mobile de long terme est positive (négative). Ce signal est donné par :

$$(20) \quad s_t(n_1, n_2) = m_t(n_1) - m_t(n_2) \quad n_1 < n_2$$

Pour les réseaux étudiés dans ce rapport  $n_1$  est fixé à 1 jour et  $n_2$  à 50 et 200 jours ( $n_1 = 1$  correspond en fait à la série de taux de change observés  $P_t$ ).

Le graphique 2 à la page i de l'annexe présente les moyennes mobiles de 50 et 200 jours pour le taux du dollar canadien par rapport au dollar américain. Chaque fois que la courbe de la moyenne mobile de court terme (ou la courbe des taux de change observés  $P_t$ ) coupe la courbe de moyenne mobile  $m_t$  de long terme vers le haut (différence positive), un signal d'achat est donné et inversement, si  $P_t$  coupe  $m_t$  vers le bas (différence négative), un signal de vente est donné.

De la même manière que Franses et Van Griensven (1998) ou Gençay (1999), la variable définie en (20) est utilisée comme input dans la deuxième version du réseau de neurones.

Le modèle devient :

$$(21) \quad r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t \quad t = p+1, \dots, T$$

Si un signal d'achat est donné en  $t-1$ ,  $s_{t-1}(n_1, n_2) > 0$ , nous devrions nous attendre à un rendement positif en  $t$ ,  $r_t > 0$ , et si un signal de vente est donné en  $t-1$ ,  $s_{t-1}(n_1, n_2) < 0$ , nous pouvons anticiper un rendement négatif en  $t$ ,  $r_t < 0$ . Un signal positif signifie une moyenne mobile de court terme supérieure à la moyenne mobile de long terme donc un changement à la hausse dans l'évolution des cours de la devise. Un signal négatif implique une moyenne mobile de court terme inférieure à la moyenne mobile de long terme donc une baisse anticipée du cours de la devise. La courbe de moyenne mobile de 200 jours est plus lisse et produit donc moins de signaux de vente-achat (voir le graphique 2 à la page xi de l'annexe pour l'exemple du taux du dollar canadien par rapport au dollar américain).

### 3) Le différentiel de taux d'intérêt :

Les deux premiers modèles que nous étudions, les réseaux de neurones avec retards des rendements ou avec signaux de vente-achat, exploitent tous le même ensemble d'information, les réalisations passées de la variable d'intérêt. Dans les applications financières, il peut s'avérer intéressant d'utiliser d'autres variables pouvant expliquer les variations des taux de change.

Les investisseurs comparent constamment les rendements qu'ils peuvent obtenir s'ils détiennent leurs fonds en une devise donnée plutôt qu'une autre. Si les taux d'intérêt américains (par exemple) sont supérieurs aux taux étrangers (toutes choses étant égales par ailleurs), les investisseurs sont susceptibles de diminuer leurs avoirs en devises étrangères en faveur du dollar américain. Cette augmentation de la demande de dollars américains entraîne une appréciation du taux de change du dollar américain par rapport aux autres monnaies (exprimé en

nombre d'unités de devises étrangères pour acheter un dollar américain). Nous pouvons anticiper alors un rendement positif la période suivante. Inversement, une diminution des taux d'intérêt américains par rapport aux taux étrangers incite les investisseurs à liquider leurs actifs en dollars américains pour se placer sur les devises étrangères. Cette demande accrue de devises étrangères a pour effet de réduire le taux de change du dollar américain par rapport aux autres devises (toujours exprimé en unités de devises étrangères pour un dollar américain) et le rendement à la période suivante devrait être négatif. Le différentiel de taux d'intérêt devrait donc expliquer (en partie) les mouvements à court terme des taux de change. Soit  $R_t$  le taux d'intérêt américain et  $R_{t,i}$  le taux d'intérêt du pays  $i$ , le différentiel des taux d'intérêt est donné par :

$$(22) \quad D_t = R_{t,i} - R_t$$

Un écart  $D_t$  positif signifie un taux d'intérêt américain  $R_t$  inférieur aux taux étrangers  $R_{t,i}$  et devrait entraîner une augmentation de la demande de monnaies du pays  $i$ , donc un rendement  $r_{t+1}$  négatif à la période suivante alors qu'un écart négatif implique un taux d'intérêt américain supérieur aux taux étrangers, une demande accrue de dollars américains et un rendement  $r_{t+1}$  positif.

Bien que la théorie de la parité non-couverte des taux d'intérêt (*Uncovered Interest Rate Parity*) décrite ici implique l'introduction du différentiel de taux d'intérêt en niveau, décrit en (22), nous préférons utiliser les premières différences de  $D_t$  comme relation empirique. Cela permet d'éviter les problèmes de non-stationnarité des données sur le différentiel (voir les graphiques 7 à 10 aux pages xiv à xvii de l'annexe) et se rapproche de la démarche suivie par Gradojevic et Yang (2000) qui utilisent toutes les variables de leurs modèles en premières différences car ils sont intéressés à prévoir le sens des changements quotidiens des taux de change et non pas le niveau des taux de change.



La variation quotidienne du différentiel du taux intérêt est définie par :

$$(23) \quad \Delta D_t = D_t - D_{t-1}$$

La troisième version du réseau de neurones étudiée utilise donc la variation quotidienne (la première différence) de l'écart entre taux intérêt nominaux de court terme comme input :

$$(24) \quad r_t = \phi(\Delta D_{t-p}, \theta) + \varepsilon_t \quad t = p+1, \dots, T$$

Il est possible de considérer combiner les trois types de variables décrites ci-haut dans un modèle de réseau de neurones. Gradojevic et Yang (2000) par exemple incluent des variables macro et micro-économiques variées pour tenter de prévoir les mouvements futurs du taux de change du dollar canadien par rapport au dollar américain. Ils introduisent de la sorte les différentiels de taux d'intérêt à court terme de même que le prix du pétrole brut et la variable micro-structurelle qui décrit les flux de transactions sur le marché des changes. Les réseaux de neurones ainsi obtenus font mieux que les divers modèles paramétriques concurrents.

Nous voulons vérifier ici si les réseaux de neurones produisent des modèles de prévision raisonnables en comparant leur performance selon l'input utilisé, les rendements passés, les règles d'analyse technique ou les différentiels de taux d'intérêt.

## V. Les Données :

Nous étudions dans ce rapport quatre séries de taux de changes quotidiens (exprimés en unités de devises étrangères pour un dollar américain) : le dollar canadien, le franc français, le Mark allemand et la Livre anglaise. Les données proviennent du site internet de "The Federal Reserve Bank of St-Louis" à l'adresse <http://www.stls.frb.org>.

Des observations quotidiennes sur les taux LIBOR à 3 mois du dollar américain, dollar canadien, franc français, mark allemand et livre anglaise sont utilisées pour calculer le différentiel de taux intérêt nominaux de court terme pour la troisième version du réseau de neurones définie à l'équation (24). Elles sont obtenues sur le site internet de la "British Banks Association" à l'adresse <http://www.bba.org.uk>. Le Taux LIBOR est le taux à court terme le plus communément utilisé par les banques et investisseurs sur les marchés de capitaux à travers le monde. Le taux LIBOR est fixé pour la majorité des grandes devises étrangères et est utilisé pour calculer les taux d'intérêt qui s'appliquent à une grande variété de contrats financiers (swaps, émissions euro-obligataires, etc...).

L'échantillon total comprend 1969 données quotidiennes qui s'étendent du 2 janvier 1991 au 31 décembre 1998. Cet échantillon est divisé en trois sous-groupes : un échantillon d'estimation (50% de l'échantillon total) qui s'étend de janvier 1991 à décembre 1994, un échantillon de validation (25% de l'échantillon total) qui s'étend de janvier 1995 à décembre 1996 et un échantillon de prévision (25% de l'échantillon total) qui s'étend de janvier 1997 à décembre 1998.

Le choix de la fréquence des observations pour les échantillons est justifié par l'utilisation de données quotidiennes par les auteurs des études consultées sur le sujet qui sont présentées à la section III. Dans un souci de continuité et pour faciliter la comparaison des résultats nous utilisons aussi des données quotidiennes.

Des statistiques descriptives sur les données sont présentées en annexe au tableau I, page xviii de l'annexe, pour les rendements quotidiens de taux de

change et au tableau II, à la page xix, pour les premières différences du différentiel des taux LIBOR à 3 mois pour les taux de change concernés. Les statistiques ont été calculées pour les quatre échantillons considérés : l'échantillon total, l'échantillon d'estimation, l'échantillon de validation et échantillon de prévision.

Les graphiques 3 à 6, pages x à xiii, présentent les niveaux et les rendements quotidiens des séries de taux de change alors que les graphiques 7 à 10, pages xiv à xvii, présentent les différentiels des taux LIBOR de même que les premières différences de ceux-ci.

#### 1) Les rendements quotidiens des taux de change :

De manière générale, les rendements quotidiens démontrent un comportement commun à toute série de données financières à haute fréquence.

La moyenne de ces séries de rendements est pratiquement nulle. L'estimateur du troisième moment (*Skewness*) montre un léger degré d'asymétrie dans la distribution de ces séries, les valeurs sont proches de zéro mais ne sont pas nulles pour toutes les séries et échantillons. L'indice d'aplatissement (*Kurtosis*) indique que la distribution des rendements quotidiens a plus de masse dans les queues relativement à la distribution Normale. Le facteur de *Kurtosis* de la distribution Normale est égal à 3, l'indice d'aplatissement présenté au tableau 1 (*Excess Kurtosis*) correspond au facteur d'aplatissement calculé pour chacune de nos séries de rendements moins 3.

Les séries de rendements semblent être toutes stationnaires comme en témoignent les graphiques des rendements. Il faut noter cependant que les rendements des taux de change du franc français, du Mark allemand et de la Livre anglaise démontrent beaucoup plus de variations tout au long de la période considérée. Fait consolidé par les valeurs des écarts-types de ces séries qui sont de l'ordre de 0.60% alors que l'écart-type du dollar canadien est de 0.30% environ.

## 2) Les changements quotidiens des différentiels des taux LIBOR :

Contrairement aux séries de rendements décrites ci-haut, le comportement des séries des premières différences des différentiels de taux LIBOR 3 mois n'est pas similaire au comportement d'autres séries de données financières à fréquence élevée.

Bien que les moyennes soient proches de zéro, les écarts-types sont bien plus élevés et ne sont pas constants à travers les différents sous-échantillons. Pour tous les différentiels, l'écart-type est plus élevé pour la période s'étendant de janvier 1991 à décembre 1994. Cet aspect est visible graphiquement, pages xiii à xvi de l'annexe. Les premières différences connaissent toutes (exception faite du différentiel entre les taux LIBOR allemands et américains) des variations importantes jusqu'à la fin 1992 et début 1993 alors que les mouvements sont relativement nuls par la suite.

Tandis que la distribution des rendements quotidiens ne manifeste pas des degrés d'asymétrie importants, les différentiels des taux LIBOR présentent tous des valeurs pour *Skewness* significativement différentes de zéro. Nous devons également noter que les indices d'aplatissement de ces séries sont toujours largement supérieurs à 3. Les distributions de ces séries de différentiels semblent avoir une masse significative dans les queues de leur distribution par rapport à la distribution Normale.

La représentation des différentiels en niveau montre clairement des signes de non-stationnarité des données alors que les premières différences sont stationnaires tout au long de l'échantillon considéré. Cet élément vient renforcer le choix des premières différences de l'écart entre les taux LIBOR comme inputs du modèle décrit en (24) plutôt que le différentiel en niveau comme le suppose la théorie de la parité des taux d'intérêt.

## VI. Les Résultats Empiriques :

Nous examinons ici la performance hors-échantillon des différents estimateurs de la moyenne conditionnelle utilisés pour prévoir les rendements quotidiens des taux de change. Nous avons entraîné quatre versions de réseaux de neurones: les réseaux avec retards des rendements, les réseaux avec règles d'analyse technique ( $n_1 = 1, n_2 = 50$ ) et ( $n_1 = 1, n_2 = 200$ ) et les réseaux avec changements quotidiens des différentiels de taux LIBOR 3 mois. Les prévisions sont calculées pour le dernier quart de l'échantillon total, les années 1997-1998.

Les différentes mesures et statistiques sont présentés en annexe (pages xx à xxxv) aux tableaux III à VI pour le dollar canadien, VII à X pour le franc français, XI à XIV pour le Mark allemand et XV à XVIII pour la Livre anglaise.

### 1) Le dollar canadien :

Pour le dollar canadien, nous obtenons des taux de succès intéressants peu importe le modèle considéré. En effet, que ce soit le réseau qui utilise les rendements passés, les règles d'analyses techniques ou celui qui utilise les différentiels de taux intérêt comme inputs, les taux de succès *SR* sont généralement supérieurs à 50 et même 55%.

La statistique de Pesaran et Timmermann *PT* permet de vérifier si nos modèles prévoient correctement le sens des variations futures des taux de change de manière significative. Pour la première version, avec les rendements passés, une seule statistique *PT* est positive et significative à un seuil de 5%, le réseau (3,3) avec un *SR* égal à 56%. Parmi les modèles qui font appel à l'analyse technique, celui qui utilise la règle ( $n_1 = 1, n_2 = 50$ ) est plus fiable que celui qui utilise la règle ( $n_1 = 1, n_2 = 200$ ). Bien que les deux réseaux produisent des taux de succès relativement élevés, seulement le premier le fait de manière significative, les réseaux (2,1) et (2,2) à un seuil de 5% et les réseaux (3,1), (3,2) et (2,3) à un seuil de 10%, tous avec des taux de succès de l'ordre de 55 et 56%. Le réseau de neurones qui utilise les changements des différentiels des taux

LIBOR obtient le nombre le plus élevé de statistiques *PT* positives et significatives, les réseaux (5,1), (1,2) et (2,3) à un seuil de 5% et les réseaux (3,2), (4,2) et (5,2) à un seuil de 10%, ici aussi avec des valeurs de *SR* supérieures ou égales à 55%.

Les réseaux de neurones produisent en général des erreurs quadratiques moyennes *MSPE* inférieures à celles de la Marche Aléatoire, le modèle qui exploite la règle d'analyse technique la plus courte (50 jours) étant le plus performant. Les statistiques de Diebold et Mariano *DM* permettent de discriminer entre deux modèles en mesurant la précision des prévisions obtenues par chacun des modèles étudiés. Les valeurs de *DM* sont rarement négatives et significatives et une fois positive et significative (pour le réseau à différentiels de taux intérêt). Les réseaux de neurones ne sont ni pires ni meilleurs que la Marche Aléatoire en termes de précision des prévisions obtenues.

Enfin, les critères de Akaike *AIC* et de Schwarz *BIC* ne permettent pas de sélectionner/valider les modèles les plus performants et ce pour les quatre versions des réseaux entraînés. Aucun des critères, qu'il soit calculé à l'étape d'estimation ou à l'étape de validation, ne sélectionne un des réseaux (*p,q*) produisant les prévisions les plus précises.

## 2) Le franc français :

Pour le franc français, les taux de succès obtenus sont inférieurs, rarement de l'ordre de 50%. Les réseaux de neurones à rendements passés produisant peut-être les meilleures valeurs pour *SR* et les modèles employant la règle technique à 200 jours les valeurs les plus petites. Les seules statistiques *PT* significatives sont négatives et correspondent en général aux modèles avec des taux de succès les plus bas, 46%.

En terme de moyenne des erreurs quadratiques, la seule version qui réussit à obtenir des valeurs comparables à celles de la Marche Aléatoire utilise les différentiels des taux LIBOR, avec des ratios de *MSPE* proches de l'unité. Les

valeurs des statistiques  $DM$  sont toutes (ou presque) négatives et la majorité sont significatives à un seuil de 5%. Seul le réseau qui introduit les réalisations passées des rendements comme inputs donne des statistiques  $DM$  non significatives. La Marche Aléatoire est en somme un meilleur modèle de prévision pour la série des rendements quotidiens du franc français comparé aux réseaux de neurones entraînés.

Les critères de sélection  $AIC$  et  $BIC$  permettent de choisir le meilleur modèle de prévision pour le réseau (1,1) pour la version avec rendements passés. Il faut noter cependant que ces deux critères sélectionnent de manière systématique le réseau (1,1) à l'étape de validation.

### 3) Le Mark allemand :

Pour le Mark allemand, les réseaux de neurones avec rendements passés et différentiels de taux LIBOR semblent être plus performants. Les premiers atteignent des taux de succès de 51% alors que les deuxièmes des taux de 53%. L'analyse technique permet de prévoir les signes des rendements futurs uniquement pour la règle à 50 jours. Cette règle permet d'obtenir des valeurs de  $SR$  qui atteignent 51% pour quelques réseaux  $(p,q)$ , tandis que pour la règle à 200 jours les  $SR$  ne dépassent pas 48%. Les statistiques  $PT$  ne sont pas significatives pour les quatre versions considérées (rendements passés, analyse technique et différentiels de taux d'intérêt) et toutes les paires  $(p,q)$  entraînés. Elles sont proches de zéro pour les réseaux qui utilisent les différentiels de taux LIBOR. Ces réseaux obtiennent les erreurs quadratiques moyennes les plus petites. En effet, il s'agit de la seule version avec des ratios  $MSPE$  constamment inférieurs à l'unité et des statistiques de  $DM$  positives même si elles ne sont pas significatives.

$AIC$  et  $BIC$  choisissent ici aussi les réseaux (1,1) tant pour l'estimation que pour la validation, quel que soit l'input utilisé. Nous ne pouvons conclure, ici non plus, à l'utilité de ces critères en tant qu'outils de sélection. En effet, les trois premières versions du réseau de neurones ne produisent pas de prévisions

fiables des rendements futurs alors que la dernière (différentiels de taux intérêt) permet de prévoir correctement les signes 52% des fois et est plus efficace que la Marche Aléatoire, mais cela est vrai pour plusieurs combinaisons  $(p,q)$ , pas le réseau (1,1) uniquement. Les critères ne facilitent pas dans le choix du modèle à utiliser à l'étape de prévision.

#### 4) La Livre anglaise :

Les réseaux de neurones introduisant les rendements passés comme inputs prévoient correctement les signes des rendements futurs de la Livre anglaise dans des proportions de 50 à 53%. Pour quatre de ces modèles, les valeurs de  $PT$  sont positives et significatives à un seuil de 10%, les réseaux (2,1), (5,1), (2,3) et (5,3). L'utilisation de variables empruntées à l'analyse technique n'améliore pas considérablement la performance des réseaux en terme de prévision de signes. La règle qui utilise la moyenne mobile de long terme de 50 jours permet d'obtenir des taux de succès de 51 ou 52% et des  $PT$  positives et significatives au seuil de 10% pour les réseaux (3,1) et (2,2). La règle qui utilise une moyenne mobile de long terme de 200 jours réussit à prévoir le bon signe 53% des fois pour le réseau (4,1), il s'agit du seul réseau qui obtient une valeur de  $PT$  positive et significative à un seuil de 5% parmi toutes les versions des réseaux  $(p,q)$  entraînés. Le différentiel de taux intérêt nominaux ne produit pas de résultats concluants, bien que certains des réseaux  $(p,q)$  estimés donnent des  $SR$  comparables aux modèles précédents. Deux d'entre eux produisent des statistiques  $PT$  significatives, les réseaux (3,1) et (3,3).

Les réseaux qui emploient les rendements passés et ceux qui ont recours à la règle technique ( $n_1 = 1, n_2 = 200$ ) donnent des erreurs quadratiques moyennes intéressantes (i.e. inférieures à celles de la Marche Aléatoire). Les réseaux qui utilisent les différentiels des taux LIBOR obtiennent des valeurs de  $DM$  négatives et significatives pour tous les couples  $(p,q)$ .

L'utilisation des critères de Akaike et Schwarz ne permet pas de choisir les modèles qui génèrent les prévisions les plus fiables.  $AIC$  et  $BIC$  sélectionnent



les réseaux (1,1) pour les quatre modèles à l'étape de validation. Les réseaux sélectionnés au moment de l'estimation ne correspondent pas non plus aux modèles  $(p,q)$  qui donnent les meilleures prévisions de signe ou les erreurs quadratiques moyennes les plus petites.

La performance des réseaux varie selon la série de taux de change étudiée. Alors que le réseau de neurones qui utilise premières différences des écarts des taux LIBOR à 3 mois semble être le modèle le plus performant pour le dollar canadien et le Mark allemand, l'analyse technique avec  $n_2 = 200$  produit les meilleurs résultats pour la Livre anglaise. Aucun modèle n'est cependant capable de fournir des prévisions fiables pour les rendements quotidiens du franc français si ce n'est la Marche Aléatoire

Bien que certains modèles prévoient correctement le sens des variations futures des taux de change, ils n'arrivent pas constamment à produire des prévisions plus précises (*MSPE* inférieures) que celles qui sont obtenues par une Marche Aléatoire. Pour le dollar canadien par exemple, les réseaux avec différentiels de taux intérêt nominaux permettent d'améliorer les prévisions de signe relativement aux réseaux avec rendements passés mais ne peuvent produire des *MSPE* inférieures à celles qui sont obtenues par les réseaux qui utilisent la règle d'analyse technique ( $n_1 = 1, n_2 = 50$ ).

Les résultats montrent aussi que le nombre de nœuds  $q$  de la couche cachée a moins d'impact sur la précision des prévisions obtenues que le choix de la variable explicative. En effet, pour un modèle (rendements passés, analyse technique ou différentiels de taux LIBOR) et un taux de change donnés, les taux de succès obtenus sont comparables quel que soit le nombre de nœuds  $q$  considéré. Les résultats sont plus sensibles au choix de l'input utilisé qu'au nombre même de nœuds de la couche cachée.

L'emploi des réseaux de neurones permet d'exploiter les non-linéarités présentes dans les processus qui décrivent les séries de rendements de taux de change quotidiens pour générer des prévisions améliorées des mouvements

futurs de ces mêmes séries. Il est possible, en effet, de construire un réseau de neurones assez flexible pour capter les relations existantes entre les variables explicatives et la variable cible de manière à obtenir des modèles de prévision assez fiables des réalisations futures de cette variable cible.

La stratégie utilisée pour valider la structure optimale ne produit pas les effets escomptés. Les critères de Akaike ou de Schwarz ne sélectionnent pas les modèles qui sont, en bout de ligne, les plus performants. Ils choisissent presque exclusivement les réseaux (1,1) quand nous utilisons les données de l'échantillon de validation. Ces critères ne sont pas plus efficaces s'ils sont appliqués à l'échantillon d'estimation, les modèles qui minimisent *AIC* et *BIC* ne sont pas les réseaux qui produisent les taux de succès plus élevés, les statistiques *PT* significatives ou des *MSPE* inférieures. Samson et White (1995) emploient le critère de Schwarz pour tenter de sélectionner les réseaux de neurones qui livrent les prévisions les plus précises alors que Franses et Van Homelen (1998) exploitent les critères de Akaike et Schwarz pour vérifier s'ils facilitent la sélection des modèles les plus performants, tous deux pour l'échantillon d'estimation.

Ces critères d'information ne constituent pas des outils fiables de sélection des réseaux ( $p, q$ ) lors de l'estimation ou à l'étape de validation.

La stratégie utilisée par Gençay (1999) pour valider les différents réseaux peut s'avérer plus efficace. Il préfère choisir le nombre de nœuds et la taille de l'échantillon d'entraînement de manière dynamique. Bien que l'approche de l'auteur soit de loin la plus longue, elle fournit les meilleurs résultats. Une approche dynamique de validation croisée pour choisir le nombre d'inputs  $p$ , de nœuds de la couche cachée  $q$  et la taille optimale pour l'échantillon d'estimation  $t$  (en même temps) pourrait donc fournir une alternative intéressante à l'utilisation des critères d'information de Akaike et de Schwarz pour déterminer les réseaux ( $p, q, t$ ) que nous voulons exploiter pour obtenir la prévision du rendement au temps  $t + 1$ .

## **VI. Conclusion :**

Plutôt que d'exploiter des techniques paramétriques classiques de modélisation de séries chronologiques, nous utilisons des méthodes non-paramétriques plus appropriées aux processus non-linéaires pour bâtir des modèles de prévisions fiables des mouvements futurs de taux de change quotidiens. Les réseaux de neurones permettent d'approximer une large classe de fonctions non-linéaires et produisent des prévisions plus précises des réalisations futures de la variable dépendante. Un bon nombre de chercheurs réussit en effet à obtenir de meilleures prévisions des rendements futurs de taux de change et à faire mieux que la Marche Aléatoire, le modèle de référence sous l'hypothèse de l'efficacité des marchés.

La première partie de rapport présente les différents éléments de modélisation des réseaux de neurones. Ces méthodes trouvant leur origine dans le domaine de la biologie du système nerveux humain, une nomenclature propre a été développée pour leur usage en économétrie financière. De manière à rendre la compréhension de nos modèles plus aisée, il a été nécessaire de faire le rapprochement entre ces techniques non-paramétriques d'estimation et les méthodes économétriques plus classiques. Les paramètres sont estimés par moindres carrés non-linéaires et le processus itératif d'entraînement est arrêté lorsque la somme des écarts au carré entre l'output du réseau et la variable d'intérêt atteint son minimum. Nous avons ensuite dressé la liste des mesures et statistiques calculées pour évaluer la performance de nos réseaux dans la prévision des rendements futurs de taux de change quotidiens.

Les méthodes non-paramétriques sont certes plus souples dans leur mise en place que leur contrepartie paramétrique mais il n'existe aucune méthode sûre pour déterminer quelle structure utiliser pour obtenir de meilleurs résultats. En effet, le nombre de retards de la variable cible à utiliser comme régresseurs ou le nombre de nœuds à introduire dans la couche intermédiaire du réseau sont inconnus. Il faut alors procéder à une étape de validation qui permettra de

sélectionner le modèle le plus performant parmi l'ensemble des modèles entraînés.

Afin d'obtenir des modèles capables de généraliser à de nouveaux exemples (à l'extérieur de l'échantillon d'estimation), nous proposons d'utiliser des critères d'information de Akaike et Schwarz comme outils de sélection, sur un nouvel échantillon n'ayant pas servi à l'entraînement des réseaux. Nous calculons ces critères pour un échantillon de validation et les modèles qui produisent les valeurs minimales pour ces critères sont utilisés pour obtenir des prévisions pour un troisième échantillon de données. Les critères de sélection de Akaike et Schwarz ne favorisent cependant pas la sélection des modèles de prévision les plus fiables. Une technique de sélection plus dynamique comme celle qui est utilisée par Gençay (1999) peut s'avérer plus efficace dans la sélection des modèles.

Les résultats sont sensibles à la série de taux de change étudiée. Certains réseaux de neurones produisent des prévisions relativement fiables pour une série de rendements donnée mais peut-être pas pour une autre. L'introduction de nouvelles variables explicatives permet d'obtenir des prévisions plus précises dans certains cas. En effet, les réseaux qui utilisent les retards de la variable cible comme inputs sont les plus efficaces pour prévoir les rendements du taux de change du dollar canadien alors que les réseaux qui utilisent les différentiels de taux d'intérêt nominaux sont les plus performants pour le taux de change du Mark allemand (tous par rapport au dollar américain).

Le choix des différentes tailles pour les sous-échantillons peut être modifié. Certains chercheurs ont en effet démontré que les résultats produits par les réseaux de neurones sont également sensibles au choix des longueurs des sous-échantillons utilisés pour les exercices d'estimation, validation et prévision. Le Baron et Weigend (1994) utilisent une technique de Bootstrap pour illustrer l'impact du choix de ces sous-échantillons sur la performance prédictive des réseaux de neurones et concluent que les résultats sont plus sensibles aux échantillons utilisés pour l'estimation, la validation et la prévision qu'à la structure

même du réseau. D'autres auteurs viennent confirmer ces conclusions et ajoutent que la technique de validation utilisée constitue, elle aussi, un élément susceptible d'influencer les résultats.

Les réseaux de neurones offrent une alternative intéressante aux chercheurs désireux de prévoir les rendements futurs de taux de change (ou tout autre réalisations futures de séries chronologiques non-linéaires). Ils permettent, par leur simplicité et flexibilité, de produire des modèles de prévisions assez fiables. Nous devons cependant être attentifs dans leur application à de nouveaux exemples. Chaque cas doit être traité sur une base individuelle car les résultats restent sensibles aux données étudiées de même qu'à la technique employée pour entraîner les réseaux.

Ces modèles prévoient le sens des variations des taux de change dans des proportions raisonnables. Ajuster les positions prises sur les marchés des changes occasionne des coûts de transactions qui viendront réduire les opportunités d'engendrer un profit. Prévoir correctement le signe du rendement de la période suivante dans des proportions de 55% ne sera peut être suffisant pour permettre à l'investisseur de couvrir les coûts reliés à ces ajustements.

## Bibliographie

Brooks C. (1996), *"Testing for non-linearity in daily sterling exchange rates"*, Applied Financial Economics 6, 307-317.

Chinn M.D. (1991), *"Some linear and non-linear thoughts on exchange rates"*, Journal of International Money and Finance, 10, 214-230.

Cybenko G. (1989), *"Approximation by superposition of a sigmoidal function"*, Mathematics of Control, Signals and Systems 2, 303-314.

Diebold F.X., Mariano R.S. (1995), *"Comparing predictive accuracy"*, Journal of Business & Economic Statistics, vol 13, No 3.

Franses P.H., Van Dijk D. (2000), *"Non-linear time series models in empirical finance"*, Cambridge University Press.

Franses P.H., Van Griensven K. (1998), *"Forecasting exchange rates using neural networks for technical trading rules"*, Studies in Nonlinear Dynamics and Econometrics, 2(4) : 109-114.

Franses P.H., Van Homelen P. (1998), *"On forecasting exchange rates using neural networks"*, Applied Financial Economics, 8, 589-596.

Gençay R. (1999), *"Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules"*, Journal of International Economics 47, 91-107.

Gradojevic N., Yang J. (2000), *"The application of artificial neural networks to exchange rate forecasting : the role of market microstructure variables"*, Bank of Canada, Working Paper 2000-23.

Hornik K., Stinchcombe M., White H. (1989), *"Multilayer feedforward networks are universal approximators"*, Neural Networks 2, 359-366.

Hsieh D.A. (1989), *"Testing for nonlinear dependence in daily foreign exchange rates"*, Journal of Business, 62, 339-68.

Kuan C-M, Liu T. (1995), *"Forecasting exchange rates using feedforward and recurrent neural networks"*, Journal of Applied Econometrics, vol 10, 347-364.

Kuan C-M., White H. (1994), *"Artificial neural networks : an econometric perspective"*, Econometric Reviews 13, 1-91.

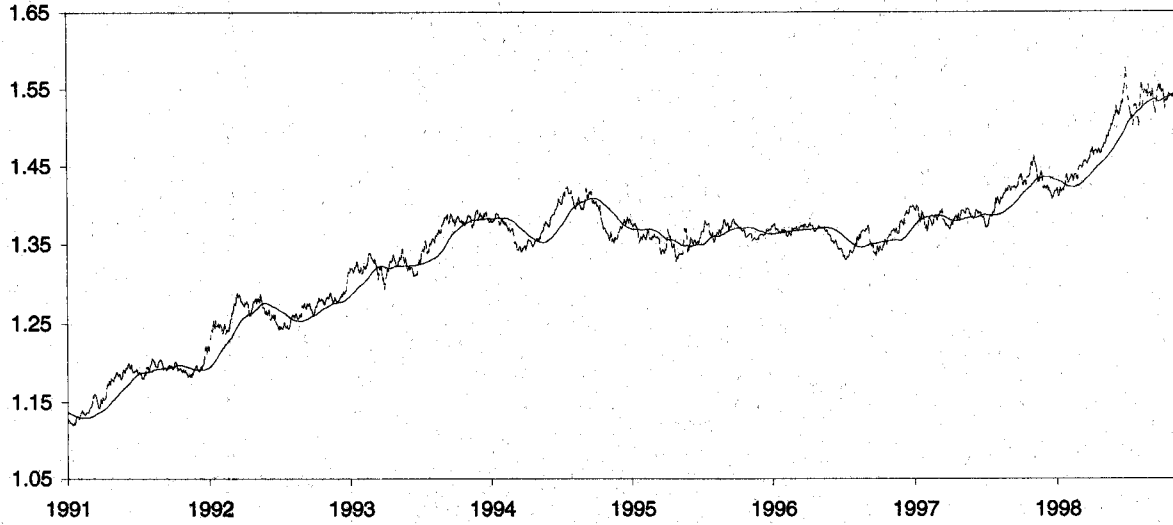
LeBaron B., Weigend A.S. (1998), "*A Bootstrap evaluation of the effect of data splitting on financial time series*", IEEE Transactions on Neural Networks, 213-220.

Pesaran M.H., Timmermann A. (1992), "*A simple nonparametric test of predictive performance*", Journal of Business & Economic Statistics, vol 10, No 4.

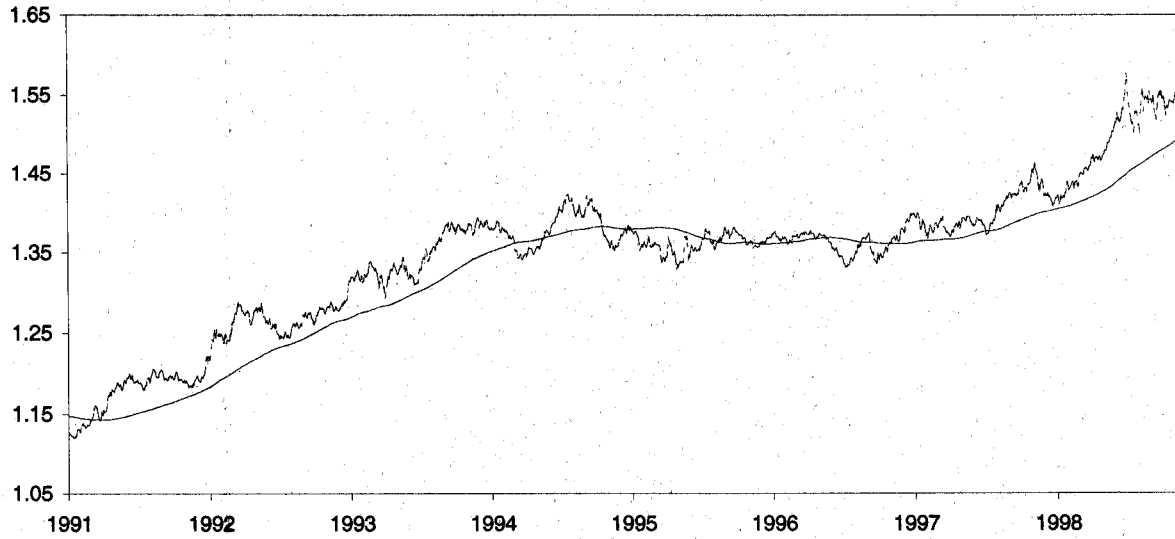
Swanson N.R., White H. (1995), "*A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks*", Journal of Business and Economic Statistics, 13, 265-75.

**Graphique 2 : Analyse Technique – Données Quotidiennes Dollars Canadiens/Dollar Américain**  
Niveaux (ligne pointillée), moyennes mobiles (lignes continues)

Moyenne Mobile 50 jours



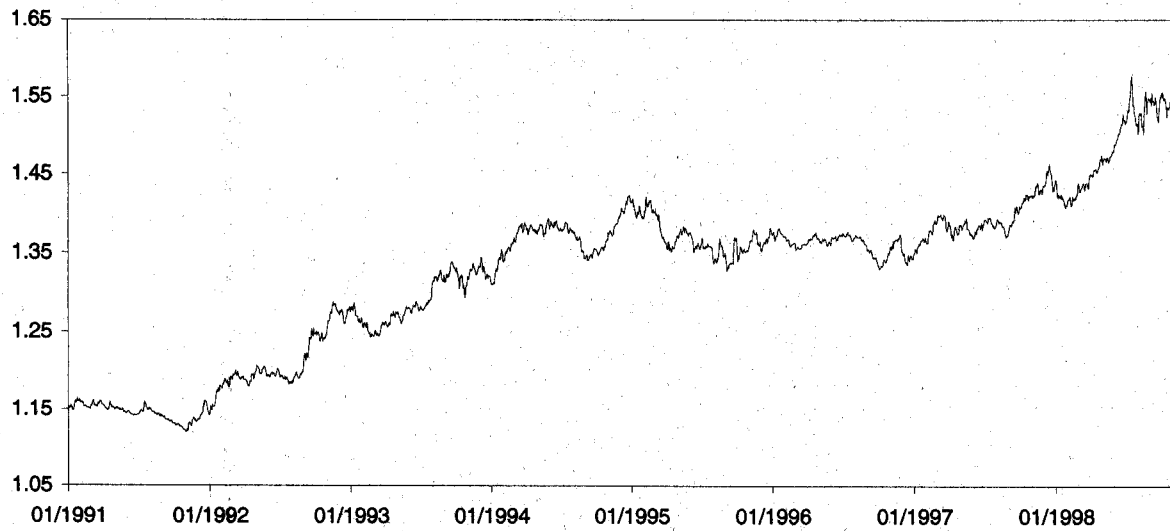
Moyenne Mobile 200 jours



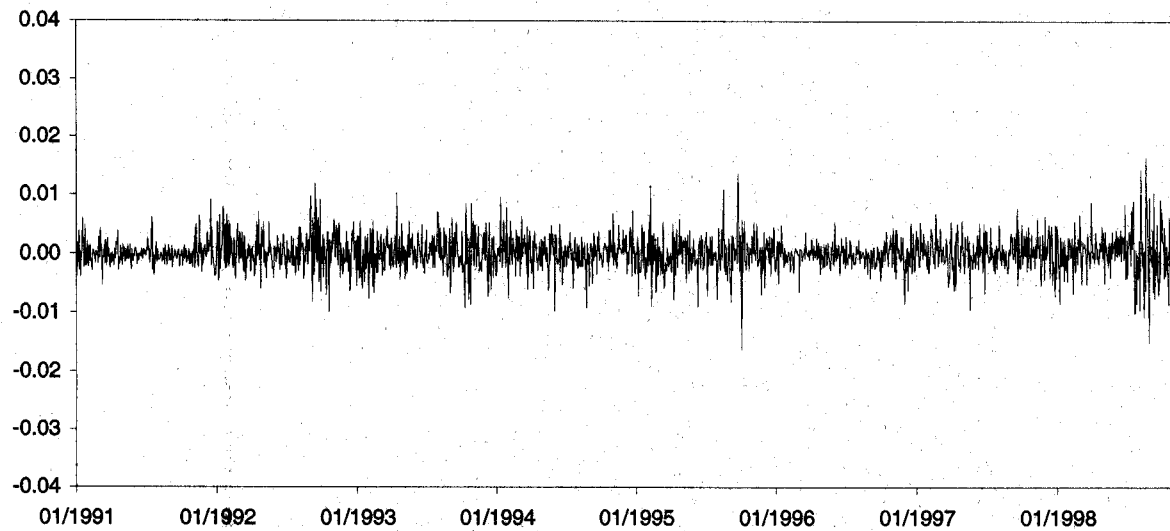


**Graphique 3 : Observations quotidiennes – Dollars Canadiens/Dollar Américain**

**Niveaux**

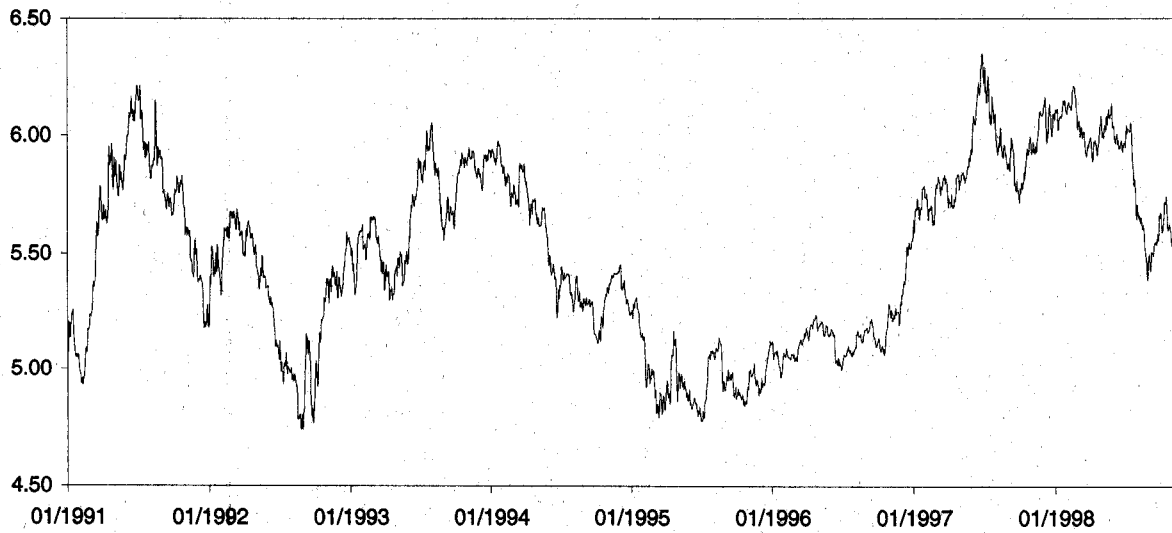


**Rendements**

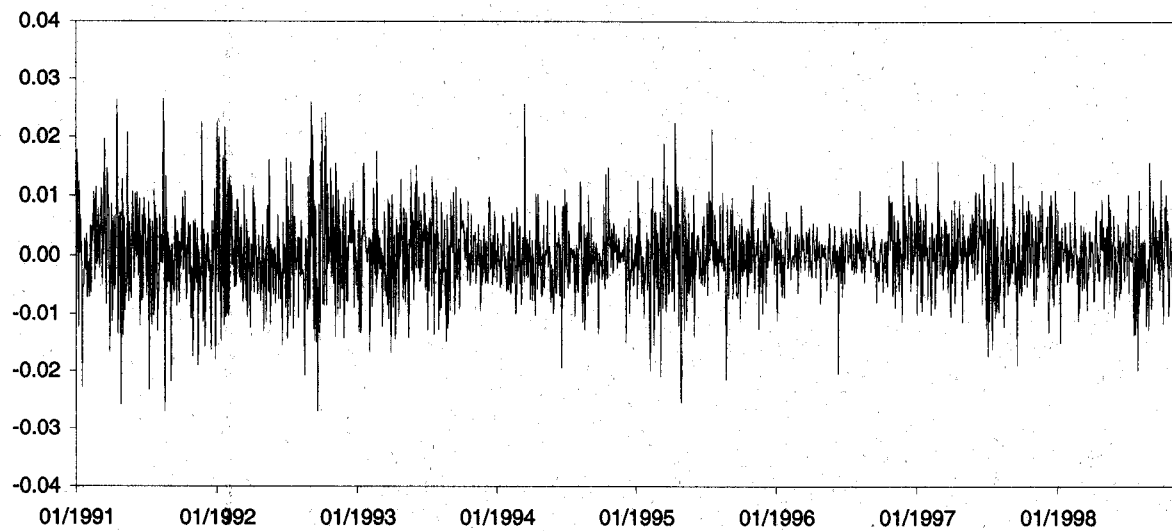


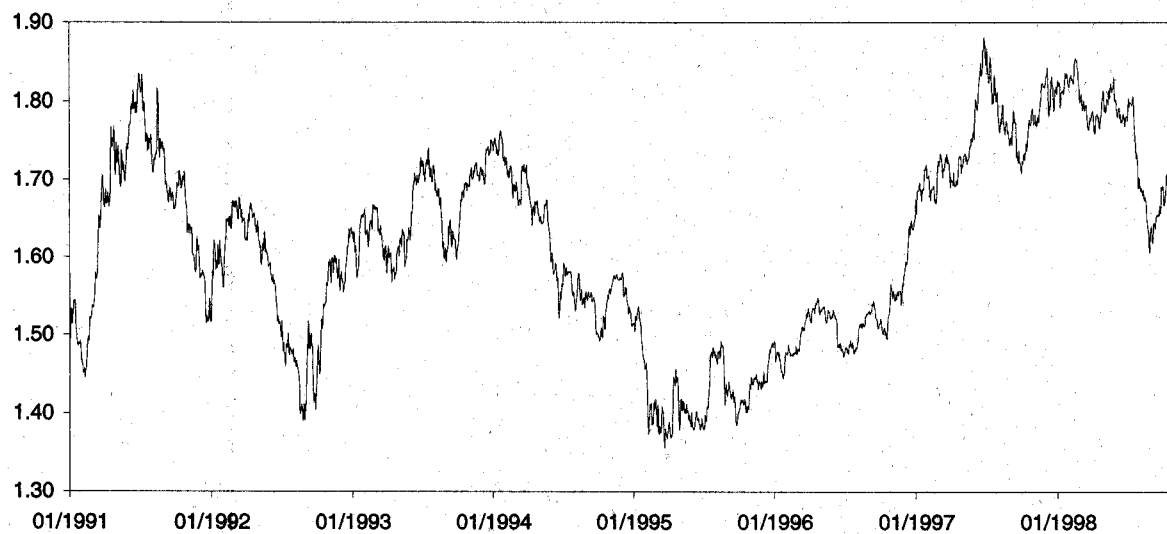
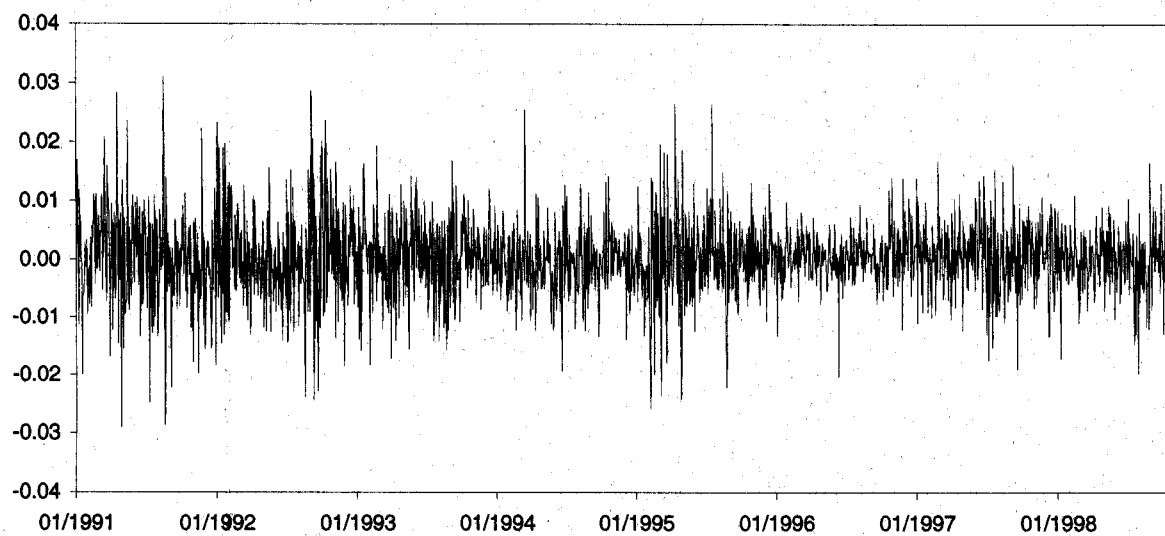
**Graphique 4 : Observations quotidiennes – Francs Français/Dollar Américain**

**Niveaux**



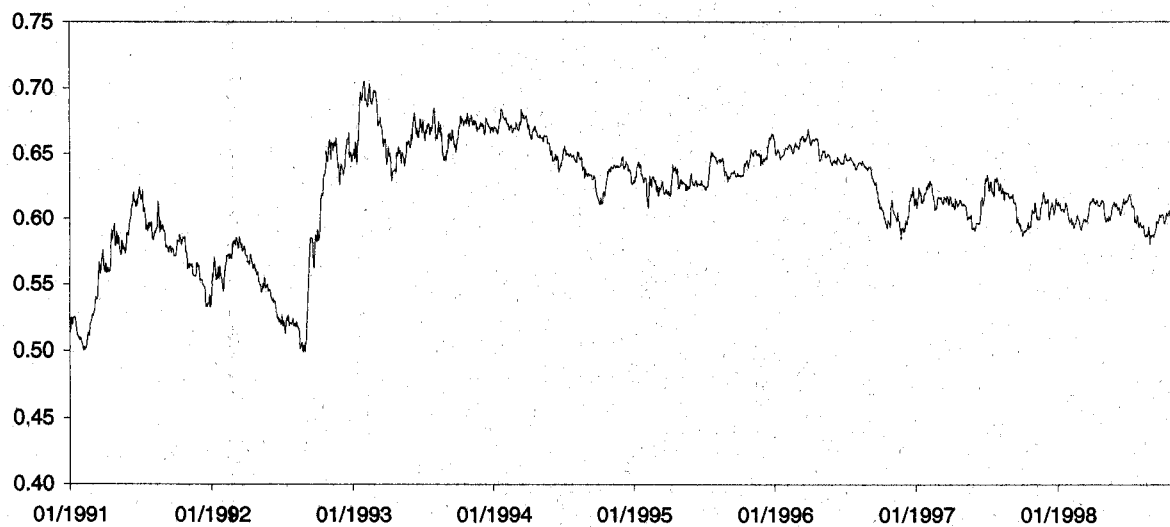
**Rendements**



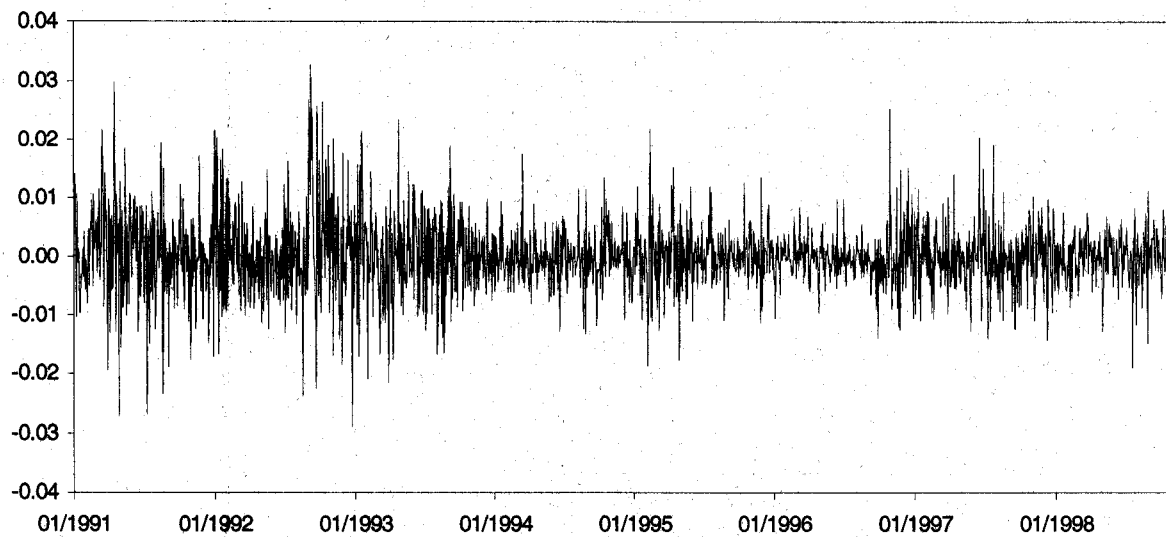
**Graphique 5 : Observations quotidiennes – Marks Allemands/Dollar Américain****Niveaux****Rendements**

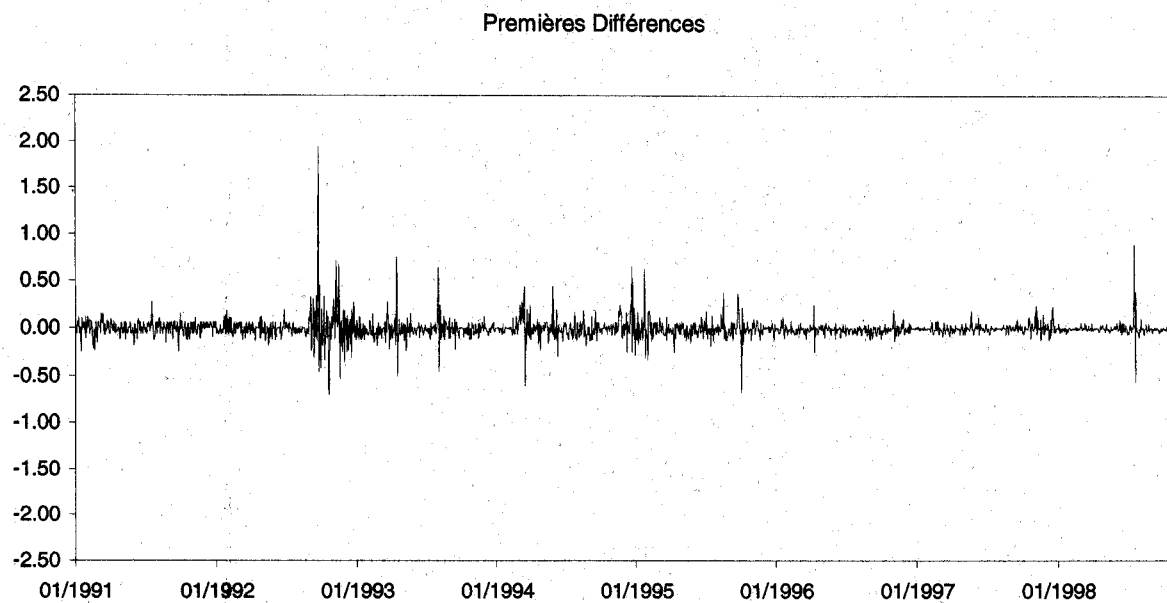
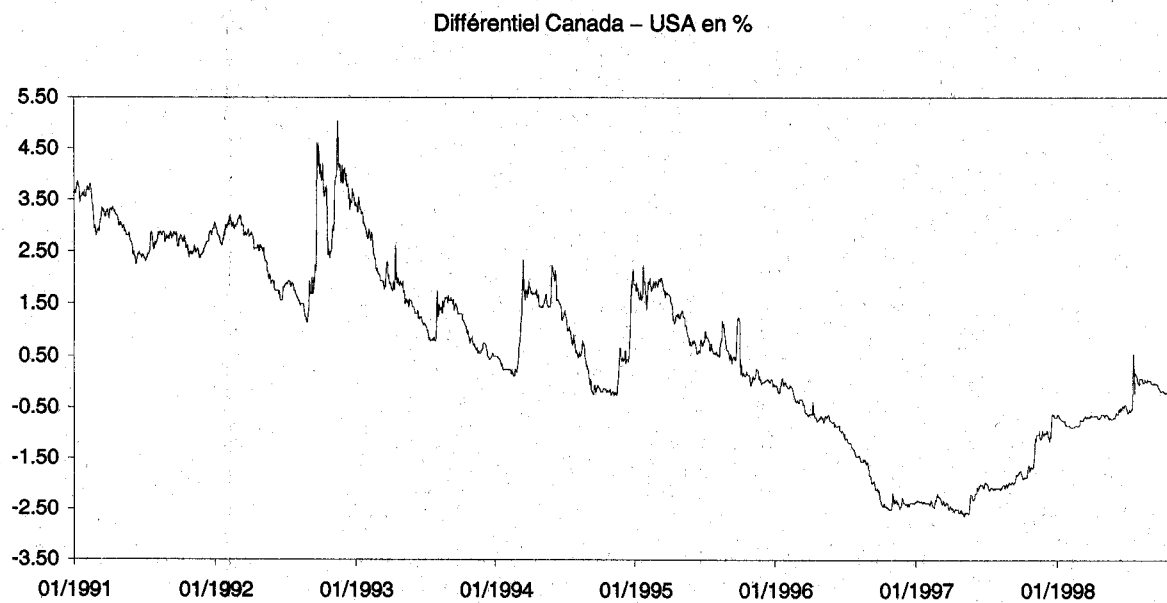
**Graphique 6 : Observations quotidiennes – Livres Anglaises/Dollar Américain**

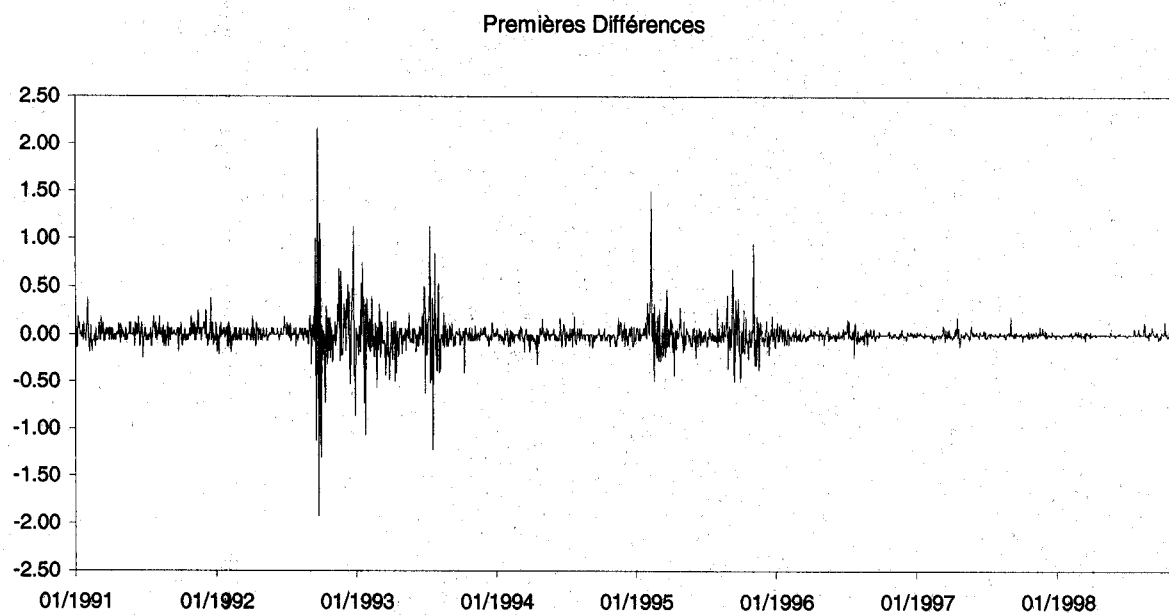
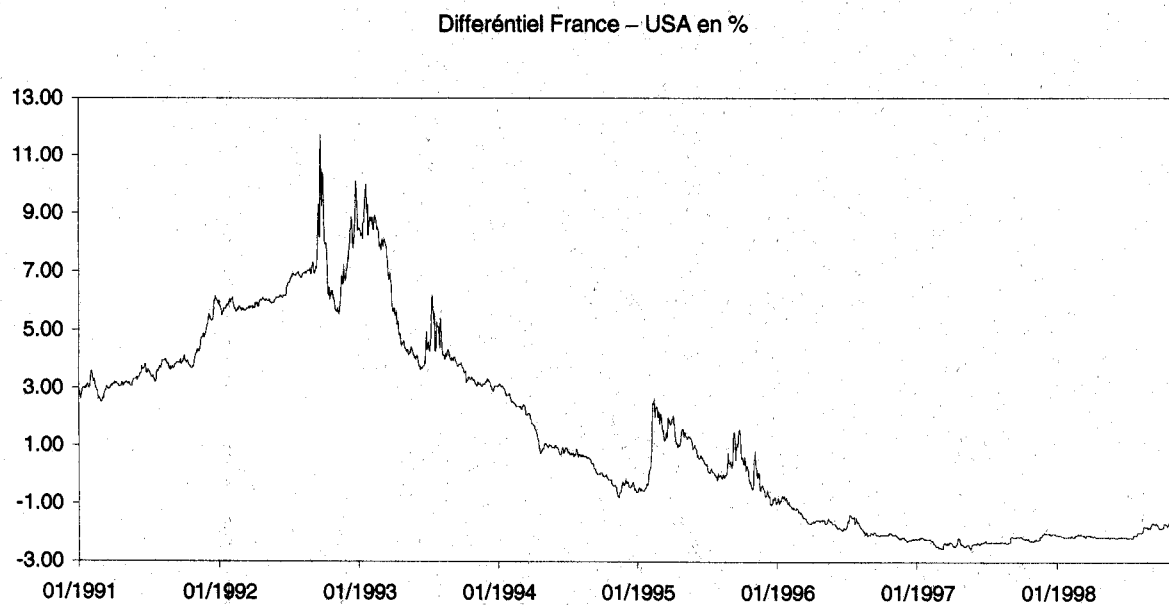
**Niveaux**



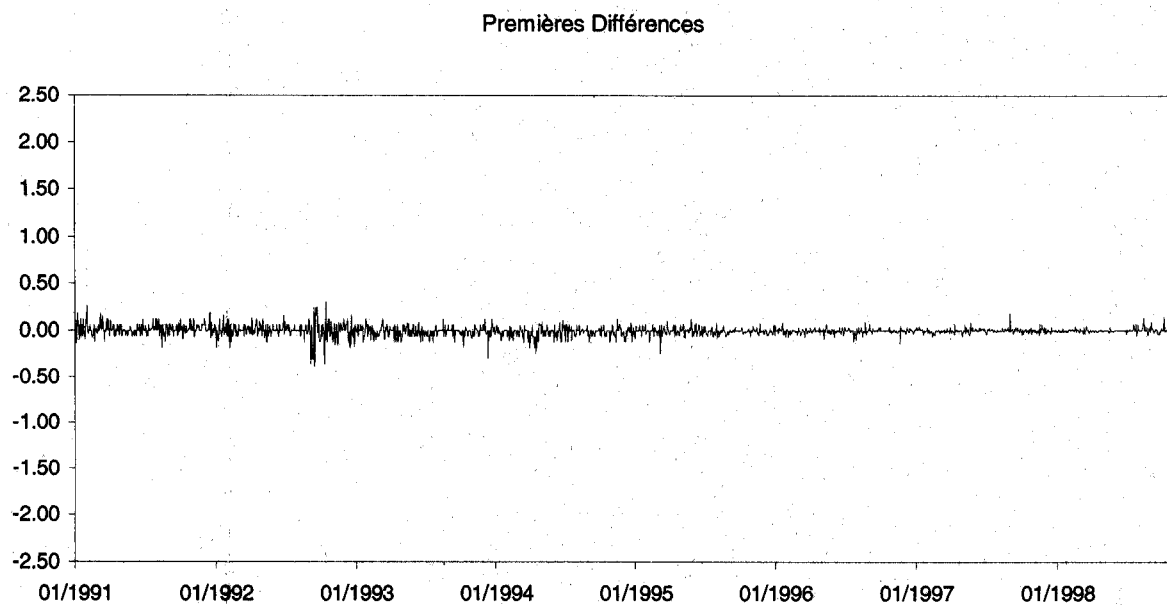
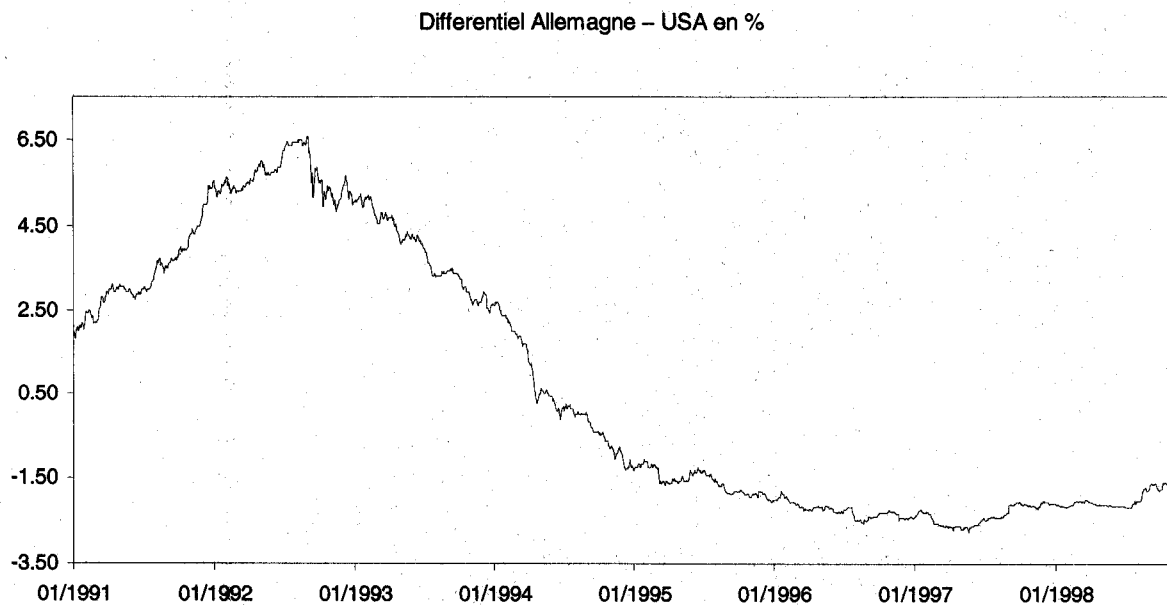
**Rendements**



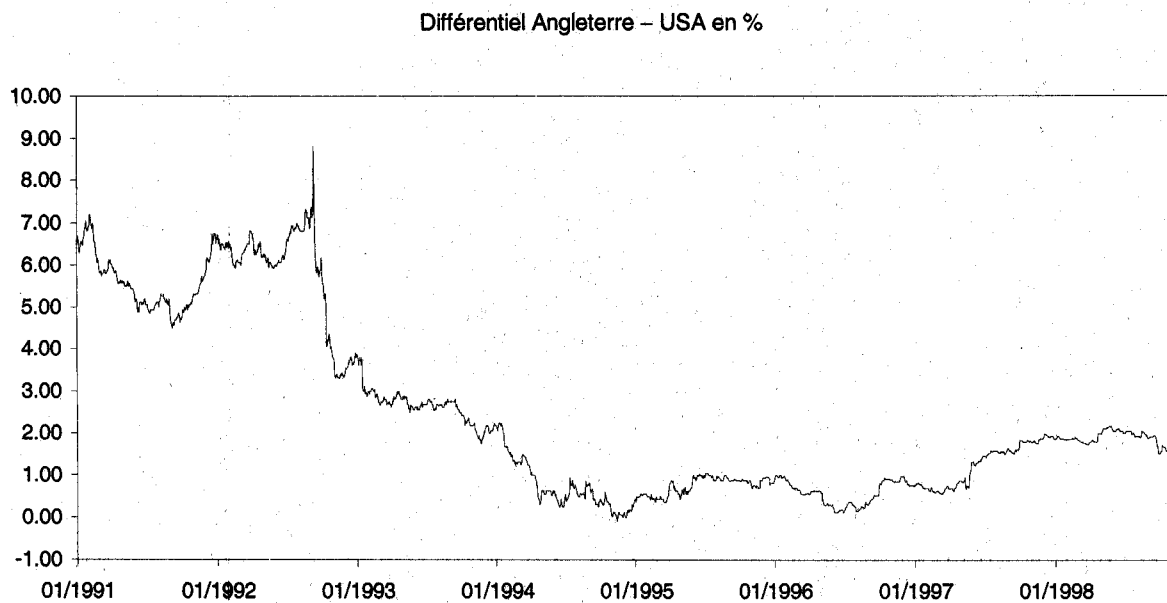
**Graphique 7 : Différentiel Libor 3 mois (CAD – USD), données quotidiennes**

**Graphique 8 : Différentiel Libor 3 mois (FF – USD), données quotidiennes**

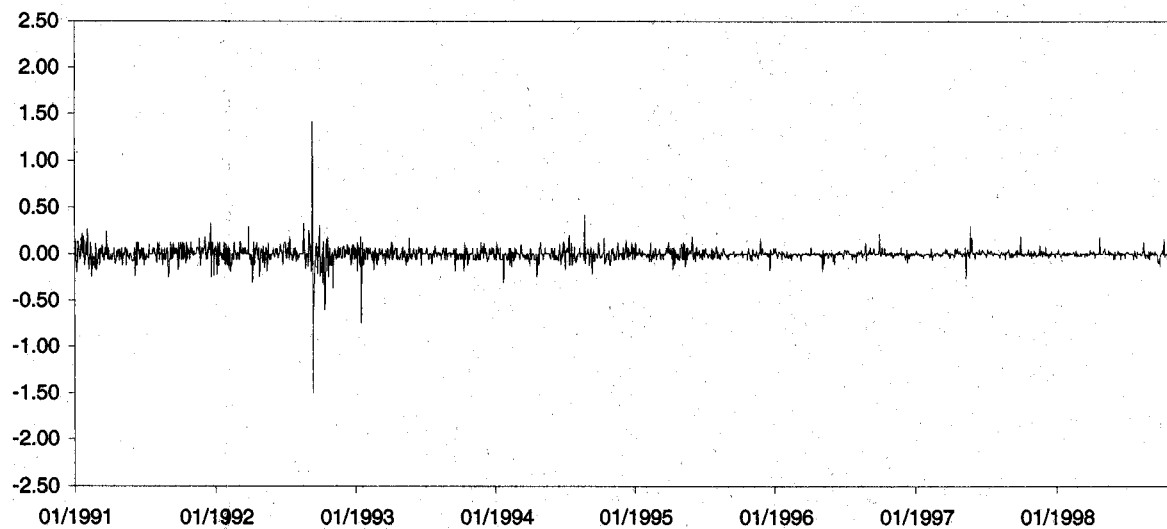
**Graphique 9 : Différentiel Libor 3 mois (DM – USD), données quotidiennes**



**Graphique 10 : Différentiel Libor 3 mois (BP – USD), données quotidiennes**



Premières Différences





**Tableau I : Statistiques - Rendements quotidiens de taux de changes**

<i>Echantillon</i>	<i>Taille</i>	<i>Moyenne</i>	<i>Ecart-Type</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Skewness</i>	<i>Excess Kurtosis</i>
<b>CAD/USD</b>							
Total	1968	0.0144	0.2916	-0.0162	0.0165	0.0420	2.8919
Estimation	984	0.0196	0.2728	-0.0099	0.0118	0.2035	1.5385
Validation	491	-0.0049	0.2756	-0.0162	0.0136	-0.2974	4.7124
Prévision	493	0.0234	0.3395	-0.0151	0.0165	0.0539	2.6907
<b>FF/USD</b>							
Total	1968	0.0049	0.6536	-0.0270	0.0265	-0.0149	1.7007
Estimation	984	0.0053	0.7300	-0.0270	0.0265	0.1316	1.2865
Validation	491	-0.0059	0.5569	-0.0256	0.0225	-0.5064	3.3632
Prévision	493	0.0148	0.5769	-0.0200	0.0160	-0.1651	0.4780
<b>DM/USD</b>							
Total	1968	0.0057	0.6880	-0.0290	0.0310	0.0258	1.7621
Estimation	984	0.0040	0.7628	-0.0290	0.0310	0.1157	1.3227
Validation	491	-0.0014	0.6245	-0.0260	0.0266	-0.1488	3.1784
Prévision	493	0.0162	0.5830	-0.0198	0.0168	-0.1655	0.4875
<b>BP/USD</b>							
Total	1968	0.0079	0.6223	-0.0289	0.0329	0.2805	2.8574
Estimation	984	0.0219	0.7337	-0.0289	0.0329	0.3155	2.0103
Validation	491	-0.0181	0.4688	-0.0187	0.0253	0.2594	3.7478
Prévision	493	0.0060	0.5028	-0.0190	0.0204	-0.0425	1.4640

*Note:* CAD/USD, FF/USD, DM/USD et BP/USD représentent les premières différences du logarithme naturel des taux de change quotidiens du dollar canadien, franc français, Mark allemand et Livre anglaise par rapport au dollar américain, tous exprimés en unités de devise étrangère par dollar américain. Les données peuvent être téléchargées sur le site internet de "The Federal Reserve Bank of St-Louis" à l'adresse [www.stls.frb.com](http://www.stls.frb.com).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prévision de janvier 1997 à décembre 1998.

Les moyennes et écarts-types sont exprimés en pourcentages.

**Tableau II : Statistiques - Changements quotidiens du différentiel du LIBOR 3 mois**

<i>Echantillon</i>	<i>Taille</i>	<i>Moyenne</i>	<i>Ecart-Type</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Skewness</i>	<i>Excess Kurtosis</i>
<b><u>LIBOR CAD-USD</u></b>							
Total	1968	-0.1888	10.8417	-0.7031	1.9453	3.6044	64.0715
Estimation	984	-0.3065	12.9459	-0.7031	1.9453	3.7418	57.3265
Validation	491	-0.6205	9.4907	-0.6719	0.6563	0.8774	16.1961
Prévision	493	0.4759	6.7077	-0.5625	0.8984	4.5882	76.9966
<b><u>LIBOR FF-USD</u></b>							
Total	1968	-0.2262	16.1895	-1.9219	2.1641	0.9476	45.4892
Estimation	984	-0.2922	20.4832	-1.9219	2.1641	0.4205	30.9418
Validation	491	-0.3882	14.1863	-0.5000	1.5000	3.0956	30.9823
Prévision	493	0.0668	2.9027	-0.1697	0.1875	1.4795	12.8869
<b><u>LIBOR DM-USD</u></b>							
Total	1968	-0.1818	5.2916	-0.3906	0.3125	-0.6837	8.7933
Estimation	984	-0.3049	6.8351	-0.3906	0.3125	-0.6357	4.8887
Validation	491	-0.2419	3.4260	-0.2500	0.1250	-0.9366	6.9656
Prévision	493	0.1239	2.6154	-0.1872	0.2120	2.1382	22.1855
<b><u>LIBOR BP-USD</u></b>							
Total	1968	-0.2713	8.4488	-1.5000	1.4219	-1.4156	99.3746
Estimation	984	-0.6455	11.2985	-1.5000	1.4219	-1.2185	60.8202
Validation	491	0.1639	4.1329	-0.1914	0.2148	0.4977	5.2890
Prévision	493	0.0421	3.6253	-0.2695	0.3008	1.3600	20.6931

*Note:* LIBOR CAD-USD, LIBOR FF-USD, LIBOR DM-USD et LIBOR BP-USD représentent les changements quotidiens des différentiels entre les taux LIBOR 3 mois du dollar canadien, franc français, mark allemand et livre anglaise et le taux LIBOR 3 mois du dollar américain. Les données quotidiennes sur les taux LIBOR peuvent être téléchargées sur le site internet de la "British Banks Association" à l'adresse [www.bba.org.uk](http://www.bba.org.uk).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prévision de janvier 1997 à décembre 1998.

Les moyennes et écarts-types sont exprimés en pourcentages.

**Tableau III : Performance du réseau ( $p, q$ ) - Prédiction des rendements quotidiens CAD/USD avec rendements passés comme inputs**

$$r_t = \phi(r_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-11.8010	-11.7761	<u>-11.7615</u>	<u>-11.7174</u>	0.54	-4.1269 <sup>∞</sup>	0.8942	0.4687
2	1	-11.8172	<u>-11.7824</u>	-11.7108	-11.6491	0.52	-0.9511	1.0273	-1.1151
3	1	-11.8077	-11.7628	-11.6597	-11.5803	0.52	-1.1431	1.0186	-1.2943 <sup>°</sup>
4	1	-11.8066	-11.7517	-11.7204	-11.6233	0.54	1.1879	0.9647	0.6674
5	1	-11.7999	-11.7350	-11.6901	-11.5754	0.52	0.0898	1.3135	-1.2615
1	2	-11.8033	-11.7635	-11.7606	-11.6900	0.54	0.0399	1.8812	-0.5315
2	2	-11.8231	-11.7683	-11.5936	-11.4966	0.52	-0.0314	5.4840	-1.0980
3	2	-11.8062	-11.7364	-11.7150	-11.5915	0.51	-0.3641	1.0183	-1.7026 <sup>∞</sup>
4	2	-11.8102	-11.7255	-11.7149	-11.5649	0.53	0.7042	0.9475	0.5507
5	2	-11.8161	-11.7163	-11.6388	-11.4623	0.50	-0.8097	1.3303	-1.2895 <sup>°</sup>
1	3	-11.8043	-11.7496	-11.7432	-11.6462	0.52	-0.9119	0.9363	0.8251
2	3	-11.8276	-11.7529	-11.5520	-11.4197	0.49	-0.9323	7.4714	-1.0895
3	3	-11.8002	-11.7055	-11.6827	-11.5151	0.56	1.9284 <sup>++</sup>	0.9205	0.6515
4	3	-11.8135	-11.6988	-11.5287	-11.3259	0.52	-1.0046	1.0231	-1.1417
5	3	<u>-11.8284</u>	-11.6937	-11.5732	-11.3350	0.50	-0.6369	3.1371	-1.6205 <sup>°</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle ( $p, q$ ) considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle ( $p, q$ ) fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>°, ∞</sup> : le modèle ( $p, q$ ) fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau IV : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens CAD/USD avec signaux de vente-achat comme inputs,  $n_1 = 1, n_2 = 50$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-11.7868	<u>-11.7609</u>	-11.7416	<u>-11.6975</u>	0.55	-0.2141	0.8424	0.7863
2	1	-11.7864	-11.7501	<u>-11.7440</u>	-11.6823	0.56	1.7855 <sup>++</sup>	0.7227	0.7354
3	1	-11.7889	-11.7422	-11.7318	-11.6524	0.55	1.5137 <sup>+</sup>	0.9700	0.7606
4	1	-11.7851	-11.7279	-11.7136	-11.6166	0.53	0.4074	0.9827	0.6066
5	1	-11.7816	-11.7140	-11.7199	-11.6052	0.53	0.7733	0.9510	0.9056
1	2	-11.7878	-11.7463	-11.7118	-11.6413	0.54	0.8971	0.8071	0.4616
2	2	-11.7933	-11.7362	-11.7020	-11.6049	0.55	1.7372 <sup>++</sup>	0.8867	0.9386
3	2	-11.7852	-11.7125	-11.7152	-11.5917	0.55	1.2979 <sup>+</sup>	0.7860	0.4691
4	2	-11.7780	-11.6897	-11.6874	-11.5374	0.51	-0.3853	0.9746	0.7161
5	2	-11.7921	-11.6881	-11.6634	-11.4869	0.52	-0.2443	1.4296	-0.6430
1	3	-11.7871	-11.7301	-11.7027	-11.6056	0.53	0.1602	0.8305	0.3614
2	3	-11.7878	-11.7100	-11.6949	-11.5626	0.55	1.3001 <sup>+</sup>	0.6899	0.6047
3	3	<u>-11.8055</u>	-11.7069	-11.5405	-11.3729	0.52	0.4275	0.7169	0.5915
4	3	-11.7966	-11.6771	-11.6792	-11.4763	0.51	-0.1502	0.9806	0.3998
5	3	-11.7927	-11.6523	-11.6281	-11.3899	0.52	-0.0251	1.5800	-0.7805

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>°</sup>, <sup>∞</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau V** : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens CAD/USD avec signaux de vente-achat comme inputs,  $n_1 = 1$ ,  $n_2 = 200$

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prévision			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-11.6565	<u>-11.6268</u>	<u>-11.7164</u>	<u>-11.6723</u>	0.54	-0.2815	1.0153	-1.0258
2	1	-11.6532	-11.6115	-11.6763	-11.6146	0.53	-0.1572	0.9370	0.9252
3	1	<u>-11.6593</u>	-11.6056	-11.6726	-11.5932	0.53	0.2437	1.0063	-1.6786 <sup>oo</sup>
4	1	-11.6477	-11.5820	-11.6966	-11.5996	0.54	0.4259	1.0365	-1.0308
5	1	-11.6329	-11.5553	-11.6785	-11.5638	0.53	-0.6599	1.0191	-1.1019
1	2	-11.6530	-11.6054	-11.6990	-11.6284	0.55	0.6484	0.9856	0.8060
2	2	-11.6492	-11.5837	-11.7048	-11.6077	0.54	0.2883	0.9893	1.0005
3	2	-11.6556	-11.5721	-11.6568	-11.5334	0.53	0.2874	1.0754	-1.1020
4	2	-11.6429	-11.5414	-11.6652	-11.5152	0.54	0.3321	1.0268	-1.0544
5	2	-11.6364	-11.5169	-11.6310	-11.4546	0.55	0.9893	1.0314	-1.1703
1	3	-11.6468	-11.5814	-11.6975	-11.6005	0.56	1.1023	1.0047	-1.0045
2	3	-11.6446	-11.5553	-11.6657	-11.5334	0.54	0.3737	0.9926	0.9754
3	3	-11.6422	-11.5290	-11.6425	-11.4749	0.54	-0.0056	0.9955	0.8507
4	3	-11.6388	-11.5016	-11.6426	-11.4398	0.54	0.3935	1.0118	-1.2019
5	3	-11.6453	-11.4840	-11.5822	-11.3440	0.55	0.8215	1.0561	-1.0792

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prévision de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau VI : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens CAD/USD avec changements quotidiens des différentiels de taux d'intérêt comme inputs**

$$r_t = \phi(\Delta D_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prévision			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-11.8163	<u>-11.7914</u>	<u>-11.7415</u>	<u>-11.6973</u>	0.55	1.1392	0.9993	0.1075
2	1	-11.8192	-11.7843	-11.7356	-11.6738	0.53	0.7248	0.9970	0.2053
3	1	-11.8171	-11.7723	-11.7285	-11.6491	0.55	0.8155	1.0068	-0.5683
4	1	-11.8200	-11.7651	-11.6972	-11.6002	0.55	0.7021	0.9888	0.9346
5	1	-11.8188	-11.7539	-11.7024	-11.5877	0.57	2.1355 <sup>++</sup>	1.0105	-0.7605
1	2	-11.8108	-11.7710	-11.7313	-11.6608	0.56	2.2182 <sup>++</sup>	1.0008	-0.1133
2	2	-11.8103	-11.7555	-11.7079	-11.6109	0.55	0.8783	0.9985	0.1762
3	2	-11.8128	-11.7430	-11.7141	-11.5906	0.55	1.4646 <sup>+</sup>	1.0115	-0.9100
4	2	-11.8202	-11.7354	-11.6618	-11.5118	0.56	1.4300 <sup>+</sup>	0.9927	0.5517
5	2	-11.8202	-11.7203	-11.6793	-11.5028	0.56	1.5725 <sup>+</sup>	1.0049	-0.5636
1	3	-11.8078	-11.7531	-11.7197	-11.6226	0.55	0.2623	0.9977	0.3072
2	3	-11.8161	-11.7414	-11.7021	-11.5698	0.55	1.8764 <sup>++</sup>	0.9821	1.5068 <sup>+</sup>
3	3	<u>-11.8255</u>	-11.7308	-11.6801	-11.5125	0.55	0.8531	1.0155	-1.3044 <sup>o</sup>
4	3	-11.8233	-11.7086	-11.6277	-11.4248	0.55	1.1724	1.0044	-0.3326
5	3	-11.8194	-11.6846	-11.5922	-11.3540	0.52	0.1469	0.9968	0.3382

*Note:* AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prévision de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau VII : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens FF/USD avec rendements passés comme inputs**

$$r_t = \phi(r_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8380	<u>-9.8131</u>	<u>-10.3697</u>	<u>-10.3256</u>	0.52	0.5376	2.1397	-1.0009
2	1	-9.8372	<u>-9.8023</u>	<u>-10.3555</u>	<u>-10.2937</u>	0.47	-1.1805	1.1225	-1.0101
3	1	-9.8393	-9.7945	-10.3500	-10.2706	0.51	0.5477	2.3089	-1.0522
4	1	-9.8385	-9.7836	-10.3172	-10.2202	0.46	-1.4106 <sup>o</sup>	1.0563	-1.0455
5	1	-9.8501	-9.7852	-10.2799	-10.1652	0.49	-0.6069	1.0194	-1.1925
1	2	-9.8408	-9.8010	-10.3542	-10.2837	0.49	-0.7650	2.2005	-1.0014
2	2	-9.8409	-9.7861	-10.3432	-10.2462	0.46	-2.0285 <sup>oo</sup>	1.5669	-1.0548
3	2	-9.8443	-9.7745	-10.2890	-10.1655	0.48	-0.8285	1.0162	-0.9769
4	2	-9.8413	-9.7565	-10.2804	-10.1305	0.47	-1.3569 <sup>o</sup>	1.0832	-1.1220
5	2	-9.8425	-9.7427	-10.3120	-10.1355	0.48	-0.5490	2.3581	-1.2171
1	3	-9.8397	-9.7850	-10.3359	-10.2388	0.48	-0.3205	2.5887	-1.0011
2	3	-9.8444	-9.7697	-10.3242	-10.1918	0.48	-0.9063	1.3504	-1.0313
3	3	-9.8490	-9.7544	-10.2679	-10.1003	0.47	-1.0124	1.0423	-1.0043
4	3	-9.8612	-9.7465	-10.2187	-10.0158	0.50	-0.1640	1.0401	-1.4962 <sup>o</sup>
5	3	<u>-9.8613</u>	-9.7266	-10.2292	-9.9910	0.49	-0.0049	1.0327	-1.7236 <sup>oo</sup>

*Note:* AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau VIII : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens FF/USD avec signaux de vente-achat comme inputs,  $n_1 = 1$ ,  $n_2 = 50$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8279	<u>-9.8020</u>	<u>-10.3529</u>	<u>-10.3088</u>	0.49	-0.2827	1.0009	-3.5398 <sup>oo</sup>
2	1	-9.8325	-9.7962	-10.3473	-10.2856	0.49	-0.2336	1.0110	-1.0269
3	1	-9.8470	-9.8003	-10.3309	-10.2515	0.48	-0.8198	1.1283	-1.4153 <sup>o</sup>
4	1	-9.8532	-9.7960	-10.2756	-10.1786	0.52	1.1649	1.1781	-2.9184 <sup>oo</sup>
5	1	-9.8415	-9.7739	-10.3197	-10.2050	0.50	-0.0374	1.5785	-2.5122 <sup>oo</sup>
1	2	-9.8289	-9.7875	-10.3440	-10.2734	0.50	0.4191	2.6114	-3.9515 <sup>oo</sup>
2	2	-9.8372	-9.7802	-10.2918	-10.1947	0.48	-0.3386	1.0257	-1.7517 <sup>oo</sup>
3	2	-9.8403	-9.7676	-10.2956	-10.1721	0.47	-1.2129	1.1625	-1.3155 <sup>o</sup>
4	2	-9.8430	-9.7547	-10.2582	-10.1082	0.49	-0.4838	1.2156	-1.6629 <sup>oo</sup>
5	2	-9.8427	-9.7387	-10.2188	-10.0424	0.49	-0.7402	1.1552	-2.9195 <sup>oo</sup>
1	3	-9.8249	-9.7679	-10.3258	-10.2287	0.48	-0.2285	2.1675	-3.9513 <sup>oo</sup>
2	3	-9.8444	-9.7666	-10.2497	-10.1174	0.47	-1.2130	1.4214	-3.8606 <sup>oo</sup>
3	3	-9.8555	-9.7568	-10.2414	-10.0738	0.49	-0.2276	1.0865	-2.0971 <sup>oo</sup>
4	3	-9.8499	-9.7304	-10.2054	-10.0025	0.48	-0.7571	1.3944	-2.9042 <sup>oo</sup>
5	3	<u>-9.8686</u>	-9.7282	-10.1495	-9.9113	0.50	-0.2697	1.1114	-1.6099 <sup>o</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.



**Tableau IX : Performance du réseau  $(p, q)$  - Prévision des rendements quotidiens FF/USD avec signaux de vente-achat comme inputs,  $n_1 = 1, n_2 = 200$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prévision			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.9213	<u>-9.8916</u>	<u>-10.3596</u>	<u>-10.3155</u>	0.46	-1.6025 °	2.5644	-4.0553 °°
2	1	-9.9090	-9.8674	-10.3480	-10.2862	0.46	-1.3455 °	1.0094	-1.1771
3	1	-9.9171	-9.8634	-10.3346	-10.2552	0.47	-0.8606	1.9575	-3.8163 °°
4	1	-9.9204	-9.8547	-10.3081	-10.2110	0.48	-0.4783	5.3311	-3.8114 °°
5	1	-9.8981	-9.8205	-10.3174	-10.2027	0.47	-0.7859	1.0225	-1.8260 °°
1	2	-9.9138	-9.8662	-10.3469	-10.2763	0.46	-1.4967 °	4.5527	-4.0563 °°
2	2	-9.9283	-9.8627	-10.3197	-10.2227	0.49	0.1369	1.0008	-0.7858
3	2	-9.9252	-9.8417	-10.2881	-10.1646	0.48	-0.4525	1.9029	-3.9645 °°
4	2	-9.9136	-9.8121	-10.2836	-10.1336	0.49	-0.0448	1.8048	-3.8411 °°
5	2	-9.9286	-9.8091	-10.2608	-10.0844	0.50	0.1042	1.3335	-3.5980 °°
1	3	-9.9332	-9.8678	-10.3234	-10.2263	0.47	-0.8915	1.0102	-3.6446 °°
2	3	-9.9137	-9.8243	-10.2850	-10.1527	0.48	-0.9506	1.3099	-7.1084 °°
3	3	-9.9335	-9.8202	-10.2277	-10.0601	0.47	-0.8062	1.0477	-1.3912 °
4	3	<u>-9.9476</u>	-9.8103	-10.2464	-10.0435	0.47	-0.7662	1.2531	-3.6096 °°
5	3	-9.9342	-9.7729	-10.2020	-9.9639	0.50	0.1546	5.1931	-3.0116 °°

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prévision de janvier 1997 à décembre 1998.

°, °° : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

°, °° : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau X :** Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens FF/USD avec changements quotidiens des différentiels de taux d'intérêt comme inputs

$$r_t = \phi(\Delta D_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8309	-9.8060	-10.3477	<u>-10.3036</u>	0.51	-1.7288 <sup>∞</sup>	1.0000	0.0464
2	1	-9.8304	-9.7956	<u>-10.3509</u>	-10.2892	0.48	0.6259	1.0005	-2.0449 <sup>∞</sup>
3	1	-9.8512	<u>-9.8063</u>	-10.2711	-10.1917	0.46	-0.7871	1.0044	-1.1502
4	1	-9.8305	-9.7757	-10.3049	-10.2078	0.54	1.0650	1.0062	-1.5511 <sup>°</sup>
5	1	-9.8504	-9.7855	-10.2501	-10.1354	0.47	-0.0694	1.0052	-1.2724
1	2	-9.8343	-9.7945	-10.3307	-10.2602	0.48	0.5207	1.0001	-1.1986
2	2	-9.8543	-9.7995	-10.2211	-10.1241	0.47	0.5192	0.9998	0.0249
3	2	-9.8468	-9.7771	-10.2167	-10.0932	0.47	-0.3120	1.0062	-1.8433 <sup>∞</sup>
4	2	-9.8491	-9.7643	-10.2319	-10.0820	0.46	-1.5444 <sup>°</sup>	1.0017	-1.0213
5	2	-9.8424	-9.7425	-10.2590	-10.0825	0.51	-0.5381	1.0380	-1.7575 <sup>∞</sup>
1	3	-9.8463	-9.7915	-10.3206	-10.2236	0.52	-1.2562	1.0000	-0.0503
2	3	<u>-9.8623</u>	-9.7876	-10.2414	-10.1090	0.48	0.7968	1.0061	-1.3295 <sup>°</sup>
3	3	-9.8441	-9.7494	-10.1858	-10.0182	0.47	-0.3279	1.0075	-1.8434 <sup>∞</sup>
4	3	-9.8460	-9.7313	-10.2269	-10.0240	0.47	0.0331	1.0104	-1.8260 <sup>∞</sup>
5	3	-9.8426	-9.7078	-10.2038	-9.9657	0.45	-1.8326 <sup>∞</sup>	1.0139	-1.8708 <sup>∞</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+, ++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>°, ∞</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XI : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens DM/USD avec rendements passés comme inputs**

$$r_t = \phi(r_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.7478	<u>-9.7229</u>	<u>-10.1374</u>	<u>-10.0933</u>	0.51	0.1823	1.5502	-1.0014
2	1	-9.7511	-9.7163	-10.0982	-10.0364	0.51	0.4808	1.0270	-1.0787
3	1	-9.7529	-9.7080	-10.0874	-10.0080	0.47	-0.7214	1.0027	-1.7729 <sup>oo</sup>
4	1	-9.7549	-9.7000	-10.0774	-9.9803	0.47	-0.9000	1.0123	-1.6075 <sup>o</sup>
5	1	-9.7589	-9.6940	-10.0714	-9.9567	0.50	0.0186	1.0461	-1.4040 <sup>o</sup>
1	2	-9.7446	-9.7048	-10.1271	-10.0565	0.51	0.5903	1.9194	-1.0015
2	2	-9.7527	-9.6980	-10.1169	-10.0199	0.49	0.3902	2.0828	-1.0304
3	2	-9.7637	-9.6939	-10.0365	-9.9130	0.46	-1.3486 <sup>o</sup>	1.0343	-1.3892 <sup>o</sup>
4	2	-9.7637	-9.6790	-10.0416	-9.8916	0.51	0.2824	1.4234	-1.3541 <sup>o</sup>
5	2	-9.7684	-9.6686	-10.0140	-9.8376	0.47	-1.1303	1.0386	-1.4334 <sup>o</sup>
1	3	-9.7385	-9.6838	-10.1144	-10.0174	0.51	0.6741	1.5359	-1.0019
2	3	-9.7565	-9.6818	-10.0899	-9.9576	0.50	0.5278	2.0986	-1.0153
3	3	-9.7658	-9.6711	-10.0227	-9.8551	0.47	-1.3397 <sup>o</sup>	1.0504	-1.0902
4	3	-9.7764	-9.6617	-10.0076	-9.8047	0.47	-1.4629 <sup>o</sup>	1.0692	-1.3921 <sup>o</sup>
5	3	<u>-9.7967</u>	-9.6620	-10.0006	-9.7624	0.50	0.3920	1.0904	-1.1735

*Note:* AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

+, ++ : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XII : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens DM/USD avec signaux de vente-achat comme inputs,  $n_1 = 1, n_2 = 50$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.7350	<u>-9.7091</u>	<u>-10.1202</u>	<u>-10.0761</u>	0.47	-1.2589	1.0000	-0.0751
2	1	-9.7387	-9.7024	-10.1131	-10.0514	0.49	-0.3712	1.0002	-0.3833
3	1	<u>-9.7410</u>	-9.6943	-10.0974	-10.0180	0.52	1.0192	1.0014	-1.1346
4	1	-9.7359	-9.6788	-10.0857	-9.9886	0.52	0.8419	1.0019	-1.2943 <sup>o</sup>
5	1	-9.7321	-9.6645	-10.0790	-9.9643	0.51	0.5994	1.0024	-1.7710 <sup>oo</sup>
1	2	-9.7286	-9.6872	-10.1076	-10.0370	0.48	-0.4903	1.0000	0.5796
2	2	-9.7301	-9.6731	-10.0962	-9.9992	0.50	0.2361	1.0001	-0.3382
3	2	-9.7306	-9.6579	-10.0756	-9.9521	0.51	0.7603	1.0018	-1.1192
4	2	-9.7224	-9.6341	-10.0617	-9.9117	0.49	-0.0461	1.0025	-1.6000 <sup>o</sup>
5	2	-9.7168	-9.6128	-10.0465	-9.8700	0.51	0.7664	1.0030	-2.0379 <sup>oo</sup>
1	3	-9.7222	-9.6652	-10.0947	-9.9977	0.47	-1.1863	1.0000	0.2321
2	3	-9.7215	-9.6437	-10.0792	-9.9469	0.50	0.0676	1.0001	-0.3045
3	3	-9.7193	-9.6207	-10.0535	-9.8859	0.51	0.8311	1.0021	-1.7016 <sup>oo</sup>
4	3	-9.7100	-9.5906	-10.0342	-9.8314	0.51	0.4749	1.0021	-1.5132 <sup>o</sup>
5	3	-9.7018	-9.5614	-10.0195	-9.7814	0.51	0.7299	1.0024	-1.7972 <sup>oo</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XIII : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens DM/USD avec signaux de vente-achat comme inputs,  $n_1 = 1, n_2 = 200$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	<u>-9.8347</u>	<u>-9.8049</u>	<u>-10.1199</u>	<u>-10.0758</u>	0.47	-0.3754	0.9999	0.2780
2	1	-9.8295	-9.7878	-10.1121	-10.0503	0.47	-0.2968	1.0000	0.0283
3	1	-9.8235	-9.7698	-10.1031	-10.0237	0.48	0.0554	1.0001	-0.2178
4	1	-9.8245	-9.7588	-10.0875	-9.9904	0.48	-0.5490	1.0035	-3.9057 <sup>oo</sup>
5	1	-9.8186	-9.7410	-10.0767	-9.9620	0.46	-1.0378	1.0021	-2.5923 <sup>oo</sup>
1	2	-9.8270	-9.7794	-10.1072	-10.0366	0.47	-0.3754	0.9999	0.2663
2	2	-9.8192	-9.7537	-10.0950	-9.9980	0.47	-0.2273	1.0001	-0.1302
3	2	-9.8098	-9.7263	-10.0818	-9.9583	0.48	0.7527	1.0004	-1.3254 <sup>o</sup>
4	2	-9.8051	-9.7037	-10.0568	-9.9069	0.47	-0.4628	1.0016	-2.1834 <sup>oo</sup>
5	2	-9.8001	-9.6807	-10.0432	-9.8668	0.46	-1.0843	1.0027	-3.0851 <sup>oo</sup>
1	3	-9.8194	-9.7539	-10.0945	-9.9974	0.47	-0.5111	1.0000	0.0862
2	3	-9.8090	-9.7197	-10.0781	-9.9457	0.47	-0.4346	1.0000	0.0117
3	3	-9.7979	-9.6846	-10.0609	-9.8933	0.48	0.0385	1.0002	-0.2900
4	3	-9.7881	-9.6509	-10.0329	-9.8300	0.47	-1.1088	1.0008	-0.9071
5	3	-9.7838	-9.6225	-10.0137	-9.7755	0.46	-1.0765	1.0037	-3.0395 <sup>oo</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XIV : Performance du réseau ( $p, q$ ) - Prédiction des rendements quotidiens DM/USD avec différentiels de taux d'intérêt comme inputs**

$$r_t = \phi(\Delta D_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.7410	-9.7161	-10.1209	-10.0768	0.52	0.0488	0.9995	1.0225
2	1	-9.7376	-9.7028	-10.1126	-10.0508	0.52	-0.2234	0.9994	0.9837
3	1	-9.7382	-9.6934	-10.1033	-10.0239	0.50	0.4471	0.9993	1.0652
4	1	-9.7362	-9.6813	-10.0861	-9.9890	0.51	0.3435	0.9996	0.3931
5	1	-9.7339	-9.6690	-10.0765	-9.9619	0.52	0.4785	0.9996	0.3931
1	2	-9.7349	-9.6951	-10.1081	-10.0376	0.52	-0.1877	0.9995	1.0345
2	2	-9.7296	-9.6748	-10.0956	-9.9985	0.52	-0.2923	0.9994	1.0189
3	2	-9.7276	-9.6578	-10.0819	-9.9584	0.49	-0.1579	0.9995	0.8779
4	2	-9.7232	-9.6384	-10.0564	-9.9064	0.50	-0.8685	0.9996	0.5806
5	2	-9.7195	-9.6196	-10.0470	-9.8706	0.48	-0.6340	0.9999	0.0908
1	3	-9.7289	-9.6742	-10.0953	-9.9983	0.53	0.2552	0.9992	0.9917
2	3	-9.7215	-9.6468	-10.0786	-9.9462	0.51	0.2656	0.9993	1.0237
3	3	-9.7177	-9.6231	-10.0610	-9.8933	0.49	0.8358	0.9994	0.9445
4	3	-9.7116	-9.5969	-10.0357	-9.8328	0.51	0.3947	0.9995	0.4799
5	3	-9.7050	-9.5702	-10.0174	-9.7792	0.50	-0.8769	1.0000	0.0419

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle ( $p, q$ ) considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

+, ++ : le modèle ( $p, q$ ) fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

°, ∞ : le modèle ( $p, q$ ) fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XV : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens BP/USD avec rendements passés comme inputs**

$$r_t = \phi(r_{t-p}, \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prévision			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8408	-9.8159	<u>-10.7795</u>	<u>-10.7354</u>	0.50	0.3137	1.0057	-0.9851
2	1	-9.8446	-9.8097	-10.7443	-10.6825	0.52	1.3085 <sup>+</sup>	0.9981	0.2076
3	1	-9.8561	-9.8112	-10.7018	-10.6224	0.51	0.7992	1.0116	-0.7274
4	1	-9.8714	<u>-9.8166</u>	-10.6641	-10.5671	0.51	0.4950	1.0476	-0.9532
5	1	-9.8562	-9.7913	-10.6665	-10.5518	0.52	1.3772 <sup>+</sup>	1.0179	-1.1282
1	2	-9.8419	-9.8021	-10.7707	-10.7001	0.51	0.3486	0.9989	0.1916
2	2	-9.8544	-9.7996	-10.7134	-10.6164	0.53	1.1236	1.0002	-0.0136
3	2	-9.8844	-9.8147	-10.6877	-10.5642	0.51	0.7668	1.0218	-1.3294 <sup>o</sup>
4	2	-9.8490	-9.7642	-10.6571	-10.5071	0.51	0.4361	1.0257	-1.5034 <sup>o</sup>
5	2	-9.8686	-9.7688	-10.6023	-10.4258	0.48	-0.8533	1.0823	-3.3940 <sup>oo</sup>
1	3	-9.8396	-9.7849	-10.7565	-10.6595	0.50	-0.5957	1.0069	-0.5841
2	3	-9.8469	-9.7723	-10.6991	-10.5667	0.53	1.4207 <sup>+</sup>	0.9989	0.0774
3	3	-9.8626	-9.7679	-10.6665	-10.4989	0.48	-1.3521 <sup>o</sup>	1.0243	-1.6665 <sup>oo</sup>
4	3	<u>-9.8862</u>	-9.7715	-10.5393	-10.3364	0.46	-1.5917 <sup>o</sup>	1.0970	-1.5769 <sup>o</sup>
5	3	-9.8701	-9.7353	-10.5547	-10.3165	0.53	1.2833 <sup>+</sup>	1.0599	-2.3567 <sup>oo</sup>

*Note:* AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prévision de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o</sup>, <sup>oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XVI : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens BP/USD avec signaux de vente-achat comme inputs,  $n_1 = 1, n_2 = 50$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8052	-9.7792	<u>-10.7972</u>	<u>-10.7531</u>	0.48	-0.6652	1.0016	-0.3515
2	1	-9.8308	-9.7945	-10.7799	-10.7182	0.49	-0.2333	1.0035	-0.8230
3	1	-9.8364	-9.7897	-10.7056	-10.6262	0.52	1.4143 <sup>+</sup>	1.0028	-0.2439
4	1	-9.8602	-9.8031	-10.6562	-10.5591	0.51	0.7637	1.0424	-1.0587
5	1	<u>-9.8728</u>	<u>-9.8052</u>	-10.6597	-10.5450	0.50	0.3822	1.0568	-1.0665
1	2	-9.8014	-9.7600	-10.7777	-10.7071	0.49	-0.1303	1.0052	-0.8313
2	2	-9.8236	-9.7665	-10.7581	-10.6611	0.48	-1.3784 <sup>o</sup>	1.0073	-1.2517
3	2	-9.8342	-9.7616	-10.7553	-10.6318	0.51	1.3368 <sup>+</sup>	0.9937	0.6227
4	2	-9.8608	-9.7725	-10.6293	-10.4793	0.51	1.0474	1.0626	-1.5498 <sup>o</sup>
5	2	-9.8459	-9.7419	-10.6296	-10.4531	0.50	0.1114	1.0763	-2.3344 <sup>oo</sup>
1	3	-9.7956	-9.7386	-10.7736	-10.6766	0.47	-0.9654	1.0011	-0.2914
2	3	-9.8280	-9.7502	-10.6843	-10.5520	0.50	0.4292	1.0187	-1.2902 <sup>o</sup>
3	3	-9.8264	-9.7278	-10.6326	-10.4650	0.50	0.0703	1.0176	-0.9887
4	3	-9.8567	-9.7372	-10.5864	-10.3835	0.49	-0.1862	1.0221	-1.3750 <sup>o</sup>
5	3	-9.8618	-9.7214	-10.5824	-10.3443	0.50	0.2353	1.0894	-1.5717 <sup>o</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+, ++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>o, oo</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.



**Tableau XVII : Performance du réseau  $(p, q)$  - Prédiction des rendements quotidiens BP/USD avec signaux de vente-achat comme inputs,  $n_1 = 1, n_2 = 200$**

$$r_t = \phi(s_{t-p}(n_1, n_2), \theta) + \varepsilon_t$$

$p$	$q$	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8829	-9.8532	<u>-10.7897</u>	<u>-10.7456</u>	0.46	-1.5873 <sup>°</sup>	0.9949	0.5154
2	1	-9.8826	-9.8409	-10.7606	-10.6988	0.47	-0.9016	0.9928	0.5832
3	1	-9.9059	-9.8523	-10.7733	-10.6940	0.50	0.5021	0.9901	0.8520
4	1	-9.9181	-9.8524	-10.7264	-10.6293	0.53	1.8870 <sup>++</sup>	1.0158	-0.6522
5	1	<u>-9.9429</u>	<u>-9.8652</u>	-10.6595	-10.5449	0.50	0.4289	1.0739	-1.1970
1	2	-9.8774	-9.8298	-10.7793	-10.7088	0.47	-0.8864	0.9939	0.6903
2	2	-9.8936	-9.8281	-10.7359	-10.6389	0.47	-1.0606	1.0054	-0.3437
3	2	-9.8996	-9.8162	-10.7431	-10.6196	0.50	0.5940	0.9870	0.9369
4	2	-9.9278	-9.8263	-10.6462	-10.4963	0.50	0.4807	1.0142	-0.4796
5	2	-9.9148	-9.7953	-10.7052	-10.5288	0.49	0.0038	1.0568	-1.5668 <sup>°</sup>
1	3	-9.8776	-9.8121	-10.7454	-10.6484	0.49	-1.1114	0.9876	0.5184
2	3	-9.8964	-9.8071	-10.5873	-10.4550	0.50	0.4136	1.0384	-1.3887 <sup>°</sup>
3	3	-9.8818	-9.7685	-10.7400	-10.5724	0.48	-0.7946	1.0078	-0.7138
4	3	-9.9386	-9.8013	-10.5929	-10.3900	0.51	1.0672	1.1032	-1.3564 <sup>°</sup>
5	3	-9.9295	-9.7683	-10.6620	-10.4238	0.50	0.3195	1.0882	-2.2434 <sup>°°</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle  $(p, q)$  considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle  $(p, q)$  fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>°</sup>, <sup>°°</sup> : le modèle  $(p, q)$  fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

**Tableau XVIII : Performance du réseau (p,q) - Prédiction des rendements quotidiens BP/USD avec différentiels de taux d'intérêt comme inputs**

$$r_t = \phi(\Delta D_{t-p}, \theta) + \varepsilon_t$$

p	q	Estimation		Validation		Prédiction			
		AIC	BIC	AIC	BIC	SR	PT	MSPE	DM
1	1	-9.8533	<u>-9.8284</u>	<u>-10.8056</u>	<u>-10.7615</u>	0.48	-0.6009	1.1261	-1.9346 <sup>∞</sup>
2	1	-9.8356	-9.8008	-10.7920	-10.7303	0.50	-2.7050 <sup>∞</sup>	1.0794	-1.6144 <sup>°</sup>
3	1	-9.8690	-9.8242	-10.7603	-10.6809	0.52	1.4621 <sup>+</sup>	1.0494	-1.7482 <sup>∞</sup>
4	1	-9.8820	-9.8271	-10.7385	-10.6414	0.50	0.5671	1.0785	-2.1230 <sup>∞</sup>
5	1	-9.8809	-9.8160	-10.7363	-10.6216	0.51	0.7210	1.0816	-2.2454 <sup>∞</sup>
1	2	-9.8516	-9.8118	-10.7910	-10.7204	0.47	-1.5042 <sup>°</sup>	1.0882	-1.7364 <sup>∞</sup>
2	2	-9.8475	-9.7927	-10.7829	-10.6859	0.46	-1.5286 <sup>°</sup>	1.0837	-2.1796 <sup>∞</sup>
3	2	-9.8682	-9.7984	-10.7426	-10.6191	0.51	1.2661	1.1090	-2.1419 <sup>∞</sup>
4	2	-9.8809	-9.7962	-10.7191	-10.5691	0.48	-0.3117	1.1097	-2.5635 <sup>∞</sup>
5	2	-9.8666	-9.7668	-10.7194	-10.5429	0.51	-0.1434	1.0842	-2.3971 <sup>∞</sup>
1	3	-9.8479	-9.7932	-10.7821	-10.6850	0.48	-0.6063	1.1261	-1.7593 <sup>∞</sup>
2	3	-9.8539	-9.7792	-10.7414	-10.6091	0.46	-1.5467 <sup>°</sup>	1.1115	-2.3088 <sup>∞</sup>
3	3	-9.8636	-9.7689	-10.7188	-10.5512	0.50	1.4999 <sup>+</sup>	1.1399	-2.4151 <sup>∞</sup>
4	3	<u>-9.8923</u>	-9.7776	-10.7085	-10.5056	0.50	1.0874	1.2123	-2.9428 <sup>∞</sup>
5	3	-9.8912	-9.7564	-10.6607	-10.4226	0.49	0.4520	1.1329	-2.5637 <sup>∞</sup>

Note: AIC et BIC sont les critères de Akaike et Schwarz définis en (11) et (12) respectivement. Les valeurs minimums sont soulignées. SR représente le taux de succès défini en (13), PT est la statistique de Peseran et Timmermann (1992) définie en (14). MSPE est le rapport entre l'erreur quadratique du modèle (p,q) considéré et l'erreur quadratique moyenne de la Marche Aléatoire. DM représente la statistique de Diebold et Mariano (1995) définie en (17).

L'échantillon d'estimation s'étend de janvier 1991 à décembre 1994, l'échantillon de validation de janvier 1995 à décembre 1996 et l'échantillon de prédiction de janvier 1997 à décembre 1998.

<sup>+</sup>, <sup>++</sup> : le modèle (p,q) fait mieux que la Marche Aléatoire aux seuils de 10% et 5%.

<sup>°</sup>, <sup>∞</sup> : le modèle (p,q) fait moins bien que la Marche Aléatoire aux seuils de 10% et 5%.

C

C