UCL Université catholique de Louvain

LOUVAIN School of Management

## Louvain School of Management

# Prédiction de la direction des taux de change à court terme : comparaison de méthodes

Mémoire recherche réalisé par
**Aurian della Faille de Leverghem**

en vue de l'obtention du titre de
**Master en ingénieur de gestion, à finalité spécialisée**

Promoteur
**Marco Saerens**

Année académique 2016-2017

*Je voudrais remercier toutes les personnes qui ont contribué à la réalisation de mon mémoire.*

*Je tiens d'abord à remercier mon promoteur, Marco Saerens, pour ses conseils, ses attentes et ses encouragements.*

*Je remercie également Sophie Béreau, professeur à la Louvain School of Management, pour son aide précieuse et ses conseils avisés.*

*Je tiens enfin à remercier mes parents pour leurs précieux commenaires et leurs encouragements*

# Table des matières

# 1 Introduction

Depuis des dizaines d'années, l'anticipation des mouvements des marchés financiers est une thématique qui n'a cessé d'intéresser tant les milieux financiers qu'académiques. La littérature sur ce sujet témoigne d'une grande richesse grâce, entre autres, au nombre important de domaines scientifiques qui sont concernés par la prédiction financière comme la finance, l'économie, la statistique ou encore l'informatique mais aussi grâce à la diversité dans le choix des actifs financiers étudiés, de l'horizon de temps et des données utilisées qui peuvent être tirées de l'analyse technique, fondamentale, de l'historique des cours de l'actif ou encore de son flux d'ordres.

Ce mémoire analysera, en particulier, la prédiction intra journalière du marché des changes aussi appelé le Forex. Avec un volume journalier moyen de 5,14 trillions de dollars en 2014 [15], le Forex occupe une place capitale dans le système économique mondial ; une variation des taux de change se répercute non seulement sur les intervenants qui traitent directement les devises mais aussi sur une série d'autres acteurs économiques, de l'investisseur particulier qui souhaiterait investir dans des actions étrangères aux multinationales qui voudraient se couvrir de façon optimale face au risque de change.

Parmi toutes les paires de devises du marché des taux de change, l'euro-dollar (EUR/USD) et le dollar-yen (USD/JPY) seront les paires étudiées. Elles sont deux candidates idéales pour la prédiction financière étant donné qu'elles occupent respectivement la première et la deuxième place en termes de volume échangé [6] et qu'elles ont un régime de change flottant. Un taux de change à régime flottant signifie que le taux de change varie librement et est déterminé par les échanges sur le marché des changes.

Le prix des taux de change liés aux devises à régime flottant est uniquement déterminé par l'offre et la demande. Or, il n'est actuellement pas possible de calculer les courbes d'offre et de demande pour ces actifs, d'où l'intérêt de développer des méthodes de prédiction financière alternatives. Ce mémoire sera l'occasion de comparer les modèles économétriques traditionnellement utilisées pour la prévision de série temporelle à d'autres techniques de prédiction telles que les algorithmes d'apprentissage automatique.

## 1.1 Objectifs du mémoire

Ce mémoire aura pour vocation de développer différents modèles de prédiction de la direction de l'euro dollar et du dollar yen à très court terme puisque, toutes les 5 minutes, une nouvelle prédiction de la direction des taux de change sera donnée. Pour calculer cette prédiction, les modèles auront besoin d'*inputs*. Dans certains cas, les *inputs* ne seront dérivés que des valeurs antérieures des taux de change tandis que, dans d'autres cas, celles-ci s'accompagneront d'un *input* « effet de surprise », lié aux publications macroéconomiques américaines.

Afin de rendre l'analyse des résultats la plus objective possible, les prédictions des différents modèles seront comparées sur base des mêmes ensembles d'observations et en utilisant majoritairement les mêmes mesures de performances. La phase d'interprétation des résultats sera l'occasion de confronter ceux-ci à la théorie sur la prédiction financière, sur les modèles autorégressifs et sur l'apprentissage automatique pour les séries temporelles financières. Concernant la théorie sur la prédiction financière, le but sera de comparer nos résultats avec les hypothèses de marché efficient développées par Eugène Fama. Tandis que pour les modèles autorégressifs et d'apprentissage automatique, l'interprétation aura notamment pour but de déterminer si certaines méthodes utilisées se démarquent des autres et, si oui, d'en comprendre les raisons.

Au-delà de l'aspect théorique développé ci-dessus, ce mémoire est susceptible d'intéresser les intervenants sur le marché des changes. Le mémoire pourrait être considéré comme une première étape dans le développement d'une ou de stratégies d'opérations sur le marché des changes ; il ne s'agirait que d'une première étape car les questions de *money management* et de *risk management* ne seront pas développées; cependant, à la fin du mémoire, la question de la rentabilité potentielle des modèles de prédiction développés sera explorée, en tenant compte au mieux des contraintes physiques et de coûts engendrées par ces différents modèles.

# 2 Le marché des changes et la prévision financière

## 2.1 Le marché des changes

Un marché financier est un mécanisme qui permet à des individus du monde entier d'acheter et de vendre des actifs financiers ou des matières premières, de faciliter la levée de capitaux pour les marchés des capitaux, le transfert de risque pour les marchés de dérivés ou encore le commerce international pour les marchés des changes [18].

Ce mémoire se concentrera sur un marché en particulier, le marché des changes, aussi appelé **le Forex**. C'est un marché de gré à gré international où les devises sont négociées par paires, 24 heures sur 24, 5 jours sur 7. Il existe deux « marchés » qui permettent de négocier des devises : le marché au comptant, où les transactions sont réalisées via le système bancaire, grâce à l'intermédiaire des courtiers électroniques ou encore d'homme à homme, et le marché des produits dérivés, qui comprend notamment les contrats à terme et les options.

Lorsque les autorités monétaires d'une zone monétaire définissent le régime de change, elles ont le choix entre deux régimes : un régime de change fixe, comme pour le yuan, ou un régime de change flottant, comme pour l'euro. Le régime de change flottant, contrairement au régime de change fixe, est déterminé en fonction de l'offre et de la demande. Les prévisions de taux de change à régime flottant devraient donc idéalement se baser sur les courbes de l'offre et de la demande. Cependant, étant donné la taille du marché, il est impossible d'établir ces courbes, d'où l'intérêt de développer des techniques alternatives de prévision financière.

Parmi toutes les paires de devises à régime de change flottant, seulement deux taux de change seront étudiés dans ce mémoire : l'euro-dollar (EUR/USD) et le dollar-yen (USD/JPY). Ces paires sont des candidates idéales de par leur place prépondérante sur

le marché des changes puisqu'elles occupent respectivement la première et la deuxième place des paires les plus négociées [6]. Une partie des modèles, développés dans ce mémoire, utilisera des données relatives aux publications d'indicateurs économiques américains. Il est donc primordial que le dollar américain y soit présent, comme devise échangée, dans chacune des deux paires.

*Figure 1 - cotation Forex provenant du site www.dukascopy.com*


L'image ci-dessus illustre la manière dont les cotations des paires de devises sont représentées auprès d'un courtier. La première colonne indique les deux monnaies traitées. La devise de gauche s'appelle la devise de base et sa valeur est toujours égale à 1. La devise de droite est la devise négociée, c'est donc sa valeur qui est cotée. Les colonnes *Bid* et *Ask* représentent respectivement le prix auquel le marché, ici représenté par le courtier, est prêt à acheter ou vendre la devise de base. Pour cet exemple, le *bid* de l'EUR/USD est à 1.13207, soit le montant en dollar américain qu'il faut payer pour acheter un euro. La dernière colonne, le *Spread*, est simplement la différence entre le prix du *bid* et du *ask* et s'exprime en termes de *pips* ou encore points de base ; le *spread* est ici égal à 0,0001 dollar pour l'euro dollar et 0,01 yen pour le dollar yen.

Les courtiers du Forex, tels que **Dukascopy**, servent d'intermédiaires entre le marché de détail et le marché interbancaire. Ils permettent aux fonds d'investissements privés, aux sociétés commerciales et aux particuliers de réaliser des transactions sur le marché des changes. La majorité des courtiers se rémunèrent grâce à un *markup,* c'est-à-dire un élargissement du *spread bid-ask* par rapport au *spread* interbancaire, et/ou grâce aux frais proportionnels aux volumes échangés par le client. Dukascopy n'a pas

de *markup*, les *spreads* fournis par ce courtier sont donc les *spreads* « interbancaires » liés aux prix proposés par ses fournisseurs de liquidités, qu'ils soient des banques, d'autres courtiers ou des fournisseurs de liquidités internes [12]. Dukascopy se rémunère donc uniquement via des commissions sur le volume de devises échangées. Pour un volume de devises échangées à hauteur d'un million de dollars US, le meilleur tarif sera de 10 dollars [13]. Cette information sera essentielle pour pouvoir déterminer si les méthodes de prédiction développées dans ce mémoire sont rentables, une fois les coûts intégrés.

## 2.2 La prévisibilité des marchés financiers

La prévisibilité du prix d'un actif financier, que ce soit une action, un taux de change, une obligation ou encore une matière première, reste aujourd'hui une des questions centrales du monde financier. Cette notion de prévisibilité est étroitement liée à l'hypothèse d'efficience du marché financier développée dans les années 70 par E. Fama.

Dans la lignée des recherches d'E. Fama, Michael C. Jensen donne une nouvelle définition de l'efficience d'un marché qui correspond mieux au sujet de ce mémoire. Il définit un marché efficient par rapport à un ensemble d'informations $X_t$ comme un marché où il serait impossible de générer un profit économique sur la base de l'ensemble $X_t$ [22].

Trois formes de marché efficient peuvent être distinguées sur base des informations contenues dans $X_t$ :

- La forme faible de l'efficience dans le cas où $X_t$ comprend seulement les prix historiques de l'actif.
- La forme semi-forte lorsque l'information contenue dans $X_t$ est étendue à toutes les informations qui sont publiquement disponibles.
- La forme forte de l'efficience pour un ensemble $X_t$ qui comprendrait non seulement toutes les informations publiquement disponibles mais aussi l'information privée.

Si un marché venait à être inefficient pour $X_t$, cela reviendrait à dire que celui-ci ne réagit pas instantanément et complètement aux nouvelles informations. Au lieu de cela, il pourrait réagir de manière excessive ou insuffisante puis graduellement s'ajuster [29].

## 2.3 Les indicateurs économiques

Les indicateurs économiques ou macro-économiques sont des données statistiques qui ont pour but de montrer la tendance générale de l'économie d'un pays ou d'une zone économique et/ou monétaire [26]. Chaque indicateur peut être classé selon 3 caractéristiques :

- Sa relation par rapport au cycle économique. Un indicateur pourra être procyclique, anticyclique ou encore acyclique s'il évolue respectivement dans le même sens que l'économie, dans la direction opposée ou n'a pas de corrélation avec la santé de l'économie. Le PIB est un indicateur procyclique tandis que le taux de chômage est un indicateur anticyclique.
- La fréquence de publication des données. Dans la plupart des pays, les chiffres du PIB sont communiqués tous les trimestres alors qu'un indicateur, comme l'*initial jobless claims* pour les Etats-Unis, sera publié chaque semaine.
- Le timing de ses variations par rapport aux changements de l'économie concernée. Les premiers indicateurs sont « les indicateurs avancés », tels que l'*unemployment insurance claim* américain, qui varient avant que l'économie ne change. La seconde catégorie est celle des « indicateurs retardés ». Les variations d'un indicateur retardé suivent celles de l'économie. Finalement, si un indicateur n'est ni avancé, ni retardé, il est appelé « coïncident », c'est-à-dire qu'il varie au même moment que l'économie.

Même si les indicateurs macroéconomiques peuvent avoir des caractéristiques différentes et des unités de mesure différentes, il y a une constance dans les données disponibles. Dans le tableau ci-dessous, les indicateurs sont regroupés en un calendrier avec,

pour chacun d'eux, la date, l'heure, le nom de l'indicateur, la période, l'estimation, la donnée réelle une fois disponible, la donnée précédente et l'éventuelle révision.

*Figure 2 - Calendrier économique Bloomberg*


| | Date Time | A | M | R | Event | Period | Surv(M) | Actual | Prior | Revised |
|---|---|---|---|---|---|---|---|---|---|---|
| 21) | 01/13 14:00 | | | | Monthly Budget Statement | Dec | $44.0B | -- | -$135.2B | -- |
| 22) | 01/14 07:30 | | | | NFIB Small Business Optimi... | Dec | 93.2 | -- | 92.5 | -- |
| 23) | 01/14 08:30 | | | | Retail Sales Advance MoM | Dec | 0.1% | -- | 0.7% | -- |
| 24) | 01/14 08:30 | | | | Retail Sales Ex Auto MoM | Dec | 0.4% | -- | 0.4% | -- |
| 25) | 01/14 08:30 | | | | Retail Sales Ex Auto and Gas | Dec | 0.3% | -- | 0.6% | -- |
| 26) | 01/14 08:30 | | | | Retail Sales Control Group | Dec | 0.3% | -- | 0.5% | -- |
| 27) | 01/14 08:30 | | | | Import Price Index MoM | Dec | 0.4% | -- | -0.6% | -- |
| 28) | 01/14 08:30 | | | | Import Price Index YoY | Dec | -0.6% | -- | -1.5% | -- |
| 29) | 01/14 10:00 | | | | Business Inventories | Nov | 0.3% | -- | 0.7% | -- |
| 30) | 01/15 07:00 | | | | MBA Mortgage Applications | Jan 10 | -- | -- | 2.6% | -- |
| 31) | 01/15 08:30 | | | | Empire Manufacturing | Jan | 3.50 | -- | 0.98 | -- |
| 32) | 01/15 08:30 | | | | PPI MoM | Dec | 0.4% | -- | -0.1% | -- |
| 33) | 01/15 08:30 | | | | PPI Ex Food and Energy MoM | Dec | 0.1% | -- | 0.1% | -- |
| 34) | 01/15 08:30 | | | | PPI YoY | Dec | 1.1% | -- | 0.7% | -- |
| 35) | 01/15 08:30 | | | | PPI Ex Food and Energy YoY | Dec | 1.3% | -- | 1.3% | -- |
| 36) | 01/15 14:00 | | | | U.S. Federal Reserve Releases Beige Book | | | | | |
| 37) | 01/16 08:30 | | | | CPI MoM | Dec | 0.3% | -- | 0.0% | -- |
| 38) | 01/16 08:30 | | | | CPI Ex Food and Energy MoM | Dec | 0.1% | -- | 0.2% | -- |
| 39) | 01/16 08:30 | | | | CPI YoY | Dec | 1.5% | -- | 1.2% | -- |

*source: http://holowczak.com/bloomberg-essentials-line-training-program/5/*

### 2.3.1 Les indicateurs économiques américains

L'étude se limitera aux indicateurs américains étant donné que le mémoire n'utilisera que ceux-ci, mais le raisonnement de cette section et de la prochaine section peut s'étendre à d'autres économies. Les indicateurs sont classés, dans le tableau ci-dessous, suivant la méthodologie utilisée par Andersen *et al.* (2003) ; ils sont d'abord classés par fréquence de publication, puis par catégorie et, dans chacune de ces catégories, les annonces sont classées par ordre chronologique de publication.

*Tableau 1 - indicateurs macroéconomiques américains*

| Announcement | Obs.[1] | Source[2] | Dates[3] | Announcement Time[4] |
|---|---|---|---|---|
| **Quarterly Announcements** | | | | |
| 1- GDP Advance | 51 | BEA | 01/92-12/02 | 8:30 |
| 2- GDP Preliminary | 50 | BEA | 01/92-12/02 | 8:30 |
| 3- GDP Final | 51 | BEA | 01/92-12/02 | 8:30 |
| **Monthly Announcements** | | | | |
| Real Activity | | | | |
| 4- Nonfarm Payroll Employment | 194 | BLS | 01/92-12/02 [5] | 8:30 |
| 5- Retail Sales | 193 | BC | 01/92-12/02 | 8:30 |
| 6- Industrial Production | 193 | FRB | 01/92-12/02 | 9:15 |
| 7- Capacity Utilization | 177 | FRB | 01/92-12/02 | 9:15 |
| 8- Personal Income | 192 | BEA | 01/92-12/02 [6] | 10:00/8:30 [7] |
| 9- Consumer Credit | 178 | FRB | 01/92-12/02 | 15:00[8] |
| Consumption | | | | |
| 10- New Home Sales | 167 | BC | 01/92-12/02 | 10:00 |
| 11- Personal Consumption Expenditures | 192 | BEA | 01/92-12/02 [9] | 10:00/8:30[10] |
| Investment | | | | |
| 12- Durable Goods Orders | 237 | BC | 01/92-12/02 [11] | 8:30/9:00/10:00[12] |
| 13- Factory Orders | 178 | BC | 01/92-12/02 [13] | 10:00 |
| 14- Construction Spending | 177 | BC | 01/92-12/02 [14] | 10:00 |
| 15- Business Inventories | 177 | BC | 01/92-12/02 | 10:00/8:30[15] |
| Government Purchases | | | | |
| 16- Government Budget | 175 | FMS | 01/92-12/02 [16] | 14:00 |
| Net Exports | | | | |
| 17- Trade Balance | 192 | BEA | 01/92-12/02 | 8:30 |
| Prices | | | | |
| 18- Producer Price Index | 193 | BLS | 01/92-12/02 | 8:30 |
| 19- Consumer Price Index | 275 | BLS | 01/92-12/02 | 8:30 |
| Forward-Looking | | | | |
| 20- Consumer Confidence Index | 138 | CB | 01/92-12/02 | 10:00 |
| 21- NAPM Index | 156 | NAPM | 01/92-12/02 | 10:00 |
| 22- Housing Starts | 269 | BC | 01/92-12/02 | 8:30 |
| 23- Index of Leading Indicators | 275 | CB | 01/92-12/02 | 8:30 |
| **Six-Week Announcements** | | | | |
| FOMC | | | | |
| 24- Target Federal Funds Rate | 175 | FRB | 01/92-12/02 | 14:15[17] |
| **Weekly Announcements** | | | | |
| 25- Initial Unemployment Claims | 600 | ETA | 01/92-12/02 | 8:30 |

*source: Andersen et al. [4]*

### 2.3.2 L' « effet de surprise » des nouvelles macroéconomiques

Les annonces de nouvelles sur les indicateurs économiques peuvent donner lieu à des ajustements de prix de divers actifs financiers, différents intervenants réexaminant leur vision de l'économie. Plusieurs études se sont intéressées aux « effets de surprise » de ces indicateurs sur différents actifs et, en particulier, sur les paires de devises. Ces recherches proposent différentes méthodes de sélection des indicateurs macroéconomiques qui peuvent avoir un impact mesurable sur les taux de change. Les indicateurs utilisés par la suite pour créer une variable explicative exogène seront sélectionnés sur base des travaux d'Andersen et *al* .

Dans ces travaux, ils tentent de déterminer si les mouvements intra journaliers des taux de change sont liés aux fondamentaux et si oui, de quelle manière.

Pour ce faire, ils utilisent le modèle linéaire suivant :

$$R_t = \beta_k S_{kt} + \varepsilon_t$$

avec,

$R_t$, le rendement sur 5 minutes d'un taux de change entre t et t+1, par exemple l'eur/usd

$\beta_k$, le coefficient du modèle

$\varepsilon_t$, le terme d'erreur

$S_{kt}$, la « nouvelle » standardisée de l'annonce pour un indicateur économique k au temps t :

$$S_{kt} = \frac{A_{kt} - E_{kt}}{\widehat{\sigma_k}}$$

où $A_{kt}$ est la valeur annoncée de l'indicateur k en t, $E_{kt}$ est la valeur attendue par le marché pour l'indicateur k en t et $\widehat{\sigma_k}$ est l'écart-type de $A_{kt} - E_{kt}$ pour l'échantillon disponible.

Sur base des résultats du modèle, Andersen et *al.* (2003, pp. 48-50) constatent que seuls les chocs fondamentaux inattendus affectent les taux de change, conformément aux prédictions de la théorie sur l'anticipation rationnelle. Ces chocs se font généralement par des ajustements très rapides des taux de change.

Les deux études démontrent également que les nouvelles macroéconomiques favorables à la croissance américaine entraînent, dans leur grande majorité, une appréciation du dollar.

Enfin, les coefficients statistiquement significatifs sont essentiellement ceux des indicateurs dont les publications paraissent le plus tôt. D'après leurs résultats, 11 indicateurs auraient un impact significatif sur les taux de change : Dix d'entre eux avec un coefficient $\beta_k$ positifs par rapport au dollar[1] et un seul avec un coefficient négatif[2].

[1] Le PIB avancé, le taux d'emploi non agricole, les ventes au détail, la production industrielle, les commandes de biens durables, l'indice manufacturier de l'*Institute of Supply Management*, les dépenses de construction, la balance commerciale, l'indice de confiance des consommateurs et le taux objectif des fonds fédéraux.

[2] Le nombre de demandes initiales d'assurance chômage

# 3 L'économétrie des séries temporelles

Une équation aux différences exprime, dans sa forme la plus générale, la valeur actuelle d'une variable comme une fonction de ces observations passées. L'économétrie des séries temporelles cherche à estimer ces équations aux différences à l'aide de méthodes de prévision adaptées aux caractéristiques des séries étudiées [14].

Les sections suivantes introduiront les concepts théoriques nécessaires à la compréhension des modèles économétriques examinés dans ce mémoire ; un « modèle moyenne » lié à la marche aléatoire et un modèle autorégressif à moyennes mobiles intégré (noté ARIMA pour *AutoRegressive Integrated Moving Average*).

## 3.1 La stationnarité

La stationnarité est un concept clé de l'analyse des séries temporelles. Il existe deux définitions de la stationnarité : la stationnarité forte ou la stationnarité faible, moins contraignante que la première. Lorsque le terme de « stationnarité » sera évoqué dans ce mémoire, il fera référence à la stationnarité faible.

Une série $\{r_t\}$ suit un processus (faiblement) stationnaire si sa moyenne, sa variance et la covariance entre $\{r_t\}$ et $\{r_{t-l}\}$ sont constantes dans le temps, avec $t$ et $l \in \mathbb{N}$ [33]

Dans la réalité, il est impossible de savoir avec certitude si une série temporelle est générée par un processus stationnaire. Mais, afin de savoir si une série semble suivre un processus stationnaire, plusieurs outils peuvent être utilisés tels que : l'analyse graphique, la fonction d'autocorrélation (notée ACF pour *AutoCorrelation Function*) et le test de Dickey-Fuller augmenté.

### 3.1.1 La fonction d'autocorrélation (ACF)

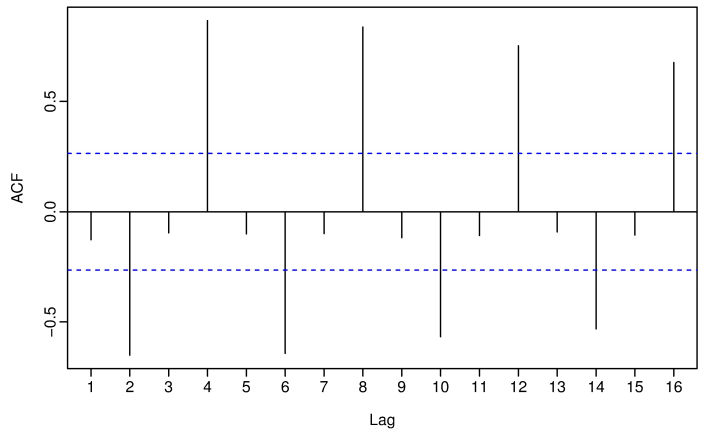
Alors que la corrélation mesure la dépendance linéaire entre deux variables, l'autocorrélation mesure la dépendance linéaire entre les valeurs décalées d'une série temporelle [5]. Pour une série $\{r_t\}$, le coefficient d'autocorrélation $\rho_l$ mesure la corrélation entre $\{r_t\}$ et les valeurs passées $\{r_{t-l}\}$ et il a pour formule :

$$\rho_l = \frac{\sum_{t=l+1}^{T}(r_t - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=1}^{T}(r_t - \bar{r})^2}$$

où $\bar{r}$ est la moyenne de la série et $T$ le nombre d'observations dans la série temporelle.

Pour faciliter l'interprétation des coefficients d'autocorrélation, ceux-ci sont généralement représentés sous la forme d'un corrélogramme appelé fonction d'autocorrélation.

*Figure 3 - Exemple d'ACF*




Une des applications des fonctions d'autocorrélation est l'identification de bruits blancs. Une séquence de données $\{\varepsilon_t\}$ est un processus appelé bruit blanc si chacune de ses valeurs a une moyenne nulle, une variance constante et n'est corrélée à aucune autre réalisation de cette séquence [14].

Théoriquement, toutes les « barres » du corrélogramme d'un bruit blanc devraient être égales à 0. En pratique, il y aura toujours des variations aléatoires qui feront que les

autocorrélations ne seront pas nulles mais proches de 0. Si la fonction d'autocorrélation a une ou plusieurs grandes « barres » ou si plus de 5% de ces « barres » sont situées au-delà des valeurs critiques $\pm 1{,}96/\sqrt{T}$ (où $T$ est le nombre de valeurs dans la séquence), alors il y a de fortes chances que la série temporelle ne soit pas un bruit blanc [5].

L'analyse d'une fonction d'autocorrélation donne également des informations sur la probabilité qu'une série temporelle suive un processus stationnaire. L'ACF d'une série temporelle qui suit un processus stationnaire aura tendance à chuter rapidement vers 0 tandis qu'elle décroîtra lentement pour un processus non-stationnaire [5].

### 3.1.2 Le test de Dickey-Fuller augmenté

Le test de Dickey-Fuller augmenté fait partie de la famille des tests de racine unitaire. Ce sont des tests d'hypothèses de stationnarité conçus pour déterminer si une série doit être différenciée. Avant de décrire plus précisément le test de Dickey-Fuller augmenté, il est judicieux de définir ce qu'est une différenciation, une racine unitaire et la stationnarité en différence.

Différencier une série temporelle non stationnaire c'est la transformer en calculant les différences entre ses observations consécutives. Une série peut être différenciée plusieurs fois avant de devenir stationnaire. La première série différenciée $\{c_t\}$ issue de la série $\{r_t\}$ aura pour observations $c_t = r_t - r_{t-1}$. Une série est considérée comme stationnaire en différence et avec une racine unitaire si la série obtenue après la première différenciation est stationnaire.

Pour en revenir au test de Dickey-Fuller augmenté, celui-ci teste l'hypothèse nulle qu'une série temporelle $\{r_t\}$ ait une racine unitaire. Une fois la longueur de décalage déterminée, la régression liée au test est estimée à l'aide d'une méthode des moindres

carrés ordinaire. Le test se termine par le calcul de la statistique ADF comparée à une valeur critique. Si la statistique du test est plus faible que la valeur critique alors l'hypothèse nulle est rejetée. En d'autres mots, si la p-valeur obtenue à la suite du test est plus petite qu'un certain niveau (habituellement 5%) alors la série temporelle n'a pas de racine unitaire.

## 3.2 Le modèle de marche aléatoire

Le modèle de marche aléatoire a été popularisé par une publication de Meese et Rogoff (1982) dans laquelle le modèle de marche aléatoire enregistre des performances supérieures à celles des autres modèles étudiés reposant sur des données économiques. Le modèle suppose qu'à chaque période, le processus suive un « pas » aléatoire depuis sa valeur précédente, les pas étant indépendants et identiquement distribués en taille [27]. Ce n'est donc pas la série en elle-même qui est aléatoire mais les variations d'une période à l'autre. Il existe deux types de modèle de marche aléatoire : avec dérive ou sans dérive.

Le modèle de marche aléatoire sans dérive suppose que les « pas » suivent une distribution de moyenne nulle. Pour un taux de change, cela revient à dire que son prix varie selon l'équation aux différences stochastiques :

$$y_t = y_{t-1} + \varepsilon_t$$

où $y_t$ est le taux de change au moment t et où la séquence des termes d'erreur $\{\varepsilon_t\}$ est un bruit blanc.

Pour le modèle de marche aléatoire avec dérive, les pas suivent non plus une distribution de moyenne nulle mais une distribution de moyenne non nulle. Cela se répercute sur l'équation par l'ajout d'un terme constant $a_0$ :

$$y_t = y_{t-1} + \varepsilon_t + a_0$$

Etant donné la condition initiale $y_0$, il est possible de réécrire l'équation ci-dessus comme :

$$y_t = y_0 + \ a_0 t + \sum_{i=1}^{t} \varepsilon_t$$

Le comportement de $y_t$ est donc dicté par deux composantes non stationnaires : une tendance linéaire et une tendance stochastique $\sum_{i=1}^{t} \varepsilon_t$.

Lorsque le modèle de marche aléatoire est utilisé comme *benchmark*, les prévisions sont en réalité calculées grâce à une méthode dite « naïve » basée sur le concept de marche aléatoire. Pour une marche aléatoire sans dérive, la méthode consiste simplement à prendre pour prévision la dernière observation enregistrée. Tandis que, pour une marche aléatoire avec dérive, la méthode naïve revient à prendre la dernière observation tout en y ajoutant une dérive (négative ou positive), dérive qui est égale à la moyenne des variations historiques.

Dans le domaine de la prévision des séries financières, le modèle de marche aléatoire et la méthode naïve sont régulièrement associés aux logarithmes du prix d'un actif. Lorsque la série étudiée n'est plus celle des prix d'un actif mais la série différenciée, le *benchmark* lié au modèle de marche aléatoire à utiliser n'est plus la méthode naïve mais une méthode « de la moyenne ». Avec cette méthode, les valeurs des prévisions sont égales à la moyenne des données historiques [5].

## 3.3 Le modèle ARIMA

Le modèle ARIMA est un des modèles économétriques les plus utilisés pour la prévision de séries temporelles. Il est composé d'une partie autorégressive (AR), d'une partie moyenne mobile (MA) et d'un certain niveau d'intégration (I).

### 3.3.1 ARMA (p,q)

Dans un modèle autorégressif univarié, la variable s'exprime sous la forme d'une combinaison linéaire de ses valeurs passées. Un modèle autorégressif d'ordre $p$ a pour équation :

$$AR(p): \quad y_t = c + \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \varepsilon_t$$

où $c$ est une constante, $\alpha_1, \ldots, \alpha_p$ sont les paramètres du modèle et $\varepsilon_t$, le terme d'erreur, est un bruit blanc.

Le modèle $AR(p)$ est donc un cas spécial de modèle régressif linéaire dans lequel la variable dépend de ses valeurs passées. Une fois $p$ fixé, les coefficients $\alpha_1, \ldots, \alpha_p$ seront facilement estimables à l'aide, par exemple, de la méthode des moindres carrés ordinaire.

Dans la forme, l'équation du modèle de moyenne mobile ressemble à celle du modèle autorégressif si ce n'est qu'au lieu d'utiliser une combinaison des valeurs passées, la formule du modèle moyenne se base sur les termes d'erreurs passées :

$$MA(q): \quad y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

où $\theta_1, \ldots, \theta_p$ sont les paramètres du modèle et $\varepsilon_t$ un bruit blanc.

Le modèle $MA(q)$ repose sur les termes d'erreur dont la valeur, non observable, dépend des coefficients du modèle. Les coefficients $\theta_1, \ldots, \theta_p$ devront donc être estimés à l'aide de méthodes d'optimisation non linéaire [27].

Le modèle $ARMA(p, q)$ n'est autre qu'une combinaison des modèles $AR(p)$ et $MA(q)$ dont l'équation générale serait :

$$ARMA(p,q): \quad y_t = c + \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Lorsqu'un modèle $ARMA(p,q)$ est utilisé pour estimer les valeurs futures d'une série temporelle, le modèle repose, entre autres, sur l'hypothèse que la série est (faiblement) stationnaire.

### 3.3.2 ARIMA (p,d,q) et la différenciation

Le modèle ARIMA est une généralisation du modèle ARMA qui permet de travailler avec des séries temporelles qui ne seraient pas stationnaires, mais pourraient être rendues stationnaires en utilisant la différenciation.

Pour faciliter la mise en équation du modèle ARIMA, le concept d'opérateur retard, noté $B$, est introduit :

$$By_t = y_{t-1} \quad \text{ou encore} \quad B(By_t) = B^2 y_t = y_{t-2}$$

Il devient dès lors très facile de représenter le processus de différenciation. Par exemple, la différenciation d'ordre 2 d'une série temporelle $y_t$ donne :

$$y_t - 2y_{t-1} + y_{t-2} = (1 - 2B + B^2) y_t = (1 - B)^2 y_t$$

L'équation représentant le modèle ARIMA est donc :

$$ARIMA(p,d,q): \quad (1 - \alpha_1 B - \ldots - \alpha_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \ldots + \theta_q B^q)\, \varepsilon_t$$

où $p, d, q$ sont respectivement l'ordre de la partie autorégressive, le degré de différenciation et l'ordre de la partie moyenne mobile. $\alpha_1, \ldots, \alpha_p$ et $\theta_1, \ldots, \theta_p$ sont les paramètres de la partie autorégressive et moyenne mobile et $\varepsilon_t$ est un bruit blanc.

# 4 L'apprentissage automatique

L'apprentissage automatique (*machine learning* en anglais) est une branche de l'informatique appliquée qui tend à construire des systèmes algorithmiques capables d'apprendre et de faire des prédictions sur des données.

On distingue traditionnellement deux types d'apprentissage automatique : l'apprentissage supervisé et l'apprentissage non supervisé. En apprentissage supervisé, l'objectif est de modéliser les relations causales entre un ensemble de variables explicatives et une variable à expliquer à partir d'un échantillon d'observations $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ d'entrées $X_i$ et de sorties $Y_i$ pour ensuite prédire de nouvelles sorties à partir des nouvelles entrées correspondantes. Tandis que pour l'apprentissage non supervisé, il n'y a pas de sortie prédéfinie. Les algorithmes cherchent alors à construire un modèle permettant de représenter les observations $X_1, \dots, X_n$ dans des groupes les plus homogènes possibles [16].

Le problème de prédiction de la direction des taux de change étudié dans ce mémoire est un problème d'apprentissage supervisé dans lequel les variables explicatives sont discrètes et continues et la variable à expliquer est binaire. Parmi les algorithmes capables de travailler sur ce type de problème et avec ce genre de données, deux ont été sélectionnés pour cette recherche : la régression logistique et le perceptron multicouche. Ce chapitre explore les aspects théoriques nécessaires à la compréhension de ces modèles.

## 4.1 Régression logistique

### 4.1.1 Description

La régression logistique est une technique de modélisation régressive dans laquelle la variable dépendante est discrète. Elle est utilisée dans de nombreux domaines tels que

la recherche biomédicale, l'écologie, les politiques sanitaires ou encore la finance. L'objectif de la régression logistique est de fournir le modèle le plus adapté pour décrire au mieux et avec parcimonie la relation entre la variable dépendante discrète et un ensemble de variables explicatives [20].

Etant donné que seul le cas d'une variable dépendante binaire sera appliqué dans ce mémoire, la théorie se focalisera uniquement sur la régression logistique binaire.

Considérons un problème de classification où la variable dépendante $Y$ a deux classes possibles {0,1} séparant la population en deux groupes distincts et où l'ensemble de variables explicatives est $X = (X_1, ..., X_p)$.

La régression logistique binaire cherche à expliquer $Y$ à partir de $X$. Plus spécifiquement, la régression logistique est utilisée pour estimer, à l'aide d'un ensemble d'observations, la probabilité inconnue qu'un individu $\omega$ vérifiant $x = X$ appartienne au groupe pour lequel $Y = 1$ :

$$p(x) = \mathbb{P}(\{Y = 1\}|\{X = x\}) \qquad x = (x_1, ..., x_p)$$

La probabilité que l'individu $\omega$ appartienne à la classe $Y = 0$ sera alors égale à la probabilité a posteriori qu'il ait comme valeur prédite $Y = 1$, ce qui donne :

$$\mathbb{P}(\{Y = 0\}|\{X = x\}) = 1 - p(x) = 1 - \mathbb{P}(\{Y = 1\}|\{X = x\}) \qquad x = (x_1, ..., x_p)$$

Comme l'explique Chesneau (2017), la probabilité a posteriori $p(x)$ ne peut être exprimée sous la forme d'une combinaison linéaire des variables explicatives de la même manière qu'avec une régression classique :

$$p(x) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

En effet, rien qu'en comparant les bornes des deux membres de l'équation ci-dessus, il y a une incohérence: $p(x) \in [0,1]$ alors que $\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$ est a priori sans bornes.

Pour formuler $p(x)$ comme une fonction de $\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$, le modèle de régression logistique passe par une transformation logit définie comme :

$$logit(y) = \ln(\frac{y}{1-y}) \in \mathbb{R}, \qquad y \in ]0,1[.$$

Formellement, le modèle de régression logistique a pour équation :

$$g(x) = logit\big(p(x)\big) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

où $\beta_0, \ldots, \beta_p$ sont des coefficients réels inconnus

$g(x)$ rassemble un grand nombre de propriétés souhaitées pour un modèle de régression linéaire. L'équation est linéaire en ses paramètres, peut être continue et peut être non bornée selon l'intervalle de $x$ [20].

Une fois $p(x)$ isolée, l'expression donne :

$$p(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots \beta_p x_p)}$$

Cette expression correspond à une fonction logistique $f\colon \mathbb{R} \to [0,1]$ définie par :

$$f(u) = \frac{\exp(u)}{1 + \exp(u)}$$

### 4.1.2 Régression logistique et classification

La partie précédente a présenté le modèle de régression logistique et a montré que celui-ci permet d'estimer la probabilité a posteriori qu'un individu appartienne à un certain groupe, chaque groupe étant représenté par une valeur de la variable à expliquer.

Seule, la régression logistique n'est donc pas un algorithme de classification. Elle sert d'algorithme de classification uniquement en la combinant avec une règle de décision.

Après avoir calculé les estimations $\hat{\beta}_0, \hat{\beta}_{01}, \dots, \hat{\beta}_p$ des coefficients $\beta_0,\ \beta_1 \dots, \beta_p$ dans le cas d'une classification binaire où la variable à expliquer $Y$ est égale à 0 ou 1, l'*output* d'un point d'observation $x = (x_1, \dots, x_p)$ sera donné par la règle de décision suivante :

$$Y = \begin{cases} 1 & si\ \ \hat{\beta}_0 + \ \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p > 0 \\ 0 & si\ \ \hat{\beta}_0 + \ \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p \leq 0 \end{cases}$$

Ramenée en probabilité a posteriori, la règle de décision ci-dessus signifie que $Y = 1$ si $p(x) > 0{,}5$ et que $Y = 0$ si $p(x) \leq 0{,}5$.

### 4.1.3 Estimation des coefficients

Cette partie décrit la méthode utilisée dans le cas d'une régression logistique pour estimer la suite des coefficients $\beta = (\ \beta_0,\ \beta_1 \dots, \beta_p)$ à partir d'un ensemble de $n$ observations $(x_i, y_i)$, $i = 1{,}2, \dots, n$, où $y_i$ désigne la valeur de la variable dépendante binaire $Y$ et où $x_i{=}(x_{1i}, \dots, x_{pi})$ désigne la série de valeurs des variables explicatives pour l'observation $i$.

Les modèles de régression logistique sont estimés à l'aide de l'estimateur du maximum de vraisemblance. Cette méthode consiste à chercher les paramètres qui maximisent la vraisemblance des $n$ observations disponibles [20]. Sachant que la vraisemblance est égale à :

$$L(\beta) = \prod_{i=1}^{n} p(x_i)^{y_i} \left[1 - p(x_i)\right]^{1-y_i}$$

Il est plus facile de travailler avec le log de l'expression ci-dessus et les paramètres qui maximisent la log-vraisemblance sont égaux à ceux qui maximisent la vraisemblance. Une fois la vraisemblance convertie en log-vraisemblance, l'expression donne :

$$l(\beta) = \sum_{i=1}^{n} y_i \log\big(p(x_i)\big) + (1 - y_i) \log([1 - p(x_i)])$$

avec les vecteurs $\mathbf{x_i} = \begin{bmatrix} 1 \\ x_{1i} \\ x_{2i} \\ \vdots \\ x_{pi} \end{bmatrix}$ et $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$, la formule de log-vraisemblance devient :

$$l(\beta) = \sum_{i=1}^{n} y_i \boldsymbol{\beta}^T \mathbf{x_i} - \log(1 + e^{\boldsymbol{\beta}^T \mathbf{x_i}})$$

S'il existe, l'estimateur du maximum de vraisemblance $\hat{\beta}$= $(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)$ est une solution de l'équation du score :

$$\nabla\ l(\beta) = \frac{\partial\ l(\beta)}{\partial \beta} = 0$$

où $\nabla\ l(\beta)$ est le vecteur gradient au point $\beta$ défini par $\nabla\ l(\beta) = \left[\frac{\partial\ l(\beta)}{\partial \beta_0} \quad \cdots \quad \frac{\partial\ l(\beta)}{\partial \beta_p}\right]^T$ sachant que :

$$\frac{\partial\ l(\beta)}{\partial \beta_j} = \sum_{i=1}^{n} x_{ji} \left(y_i - p(x_i)\right)$$

ce qui donne comme équation du score :

$$\nabla\ l(\beta) = \sum_{i=1}^{n} \mathbf{x_i} \left(y_i - p(x_i)\right) = 0$$

Il ne sera pas possible de résoudre analytiquement ce système d'équations non linéaires en $\beta$ [10]. Cependant les valeurs de $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$ pourront être approchées à l'aide de

méthodes numériques telles que l'algorithme IRLS (pour *Iterative Reweighted Least Squares* ou en français, Moindres Carrés Pondérés Itérés).

L'algorithme IRLS fonctionne par itération. Lors de chaque itération, l'estimateur $\hat{\beta}=(\hat{\beta}_0,\hat{\beta}_1,\dots,\hat{\beta}_p)$ est réévalué jusqu'à convergence. L'algorithme se résume de la manière suivante :

1. Choix d'un point de départ $\hat{\beta}^0$
2. Pour $k = 0,1,2,\dots,$
   a. Calcul de $\hat{\beta}^{k+1}$ à partir de $\hat{\beta}^k$

$$\hat{\beta}^{k+1} = \hat{\beta}^k + A^k \nabla\, l\left(\hat{\beta}^k\right)$$

   où $\nabla\, l\left(\hat{\beta}^k\right)$ est le gradient au point $\hat{\beta}^k$ et $A^k = -\left(\nabla^2\, l\left(\hat{\beta}^k\right)\right)^2$ est la matrice de « pas » de l'algorithme (l'inverse du hessien de $l$ au point $\hat{\beta}^k$) [30].

   b. Si $\hat{\beta}^{k+1} \approx \hat{\beta}^k$ et/ou $l\left(\hat{\beta}^{k+1}\right) \approx l\left(\hat{\beta}^k\right)$ alors il y a convergence.

   Dans le cas contraire, on passe à l'itération $k+1$ et le point 2 est répété

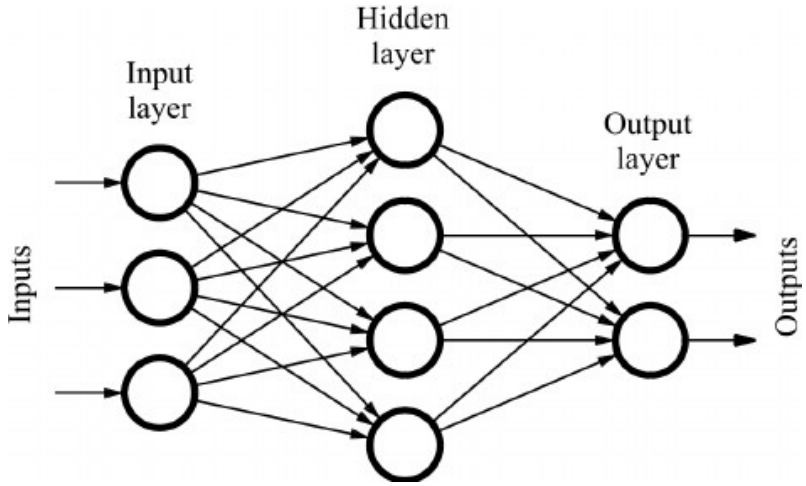
## 4.2 Perceptron multicouche

### 4.2.1 Description

Le perceptron multicouche fait partie de la famille des réseaux de neurones artificiels. Un réseau de neurones artificiels est un modèle non linéaire de calcul basé sur la structure et la fonction d'un réseau de neurones biologiques. L'information qui circule au sein du réseau affecte la structure de l'ANN qui « apprend » sur base des entrées et sorties de données [32] . Les réseaux de neurones artificiels ont été appliqués dans de nombreux domaines tels que la reconnaissance de caractères, la compression d'images ou la prédiction d'actifs financiers.

Un perceptron multicouche est composé de plusieurs couches contenant des nœuds. La couche inférieure est la couche d'entrée dans laquelle les données extérieures sont reçues. La couche supérieure est la couche de sortie, couche qui émet la prédiction du réseau. Entre ces deux couches, il peut y avoir une ou plusieurs couches intermédiaires, appelées les couches cachées.

*Figure 4 - Représentation d'un Percpetron multicouche avec une couche cachée*




source: Davim [11]

Les connexions entre neurones leur permettent de se transmettre des signaux alors que l'information est traitée. Le perceptron multicouche est un réseau *feed-forward.* Dans un réseau *feed-forward*, les neurones sont uniquement connectés aux neurones de la couche supérieure avec pour conséquence que l'information se déplace uniquement depuis les neurones de la couche d'entrée vers ceux de la couche de sortie.

Chaque connexion entre deux neurones a un certain poids. Si ce poids est égal à 0, cela signifie qu'il n'y a pas de connexion entre ces deux neurones. Pendant la phase d'apprentissage, le réseau apprend en ajustant les poids de connexion pour minimiser une fonction de coût. Pour bien comprendre cette phase d'apprentissage, il est nécessaire

d'expliquer les notions de fonction d'activation, de fonction de coût et d'algorithme d'apprentissage.

Considérons un neurone artificiel avec $n$ entrées représentées par $x_i$ pour $i =$ (1,2, ... , $n$), $n$ poids $(w_1, w_2, \dots, w_n)$ associés aux $n$ entrées et un terme de biais $w_0$ pour $x_0 = 1$. La valeur effective de la sortie $y$ aura pour formule :

$$y = F\left(\sum_{i=1}^{n} w_i x_i + w_0\right) = F\left(\sum_{i=0}^{n} \mathbf{w}_i \mathbf{x}_i\right) = F[\mathbf{w}^T \mathbf{x}]$$

où $\mathbf{x} = [1, x_1, x_2, \dots, x_n]^T$ avec $x_0 = 1$

$\mathbf{w} = [w_0, w_1, w_2, \dots, w_n]^T$

et $F(\ )$ est la fonction d'activation.

Il existe différents types de fonctions d'activation. La fonction d'activation des neurones de la dernière couche d'un réseau sera choisie en fonction du type de problème auquel il faut répondre. Ci-dessous, sont regroupées les fonctions d'activation les plus courantes :

- Une fonction identité : $y = \mathbf{w}^T \mathbf{x}$

  La fonction d'activation est alors linéaire.

- Une fonction de « *Heaviside* » : $y = \begin{cases} 0 \ si \ \mathbf{w}^T \mathbf{x} < 0 \\ 1 \ si \ \mathbf{w}^T \mathbf{x} \geq 0 \end{cases}$

  avec une sortie binaire.

- Une fonction sigmoïde ou logistique :

  $$y = \frac{1}{1 + \exp[-(\mathbf{w}^T \mathbf{x})]}$$

  avec une sortie continue, $0 \leq y \leq 1$.

- Une fonction « *softmax* » :

  $$y = \frac{\exp(\mathbf{w}^T \mathbf{x})}{\sum_{j=1}^{q} \exp(w_j^T \mathbf{x})}$$

  avec des sorties positives, $0 \leq y \leq 1$ et dont la somme de ces sorties est 1.

- Une fonction tangente hyperbolique :

$$y = \frac{\exp(\boldsymbol{w}^T\boldsymbol{x}) - \exp[-(\boldsymbol{w}^T\boldsymbol{x})]}{\exp(\boldsymbol{w}^T\boldsymbol{x}) + \exp[-(\boldsymbol{w}^T\boldsymbol{x})]}$$

avec une sortie continue, $-1 \leq y \leq 1$.

### 4.2.2 L'algorithme de rétropropagation

Pour rappel, l'objectif de la phase d'apprentissage est d'obtenir les poids de connexion qui minimisent une fonction de coût prédéfinie. Dans ce mémoire, l'algorithme d'apprentissage utilisé à cette fin est l'algorithme de rétropropagation. L'algorithme de rétropropagation est itératif. Les poids de connexion sont ajustés lors de chaque itération jusqu'à ce que certaines conditions soient remplies. Une condition de fin peut, par exemple, être la fonction de coût qui atteint un minimum ou le nombre d'itérations qui dépasse un certain seuil.

L'algorithme de rétropropagation est utilisé conjointement avec une méthode d'optimisation qui cherche à résoudre un problème de minimisation de la fonction de coût par rapport aux poids de connexion [37]. Plusieurs méthodes existent, la plus connue étant l'algorithme du gradient (*Gradient Descent* en anglais), mais d'autres méthodes telles que la méthode du gradient conjugué ou encore l'algorithme **Broyden-Fletcher-Goldfarb-Shanno** (**BFGS**) sont souvent utilisées dans la pratique.

Soit l'itération $t$ d'un algorithme de rétropopagation associé à un algorithme BFGS, les différentes étapes de l'itération $t$ sont :

1. Activation des couches d'entrée (couche 0) avec le vecteur caractéristique $\mathbf{y}^{(0)}(k) = \mathbf{x}_{\boldsymbol{k}}$ pour l'observation $k$.

2. Les activations de toutes les unités des couches cachées et de la couche de sortie sont calculées en propageant l'activation de la couche d'entrée vers la couche de sortie :

- L'activation de l'unité $i$ du niveau $l$ (avec $p(l)$ unités pour la couche $l$) pour une observation $k$ est égale à :

$$\widehat{y}_i^{(l)}(\mathbf{x_k}) = \widehat{y}_i^{(l)}(k) = F\left[\sum_{j=1}^{p(l-1)} w_{ji}^{(l)} \widehat{y}_j^{(l-1)}(k) + w_{0i}^{(l)}\right]$$

- L'activation de l'unité $i$ de la couche de sortie $L$ (avec $q = p(L)$ unités pour la couche $L$ pour une observation $k$ est égale à :

$$\widehat{y}_i^{(L)}(\mathbf{x_k}) = \widehat{y}_i^{(L)}(k) = S\left[\sum_{j=1}^{p(L-1)} w_{ji}^{(L)} \widehat{y}_j^{(L-1)}(k) + w_{0i}^{(L)}\right]$$

où $w_{ji}^{(l)}$ est le poids de connexion entre l'unité $j$ de la couche $(l-1)$ et l'unité $i$ de la couche $l$ et où $w_{0i}^{(l)}$ est le terme de biais pour l'unité $i$ de la couche $l$.

3. Une fois les valeurs de sortie des neurones de la couche $L$ connues, l'algorithme calcule la fonction de coût $C$ appelée aussi fonction d'erreur. Il existe une large gamme de fonctions d'erreurs comme, par exemple, l'erreur quadratique ou encore l'entropie croisée parmi les plus connues. Selon le type de problème et la fonction d'activation des neurones de la couche de sortie, une fonction de coût sera privilégiée par rapport aux autres.

4. L'algorithme détermine ensuite la contribution relative à l'erreur $C$ de chaque neurone du réseau. On représente « l'erreur » du neurone $i$ pour la couche $l$ par $\delta_i^{(l)}(k)$. Les erreurs $\delta_i^{(l)}(k)$ sont calculées par rétropropagation depuis la couche de sortie vers la couche d'entrée.

   L'erreur du neurone $i$ pour la couche de sortie $L$ a pour équation :

$$\delta_i^{(L)}(k) = \frac{\partial C}{\partial \widehat{y}_i^{(L)}(k)} {F'}_i^{(L)}(k)$$

avec ${F'}_i^{(L)}(k) = \frac{\partial S_i^{(L)}\left(In_i^{(L)}(k)\right)}{\partial In_i^{(L)}(k)}$, la dérivée de la fonction d'activation du neurone $i$ de la couche de sortie par rapport à la somme pondérée de ses entrées $In_i^{(L)}(k)$.

L'erreur du neurone $i$ pour la couche de sortie $l$ a pour équation :

$$\delta_i^{(l)}(k) = {F'}_i^{(l)}(k) \sum_{j=1}^{p(l+1)} w_{ij}^{(l+1)} \delta_j^{(l+1)}(k)$$

avec ${F'}_i^{(l)}(k) = \frac{\partial F_i^{(l)}\left(In_i^{(l)}(k)\right)}{\partial I \ {}_i^{(l)}(k)}$

5. Les contributions relatives à l'erreur et les activations de chaque neurone étant connues, l'algorithme dispose de toutes les informations pour trouver le gradient nécessaire au processus de mise à jour des poids de connexion. Le gradient de $w_{ji}^{(l)}$ noté $\frac{\partial C}{\partial w_{ji}^{(l)}}$ sera égal à la somme des gradients de toutes les observations $k$.

$$\frac{\partial C}{\partial w_{ji}^{(l)}} = \sum_{k=1}^{n} \delta_i^{(l)}(k)\, \hat{y}_{ij}^{(l-1)}(k)$$

6. Soit $w_t$ le vecteur des poids de connexion et $\nabla f(w_t)$ le vecteur gradient de l'itération $t$. La matrice $\mathrm{H}_t$, une approximation du Hessien pour l'itération $t$ est déterminée à l'aide de la méthode BFGS détaillée dans l'annexe **a**. Elle est calculée avec l'égalité suivante :

$$\mathrm{H}_t = \mathrm{H}_{t-1} + \frac{yy^T}{y^T s} - \frac{\mathrm{H}_{t-1} s s^T \mathrm{H}_{t-1}}{s^T \mathrm{H}_{t-1} s}$$

avec

$s = w_t - w_{t-1}$

$y = \nabla f(w_t) - \nabla f(w_{t-1})$

7. Calcul de $\alpha_t$ la longueur de pas de l'itération $t$ en utilisant un algorithme de recherche linéaire

8. Mise à jour du vecteur des poids de connexion. La formule de mise à jour est :

$$w_{t+1} = w_t + \alpha_t d_t$$

avec $d_t = -\mathrm{H}_t \nabla f(x_t)$, la direction de recherche en $t$.

Pour la première itération $t = 0$, les poids de connexions sont généralement choisis aléatoirement et l'approximation du Hessien $\mathrm{H}_0$ est égale à la matrice identité.

## 4.3 Comparaison des deux méthodes d'apprentissage

D'un point de vue technique, les perceptrons multicouches et les régressions logistiques se différencient sur plusieurs points mais ont également certaines caractéristiques communes. D'ailleurs, une régression logistique n'est autre qu'un cas particulier de perceptron multicouche sans couche cachée et dont la fonction d'activation est une fonction sigmoïde.

Grâce aux couches cachées, les perceptrons sont capables de modéliser des relations entre les variables explicatives et dépendantes plus complexes que les régressions logistiques. En effet, avec les régressions logistiques, la relation entre les variables indépendantes et la variable cible est déjà définie [7].

Du fait de leurs complexité, les perceptrons sont plus sensibles au surapprentissage[3]. Il existe plusieurs techniques de régularisation pour gérer l'*overfitting*:

- Modifier la taille de l'échantillon d'apprentissage.
- Choisir le nombre de couches cachées que contient un réseau de neurones artificiel. Une couche est généralement suffisante pour la grande majorité des problèmes [2].
- Varier le nombre de neurones dans les couches cachées.

[3] Le surapprentissage ou l'*overfitting* survient lorsqu'un modèle apprend non seulement les relations entre variables mais en plus le bruit que contiennent les données [2]. Si c'est le cas, le modèle aura de mauvaises performances prédictives avec de nouvelles données.

- Ajuster le nombre d'itérations des algorithmes.
- Introduire une modération de poids (*weight decay* en anglais) pour les perceptrons multicouches. L'idée, derrière la modération de poids, est de donner à chaque poids de connexion une tendance à décliner vers 0 afin que la connexion entre deux neurones disparaisse sauf si elle contribue à réduire l'erreur du réseau [2].

De manière générale, les paramètres ci-dessus seront ajustés à l'aide d'une étape de validation durant laquelle les performances de plusieurs modèles avec différentes combinaisons de paramètres seront comparées sur base d'un nouvel échantillon de données.

Etant donné que la régression logistique est un problème d'optimisation convexe, elle a l'avantage de n'avoir qu'un minimum pour sa fonction d'erreur. Des méthodes d'optimisation locale comme l'algorithme du gradient ou encore l'algorithme BFGS peuvent même être utilisées pour trouver ce minimum global. Au contraire, la fonction d'erreur d'un réseau de neurones artificiels peut avoir plusieurs minima locaux. C'est pourquoi il est généralement recommandé d'ajuster plusieurs perceptrons et de calculer la performance moyenne [19].

Enfin, la complexité des modèles, les techniques de régularisation pour éviter l'*overfitting* et l'ajustement de plusieurs modèles pour répondre aux problèmes des minima locaux font que le temps d'apprentissage des perceptrons multicouches est plus lent que l'apprentissage des régressions logistiques. Cette réalité devra être prise en compte lors du développement des méthodes de prédiction.

# 5 Elaboration des series temporelles

Ce chapitre explique les étapes qui permettent de passer des données brutes extraites de Bloomberg et Dukascopy aux séries de données adaptées à la prévision financière, en respectant les éventuelles contraintes et en facilitant la comparaison des performances des différents modèles.

## 5.1 Collecte des données

### 5.1.1 Les taux de change

Les observations de l'euro-dollar et du dollar-yen ont été extraites du site www.dukascopy.com, qui permet de télécharger depuis 2007 les cotations de paires de devises dans une fréquence supérieure ou égale à 1 minute. Dans ce mémoire, c'est la fréquence de 5 minutes qui sera sélectionnée. Le fichier téléchargé indique la date et l'heure, le prix le plus haut/bas du *bid* et de l'*ask*, le prix d'ouverture/de fermeture du *bid* et de l'*ask* et le volume échangé au *bid* et à l'*ask* pour chaque période d'observation.

Toutes les informations ne sont pas utiles pour les séries temporelles finales. Seuls la date et l'heure, le prix d'ouverture/de fermeture du *bid* et de l'*ask* seront conservés, tandis que les observations où le volume est nul sont éliminées de la série. Après analyse, les observations supprimées correspondent à tous les week-ends ou jours fériés, pendant lesquels le marché des changes n'est pas ouvert.

*Figure 5 - Echantillon d'observations de l'euro-dollar*

| Time | | EUR/USD ASK | Open | High | Low | Close | Volume | EUR/USD Bid | Open | High | Low | Close | Volume |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 01-03-07 | 00:00:00 | | 1,32271 | 1,32335 | 1,32267 | 1,3231 | 3974,3 | | 1,32255 | 1,32325 | 1,32249 | 1,323 | 4580,6 |
| 01-03-07 | 00:05:00 | | 1,32304 | 1,32331 | 1,32264 | 1,32304 | 3158,6 | | 1,32288 | 1,32316 | 1,32255 | 1,32294 | 4028,6 |
| 01-03-07 | 00:10:00 | | 1,3228 | 1,32307 | 1,32255 | 1,32271 | 2833 | | 1,32275 | 1,32295 | 1,3224 | 1,32256 | 4413 |
| 01-03-07 | 00:15:00 | | 1,32276 | 1,32295 | 1,32228 | 1,32264 | 3481,6 | | 1,3226 | 1,32282 | 1,32211 | 1,32244 | 4772,3 |
| 01-03-07 | 00:20:00 | | 1,32289 | 1,32314 | 1,32261 | 1,32281 | 3233 | | 1,32285 | 1,32294 | 1,32255 | 1,32271 | 3959,4 |
| 01-03-07 | 00:25:00 | | 1,32267 | 1,32286 | 1,32228 | 1,32232 | 2847,6 | | 1,32257 | 1,32275 | 1,32212 | 1,32222 | 3362,4 |
| 01-03-07 | 00:30:00 | | 1,3224 | 1,3226 | 1,32217 | 1,32241 | 3605,6 | | 1,3223 | 1,32244 | 1,32207 | 1,32231 | 3273,8 |
| 01-03-07 | 00:35:00 | | 1,32235 | 1,32298 | 1,3223 | 1,32293 | 3242,7 | | 1,32225 | 1,32285 | 1,32217 | 1,32283 | 4442,3 |
| 01-03-07 | 00:40:00 | | 1,32297 | 1,32297 | 1,32254 | 1,32264 | 3564,5 | | 1,32281 | 1,32281 | 1,32241 | 1,32254 | 4048,3 |
| 01-03-07 | 00:45:00 | | 1,32272 | 1,32315 | 1,32268 | 1,32285 | 3239,3 | | 1,32262 | 1,32301 | 1,32254 | 1,32269 | 4062,5 |
| 01-03-07 | 00:50:00 | | 1,32285 | 1,32301 | 1,32239 | 1,32245 | 3352 | | 1,32275 | 1,32286 | 1,32222 | 1,32241 | 3602,6 |
| 01-03-07 | 00:55:00 | | 1,32237 | 1,32354 | 1,32237 | 1,32346 | 2895,4 | | 1,32221 | 1,32339 | 1,32221 | 1,3233 | 3734,2 |
| 01-03-07 | 01:00:00 | | 1,32353 | 1,32353 | 1,3227 | 1,32273 | 3623,4 | | 1,32343 | 1,32343 | 1,32255 | 1,32258 | 4786,4 |
| 01-03-07 | 01:05:00 | | 1,32292 | 1,32312 | 1,32279 | 1,32297 | 3386,9 | | 1,32282 | 1,32297 | 1,32261 | 1,32287 | 4275,8 |

Une fois modifié, le fichier contient 639.700 observations entre le 01/03/2007 à 00:00 et le 10/09/2015 à 16:45.

### 5.1.2 Les données macroéconomiques

A partir de la plateforme Bloomberg et pour les onze indicateurs macroéconomiques sélectionnés au point 2.3.2, il est possible d'extraire un tableau contenant une série de données pour chaque publication macroéconomique depuis 1996. Les seules informations retenues sont la date et l'heure des publications, la moyenne des prévisions d'un ensemble d'analystes et la valeur réelle de la statistique économique obtenue à l'instant précis de la publication.

Lorsque des modèles dont l'objectif est de prédire des séries temporelles financières intègrent des données fondamentales comme variables, plusieurs aspects doivent être pris en compte [23]. Il ne faut prendre que des données non-révisées. Le chercheur doit s'assurer que les informations soient encore disponibles à l'avenir. Le dernier point et sans doute le plus important, est de s'assurer que l'heure de publication soit cohérente avec la série temporelle étudiée. Si, par exemple, il y a une publication du PIB américain le 30 octobre 2008 à 12 heures 30 UTC, cette information doit être reflétée ni avant ni après, mais exactement au même moment dans le modèle.

## 5.2 Les séries temporelles

Trois séries temporelles forment l'ensemble des données nécessaires à la construction, l'ajustement et la prédiction des modèles.

La série temporelle « **log-rendement du taux de change** » est la série principale et indispensable. Les modèles autorégressifs seront calculés sur base de ses valeurs et les algorithmes d'apprentissage automatique utiliseront ses valeurs passées comme *inputs*.

Dans un second temps, la série nommée « **effet de surprise des publications** » sera intégrée aux modèles sous la forme d'une variable exogène dans le processus autorégressif et comme *input* supplémentaire pour l'apprentissage automatique. Le but de cette intégration sera de voir si l'ajout d'informations concernant les publications macroéconomiques permet d'augmenter la prédictibilité des taux de change post-annonce macroéconomique.

La dernière série « **direction du taux de change** » servira à déterminer la qualité des prédictions. Dans le cas des algorithmes d'apprentissage, la série sera utilisée comme *output*. Tandis que pour mes processus autorégressifs, elle sera comparée à la direction des estimations des taux de change.

Il est à noter que toutes les transformations et explications qui suivent sont valables aussi bien pour l'euro dollar que le dollar yen et pourraient être étendues à d'autres paires de devises et d'autres unités de temps.

## 5.3 Construction de la série « log-rendement du taux de change »

Avec $OPENBID_t$, $OPENASK_t$, $CLOSEBID_t$, $CLOSEASK_t$, respectivement le prix du *bid* et de l'*ask* à l'ouverture de la période $t$, ainsi que le prix du *bid* et de l'*ask* à la fermeture de la période $t$, la valeur de la série « log-rendement du taux de change » en $t$ est :

$$\boldsymbol{LR_t} = \ln\left(\frac{CLOSEBID_t + CLOSEASK_t}{OPENBID_t + OPENASK_t}\right)$$

Les avantages du log-rendement pour la prévision financière sont multiples, notamment par rapport à l'utilisation du prix d'un actif comme les taux de change. Premièrement, le rendement d'un actif résume les opportunités d'investissement de manière complète,

indépendamment de la grandeur de l'actif sous-jacent. Ensuite, les propriétés statistiques des séries de rendements les rendent plus faciles à manipuler que les séries de prix [33].

Beaucoup de méthodes économétriques et statistiques sont adaptées pour des séries stationnaires, c'est-à-dire des séries dont la moyenne et la variance sont stables dans le temps. Une rapide analyse des graphiques et fonctions d'autocorrélation de sous-échantillons de la série des taux de change entre l'euro et le dollar ci-dessous et en annexe **b** confirme le caractère non stationnaire des taux de change. En effet, si on comparait la moyenne de différents échantillons dans le temps, elle ne serait pas stable. De plus, la fonction d'autocorrélation diminue lentement, comportement typique des fonctions d'autocorrélation de séries non stationnaires [5]. Cette analyse est renforcée par les résultats du test de Dickey-Fuller réalisé sur les mêmes sous-échantillons : les p-valeurs sont de 0,801, 0,1821 et 0,4107 (voir annexe **c**). L'hypothèse nulle ne peut pas être rejetée. Il est donc vraisemblable que la série des taux de change soit non stationnaire en différence.

*Figure 6 - Evolution d'un sous-échantillon de la série euro-dollar*

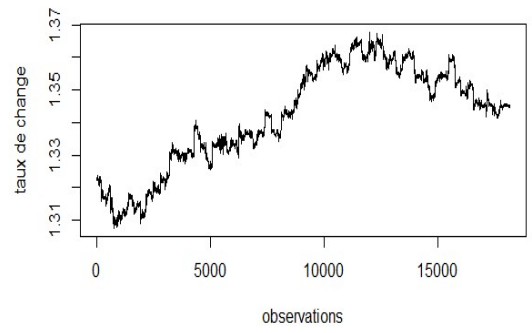



*Figure 7- ACF d'un sous-échantillon de la série euro-dollar*

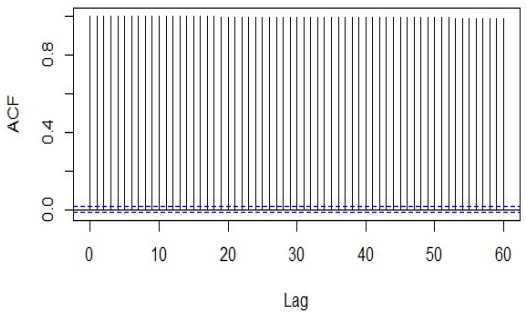


*Figure 8 - Evolution d'un sous-échantillon de la série log-rendement*

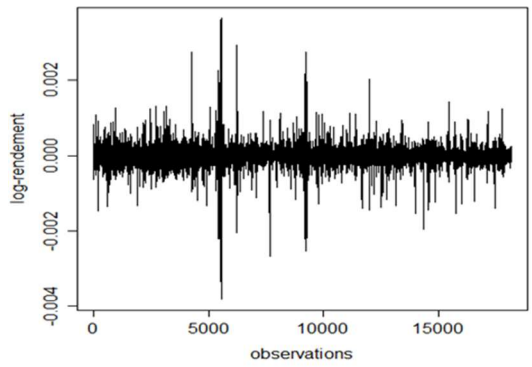
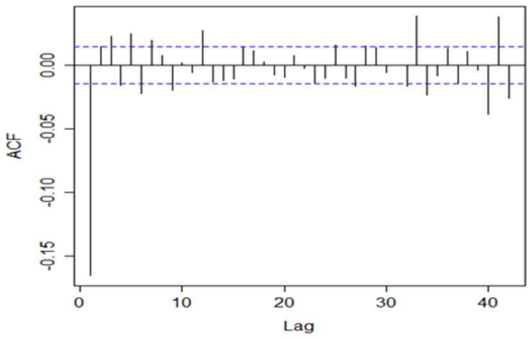



*Figure 9 - ACF d'un sous-échantillon de la série log-rendement*




Les caractéristiques des sous-échantillons de la série log-rendements ci-dessus et en annexe **d** se rapprochent déjà beaucoup plus de celles d'une série stationnaire. La fonction d'autocorrélation de la série log-rendement chute rapidement vers 0, les log-rendements évoluent autour de 0 et la variance semble assez stable malgré de potentielles « grappes » de volatilité *(*en anglais *volatility clustering)*. Contrairement à la série « brute » des taux de change, les p-valeurs des tests de Dickey-Fuller augmentés, disponibles en annexe **e**, sont inférieures à 0,01 pour les 3 sous-échantillons. L'hypothèse nulle est donc rejetée, ce qui indique que la série log-rendement ne nécessite plus de différenciation.

Un autre élément dans la construction de la série log-rendement qui mérite justification est le choix de prendre le prix d'ouverture et de fermeture pour calculer le log-rendement plutôt que, comme beaucoup d'études, utiliser les prix de fermeture (le rendement en $t$ est alors calculé grâce aux prix de fermeture en $t$ et $t$-1). Un modèle qui veut prévoir le rendement en $t$ doit être cohérent avec le rendement dont bénéficierait un acteur financier qui suivrait le signal donné. La prédiction du rendement en $t$ ne pourra se faire qu'une fois le prix de fermeture en $t$-1 connu. Si le rendement était basé sur les

prix de fermeture, cela voudrait dire que l'acteur financier achèterait ou vendrait instantanément à la fermeture de $t$-1 selon le signal donné, scénario qui n'est pas réaliste. Par contre, en utilisant le prix d'ouverture, il y a une marge entre le moment où la prévision est calculée et le moment où l'ordre est passé sur le marché.

### 5.3.1 Construction de la série « effet de surprise des publications »

Dans la partie 2.3.2, onze indicateurs macroéconomiques ont été mis en évidence pour leur impact significatif sur les taux de change. C'est à partir de leurs données pour chaque publication entre le 01/03/2007 et le 10/09/2015 que la série temporelle « effet de surprise des publications » va être développée.

Soit $A_{kt} - E_{kt}$, la « nouvelle » de l'annonce pour un indicateur économique k au temps t où $A_{kt}$ est la valeur annoncée de l'indicateur k en t et $E_{kt}$ est la valeur attendue par le marché de l'indicateur k en t. La première étape est de calculer les « nouvelles » pour toutes les publications des onze indicateurs.

Pour les dix indicateurs avec un coefficient positif face au dollar[4], si la « nouvelle » de l'annonce $A_{kt} - E_{kt}$ est positive, alors la « surprise » de l'annonce liée à l'indicateur k en t $s_{kt}$ est égale à 1 tandis que, si la « nouvelle » de l'annonce est négative, $s_{kt}$ est égale à -1 et, si la « nouvelle » est nulle, $s_{kt}$ est également nulle.

Pour le seul indicateur significatif qui a un coefficient négatif avec le dollar[5], $s_{kt}$ aura la valeur de 1 si la « nouvelle » de l'annonce est négative et $s_{kt}$ aura la valeur de -1 si la « nouvelle » de l'annonce est positive.

---

[4] Le PIB avancé, le taux d'emploi non agricole, les ventes au détail, la production industrielle, les commandes de biens durables, l'indice manufacturier de l'*Institute of Supply Management*, les dépenses de construction, la balance commerciale, l'indice de confiance des consommateurs et le taux objectif des fonds fédéraux.

[5] Le nombre de demandes initiales d'assurance chômage

Une fois toutes ces données collectées, il est facile de calculer $\boldsymbol{P_t}$ pour chacune des 639.700 observations de la série finale « effet de surprise des publications » :

$$\boldsymbol{P_t} = \begin{cases} 1 \ si \ \sum_{k=1}^{11} s_{kt} > 0 \\ -1 \ si \ \sum_{k=1}^{11} s_{kt} < 0 \\ 0 \ si \ \sum_{k=1}^{11} s_{kt} = 0 \end{cases}$$

La grande majorité des observations auront $\boldsymbol{P_t = 0}$, puisque les observations où une publication macroéconomique a lieu sont assez rares (1.156 pour 639.700 observations toutes les 5 minutes entre le 01/03/2007 et le 10/09/2015). Parmi ces 1.156 observations, seules trois observations enregistrent trois publications simultanées de différents indicateurs macroéconomiques, 202 observations ont deux publications simultanées et le reste ne compte qu'une publication à la fois. 470 observations se retrouvent avec $\boldsymbol{P_t} = -1$, 492 avec $\boldsymbol{P_t} = 1$ et 194 avec $\boldsymbol{P_t} = 0$. Il y a donc une assez bonne répartition entre les $\boldsymbol{P_t} = -1$ et les $\boldsymbol{P_t} = 1$.

La structure de la série finale est le résultat d'un compromis entre l'information qu'apportent les valeurs de la série et la bonne intégration de la série aux modèles de prévision et de prédiction. En exprimant « l'effet de surprise » d'une publication par 3 valeurs possibles (1,-1 et 0), il y a une perte d'information, puisque la grandeur de l'effet n'est plus prise en compte. Mais cette perte est nécessaire pour pouvoir exprimer, dans une même variable, « l'effet de surprise » de différents indicateurs dont les données ne sont ni de même taille ni de même distribution. De plus, faire onze variables pour onze indicateurs aurait été mathématiquement beaucoup plus lourd et moins cohérent vu les différences de fréquence de publications entre certains indicateurs.

### 5.3.2 Construction de la série « direction du taux de change »

Une fois le log-rendement du taux de change connu, le calcul des valeurs de cette nouvelle série est direct : Si $\boldsymbol{D_t}$ est la valeur de la série en t et $\boldsymbol{LR_t}$ le log-rendement du taux de change, alors :

$$\boldsymbol{D_t} = \begin{cases} 1 \; si \; \boldsymbol{LR_t} > 0 \\ 0 \; si \; \boldsymbol{LR_t} \leq 0 \end{cases}$$

Comme son nom l'indique, la série précise la direction du taux de change pour les 639.700 observations disponibles : lorsque le taux de change est positif, la direction est haussière et la valeur de $\boldsymbol{D_t}$ est 1, tandis que, si le taux de change est nul ou négatif, la valeur de $\boldsymbol{D_t}$ est 0.

Contrairement aux deux séries précédentes, la série « direction du taux de change » n'est pas utilisé comme *input*. Les valeurs de cette série seront comparées aux prédictions des modèles afin de déterminer leurs capacités à prévoir la future direction des taux de change. De plus amples détails sur le calcul des performances seront donnés dans le chapitre suivant.

# 6 Conception des modèles

Après avoir sélectionné les données adaptées à la prédiction financière, ce nouveau chapitre expliquera comment les différents modèles sont construits et comment calculer les prédictions *out-of-sample* à partir des modèles appris.

Dans un premier temps, le concept de fenêtre glissante sera introduit. Ensuite, les processus de construction des modèles prévisionnels et prédictifs seront décrits un par un. On commencera par le développement du modèle « moyenne » qui servira de modèle *benchmark* lors de l'analyse des résultats. En second lieu, on analysera la méthodologie qui permet de sélectionner les meilleurs modèles ARIMA et d'en calculer les prévisions. Le dernier point détaillera comment générer des prédictions à partir des modèles de régression logistique et de réseau de neurones artificiels. Le tout sera implémenté à l'aide du logiciel R[6].

## 6.1 Concept de fenêtre glissante

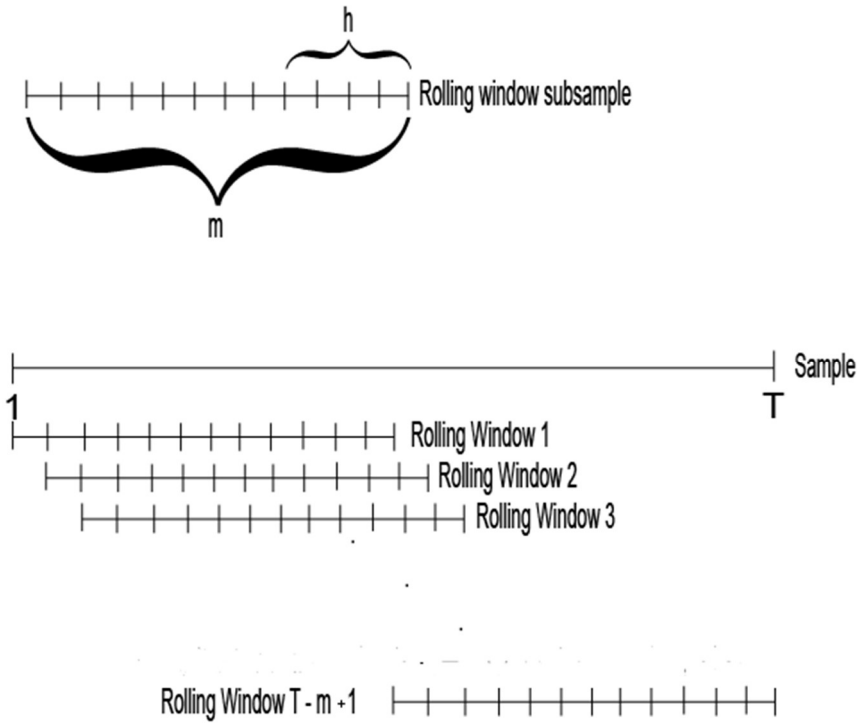
En analyse des séries temporelles, la prévision par fenêtre glissante est un processus qui utilise une fenêtre avec un nombre fixe d'observations se déplaçant de manière chronologique sur toute la série temporelle. A chaque fois que la fenêtre est déplacée, les différents paramètres du modèle pour lequel on applique une fenêtre glissante sont réajustés sur base des nouvelles observations comprises dans la nouvelle fenêtre et une ou plusieurs nouvelles prévisions *out-of-sample* peuvent être considérées.

Pour mettre en place une prévision par fenêtre glissante [24], il faut d'abord choisir la taille de la fenêtre glissante $m$ (c'est-à-dire le nombre d'observations consécutives dans une fenêtre glissante). Cette taille va dépendre des contraintes du problème et de la

[6] R est un logiciel libre de traitement de données et d'analyses statistiques mettant en œuvre le langage de programmation S [36]. Plus précisément, c'est RStudio qui sera employé. RStudio qui est un environnement de développement lié à R.

taille de l'entièreté de l'échantillon disponible. Ensuite, un horizon de prévision $h$ est déterminé, horizon qui représente le nombre de prévisions *out-of-sample* calculées à chaque nouvelle fenêtre. Le déplacement de la fenêtre glissante se fait par incrément. Cet incrément correspond à un nombre fixe d'observations. Un incrément composé d'une observation signifie que, pour chaque nouvelle fenêtre, l'observation la plus ancienne de la fenêtre précédente est enlevée du nouvel échantillon et l'observation la plus ancienne qui n'était pas encore comprise dans l'échantillon y est ajoutée.

*Figure 10 - Fonctionnement des fenêtres glissantes*




*source: http://nl.mathworks.com/help/econ/rolling-window-estimation-of-state-space-models.html*

La construction d'un modèle statistique, appliqué à une série temporelle pour en estimer les valeurs futures, implique l'hypothèse fondamentale selon laquelle les paramètres de ce modèle restent constants sur toute la période étudiée [34]. Dans ce mémoire, les observations des séries temporelles financières étudiées s'étalent sur plusieurs années. Utiliser un modèle de prévision qui engloberait la totalité des observations irait selon

toute vraisemblance à l'encontre de l'hypothèse fondamentale énoncée ci-dessus. En effet, il n'est pas cohérent de supposer que les paramètres du modèle sont constants sur le long terme, alors que le contexte économique et l'interprétation de ce contexte par les acteurs financiers changent régulièrement.

Plus particulièrement, les relations entre les variables du modèle (rendements à prévoir, rendements antérieurs et « effet de surprise » des annonces macroéconomiques) auront certainement évolué au cours de cette période. Par exemple, l'interprétation des publications des indicateurs macroéconomiques par les acteurs financiers pourra varier en fonction du contexte. En général, des annonces économiques avec de bons résultats feront apprécier la valeur de la monnaie concernée. Mais, dans certains contextes de marché, les nouvelles macroéconomiques pourront avoir un effet inverse sur la valeur de la monnaie en question: Si une monnaie est considérée comme une valeur refuge telle le dollar américain, et que le contexte des marchés met en avant la crainte d'un ralentissement de l'économie mondiale, de mauvais résultats concernant l'économie américaine attiseront la peur d'un ralentissement économique globalisé et inciteront les investisseurs à acheter des dollars américains pour leur qualité de valeur refuge [31].

Le mécanisme de fenêtre glissante devrait donc permettre au modèle de faire varier ses paramètres pour être plus proche de la réalité des marchés tout en respectant mieux l'hypothèse fondamentale selon laquelle les paramètres sont constants puisque le modèle se base sur un échantillon d'observations plus court pour ajuster ses paramètres.

### 6.1.1 Spécification de la fenêtre

Comme expliqué dans la partie 6.1., pour pouvoir appliquer un mécanisme de fenêtre glissante à un échantillon, il faut déterminer la taille $m$, l'horizon $h$ et l'incrément des fenêtres glissantes ainsi que la taille $T$ de l'échantillon complet. $T$ est déjà connu tant pour l'euro dollar que le dollar yen et il est composé de 639.700 observations. Il ne reste

donc plus qu'à choisir et justifier les valeurs de $h$ et $m$ ainsi que l'incrément. Le processus permettant de les déterminer va dépendre d'un ensemble de contraintes, parfois contradictoires, qui devront être respectées le mieux possible.

La première contrainte est d'ordre technique et concerne l'incrément ; plus celui-ci est petit, plus il y aura de fenêtres ce qui implique davantage de réajustement des modèles et donc un plus long temps de calcul.

La périodicité des données a aussi un impact sur le choix de $h$ et $m$. La taille et l'horizon de la fenêtre glissante doivent être suffisamment grands afin de représenter un ensemble cohérent de variations des marchés tout en étant limités pour éviter au maximum de capturer des observations dans des contextes économiques différents, comme expliqué précédemment.

Enfin, les valeurs de $h$ et $m$ dépendent également de l'objectif poursuivi au travers des modèles de prévision. Un des points clés du mémoire est l'intégration de données relatives aux publications macroéconomique pour améliorer les performances de prévision. Les publications macroéconomiques étant assez espacées dans le temps, la taille de la fenêtre devra comprendre suffisamment de données sur ces publications pour permettre aux modèles de prévision d'estimer au mieux le paramètre lié à la variable « effet de surprise » des annonces macroéconomiques.

Une fois toutes ces contraintes prises en compte, la taille de la fenêtre glissante $m$ a été fixée à 21.775 observations, l'horizon $h$ à 3.625 et l'incrément également à 3.625 observations ce qui donne un total de 171 fenêtres pour l'échantillon complet de 639.700 observations. Ce choix est le fruit d'un compromis entre le temps de calcul et la problé-

matique de changement du contexte économique. La différence entre la taille et l'horizon est égale à 18.150 observations ce qui correspond plus ou moins à une période de trois mois sur le marché des changes, période la plus petite comprenant au moins une fois une publication de chaque indicateur économique utilisé pour créer la variable « effet de surprise » des publications[7]. L'horizon contient délibérément le même nombre d'observations que l'incrément et ce, afin d'avoir, pour chaque modèle, une série continue d'une et une seule prévision *out-of-sample* sur l'ensemble des observations, excepté les 18.150 premières observations de la première fenêtre glissante.

Les modèles prévisionnels se basent sur des ensembles de données appelés *training set*, *test set* et *validation set*[8]. A chaque nouvelle fenêtre glissante, les modèles se voient attribués de nouveaux *training sets*, *test sets* et *validation sets* dont la grandeur dépend de la taille et de l'horizon de la fenêtre glissante. Le *test set* est équivalent à l'horizon, il sera toujours composé des 3 .625 dernières observations de la fenêtre. Pour les méthodes régressives, le *training set* comprend toutes les observations présentes dans la fenêtre hormis celles déjà comprises dans le *test set*, c'est-à-dire 18.150 observations. Dans le cas des réseaux de neurones artificiels, il y aura un *validation test* correspondant à la suite d'observations précédant le *test set* et de même grandeur que celui-ci. Le *training set* ne contiendra donc plus que de 14.525 observations.

## 6.2 Développement du modèle « moyenne »

Comme expliqué au point 3.2, le modèle « moyenne » est lié à la théorie de la marche aléatoire ; Les données étudiées ne sont pas les logarithmes des taux de change mais les

[7] Parmi les indicateurs macroéconomiques étudiés, c'est le PIB américain qui a la plus faible fréquence de publication avec une publication trimestrielle.

[8] Le *training set* est le jeu de données utilisé pour ajuster les paramètres d'un modèle. Le *validation set* est utilisé pour ajuster les éventuels hyperparamètres. Le *test set* permet d'évaluer les performances du modèle.

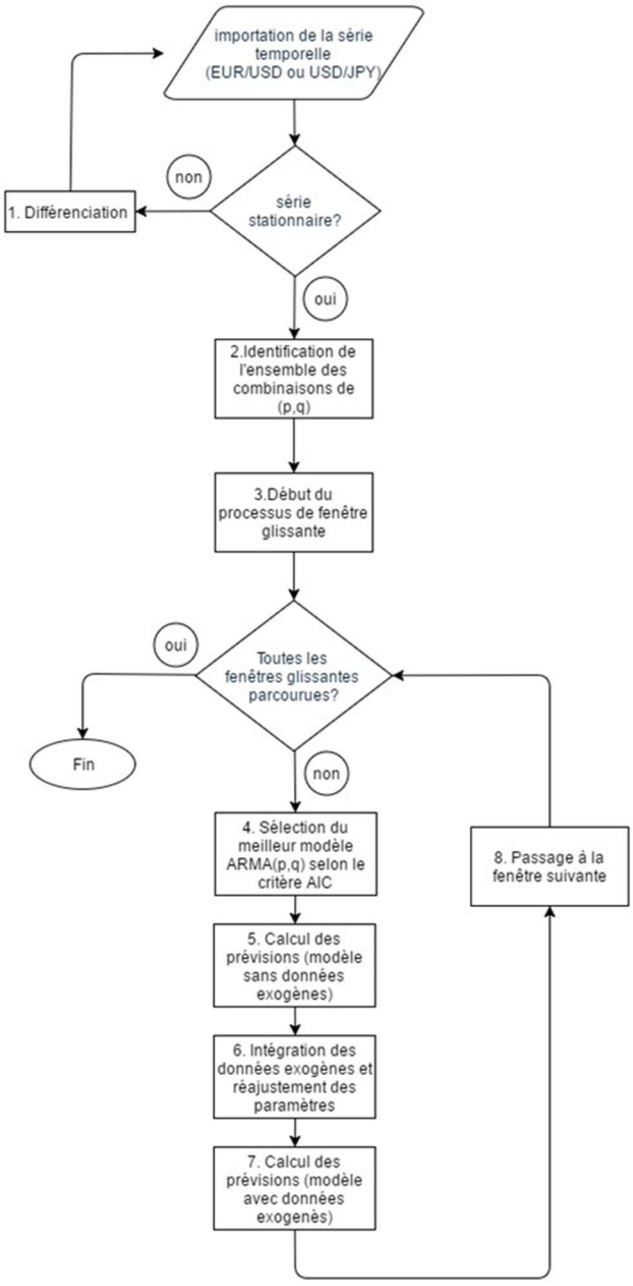
logarithmes des variations entre taux de change, la méthode à privilégier pour représenter la marche aléatoire est donc la méthode « moyenne » qui représente le mieux la différence première d'une série temporelle suivant une marche aléatoire.

Avec le modèle « moyenne », les prévisions *out-of-sample* d'une série temporelle sont égales à la valeur moyenne des observations du *training set.* Pratiquement, la série temporelle des log-rendements de l'euro-dollar ou du dollar-yen sera transformée en 171 « sous-séries » grâce au mécanisme de fenêtre glissante expliqué dans la partie 6.2. A chaque fenêtre, une nouvelle moyenne sera calculée sur base du *training set* et attribuée comme prévisions *out-of-sample* des 3.625 dernières observations de la fenêtre. Les performances pourront être calculées de manière ciblée en comparant uniquement les prévisions *out-of-sample* d'une fenêtre spécifique au *test set* correspondant ou de manière « globale » en établissant les performances sur base des 171 échantillons de prévisions.

## 6.3 Développement des modèles ARIMA (p,d,q)

Cette nouvelle partie décrit la modélisation ARIMA(p,d,q) des séries de taux de change adaptée au système de fenêtres glissantes et capable , dans un second temps, d'intégrer la variable exogène « effet de surprise » des publications macroéconomiques. Le diagramme de flux ci-dessous synthétise les différentes étapes.

*Figure 11 - Processus de prédiction de la direction des taux de change sur base des modèles ARIMA*



La première étape de la modélisation consiste à différencier la série temporelle jusqu'à ce qu'elle soit stationnaire. Lors de la construction de la série « log-rendement du taux de change » au point 5.2.1, le caractère stationnaire des séries log-rendements pour l'euro-dollar et le dollar-yen avait déjà été démontré par une analyse graphique renforcée par les résultats des tests de Dickey-Fuller augmenté. Il n'est donc plus pertinent de parler de modélisation ARIMA (p,d,q) mais plutôt de modélisation ARMA(p,q) étant donné qu'aucune différenciation ne sera appliquée aux séries temporelles.

Pour chaque fenêtre glissante, le meilleur modèle ARMA (p,q) sera sélectionné parmi les modèles « candidats » sur base des critères de sélection préétablis. A chaque « candidat » correspond une combinaison (p,q) spécifique. Le point 2 du diagramme représente l'étape lors de laquelle l'ensemble de ces combinaisons est déterminé en fixant les valeurs minimales et maximales que pourront prendre les ordres $p$ et $q$.

La grandeur maximale imposée aux ordres $p$ et $q$ est de 5. Le meilleur modèle sélectionné par fenêtre glissante ne pourra donc pas avoir un $p$ ou un $q$ plus grand qu'ARMA (5,5). Ce choix s'explique notamment par le contexte de la recherche et les données utilisées : Au-delà d'un certain point, il apparaît peu probable que les rendements des taux de change passés aient un impact significatif sur le rendement à prévoir. Ce choix a été aussi pensé en fonction des contraintes techniques. Les valeurs maximales fixées pour $p$ et $q$ sont les plus grandes valeurs permettant de garder un temps de calcul raisonnable.

La spécification des valeurs minimales de $p$ et $q$ dépend indirectement de la manière dont sera calculé le réseau de neurones artificiels. En effet, le réseau de neurones artificiels ne fonctionne pas sous R sans un minimum de 2 inputs. Or, le réseau de neurones aura comme inputs, à chaque fenêtre glissante, les $p$ valeurs passées avec $p$ égal à

l'ordre autorégressif du modèle ARMA (p,q) sélectionné. C'est pourquoi un premier algorithme sélectionnera le meilleur modèle parmi tous les candidats modèles ARMA (p,q) avec $2 \leq p \leq 5$ et $0 \leq q \leq 5$. Un second algorithme choisira le meilleur modèle parmi tous les candidats modèles ARMA (p,q) avec $0 \leq p \leq 5$ et $0 \leq q \leq 5$ et leurs résultats pourront être comparés. Il sera notamment intéressant de voir s'il y a une différence significative entre les performances des deux processus.

Le processus de fenêtre glissante représenté au point 3 du diagramme de flux ayant déjà été détaillé dans la partie 6.1, on peut donc aborder directement le point 4.

Le meilleur modèle ARMA (p,q) est sélectionné automatiquement lors de chaque fenêtre glissante à l'aide d'un critère de sélection. Le critère choisi est le critère d'information d'Akaike (AIC) défini dans l'annexe **f**.

Une fois le meilleur modèle ARMA (p,q) connu, il ne reste plus qu'à calculer les prévisions *out-of-sample* de chaque fenêtre, comme le précisent les points 5 à 7 du diagramme de flux. Les premières prévisions calculées sont celles du modèle ARMA (p,q) sans variables exogènes. Pour la deuxième série de prévisions, le modèle ARMA (p,q) intègre la série « effet de surprise » comme variable exogène. Le modèle garde les mêmes ordres $p$ et $q$ mais les coefficients du modèle sont réajustés avant de calculer les prévisions.

En fonction des données autorégressives[9], de l'intégration de la variable exogène[10] et des limites imposées aux ordres $p$ et $q$[11], 8 variations du modèle ARMA (p,q) seront possibles. Le processus de prévision représenté par la figure 11 sera appliqué 4 fois avec,

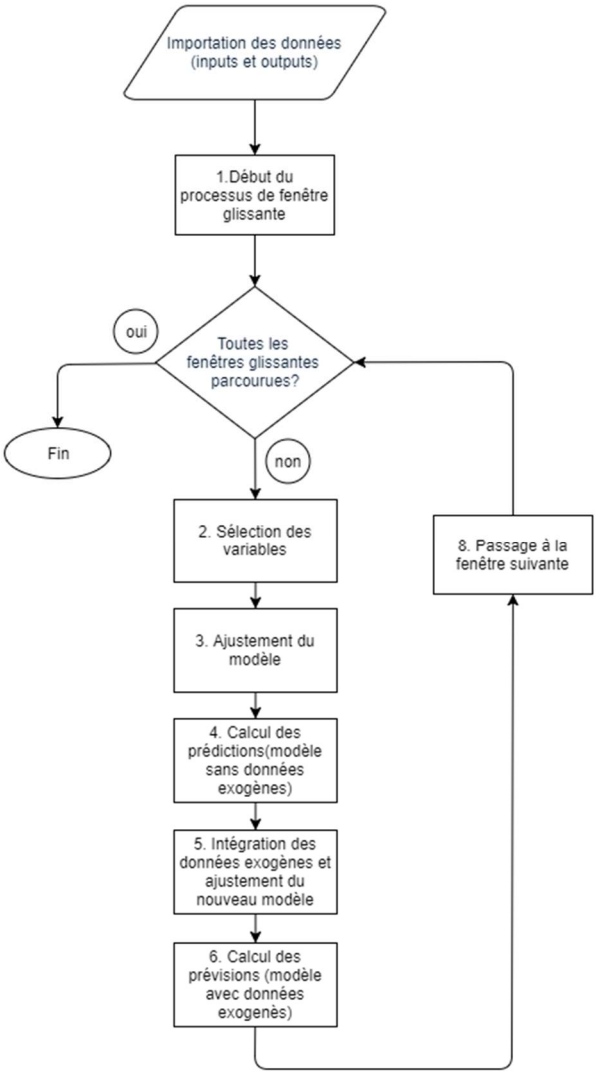
[9] Log-rendements en 5 minutes de l'euro-dollar OU du dollar-yen
[10] Variable « effet de surprise » des publications
[11] Deux possibilités : $2 \leq p \leq 5$ et $0 \leq q \leq 5$ OU $0 \leq p \leq 5$ et $0 \leq q \leq 5$

comme résultat, 8 séries de prévisions, prévisions qui seront comparées aux observations correspondantes pour établir les performances de chaque modèle.

## 6.4 Développement des modèles de régression logistique

Comme pour les modèles autorégressifs, le diagramme de flux ci-dessous reprend les différentes étapes qui permettent aux modèles de régression logistique de générer les prédictions *out-of-sample*.

*Figure 12 - Processus de prédiction de la direction des taux de change sur base des régressions logistiques*




Une fois les données importées, l'algorithme passe à la première fenêtre glissante.

Pour une fenêtre déterminée, les variables explicatives de la régression logistique correspondent aux *inputs* de la partie autorégressive du modèle ARMA (p,q). L'algorithme a donc besoin de connaître la valeur de l'ordre $p$ du modèle ARMA (p,q) pour ensuite préciser les *inputs* de la régression logistique qui sont les $p$ log-rendements des taux de change passés. La variable dépendante est la direction du taux de change décrite dans le chapitre précédent, la variable est donc dichotomique et la régression logistique binaire.

Après avoir divisé les observations en un *training set* et un *test set* dont les tailles ont été précisées dans la partie 6.1.1., l'algorithme ajuste le modèle de régression logistique à l'aide du *training set*. Pour chercher les paramètres optimaux, le modèle se base sur le modèle *glm* du paquet *caret* de R comprenant un algorithme des moindres carrés pondérés itérés, qui a été présenté dans la partie théorique de la régression logistique. Il ne reste plus qu'à calculer les prédictions et les comparer aux valeurs de la variable dépendante du *test set*.

Pour calculer les prédictions de la régression logistique qui intègre l'« effet de surprise » des publications macroéconomiques, il suffira simplement d'ajouter ces données sous forme d'une variable explicative supplémentaire. La régression logistique sera ajustée avec son nouvel ensemble de variables explicatives et les nouvelles prédictions seront calculées de la même manière qu'avec le modèle qui ne prend pas en compte l'« effet de surprise » des publications macroéconomiques.

Ces opérations sont répétées jusqu'à ce que toutes les fenêtres glissantes aient été parcourues.

Etant donné que l'algorithme représenté par le diagramme de flux ci-dessus sera exécuté une fois pour l'euro-dollar et une fois pour le dollar-yen, 4 ensembles de prédictions basés sur la régression logistique seront disponibles lors de l'analyse des résultats.

## 6.5 Développement des réseaux de neurones artificiels

Le diagramme de flux élaboré pour la régression logistique représente également la méthodologie utilisée pour calculer les prédictions à partir des réseaux de neurones artificiels.

Les données sont d'abord importées, la première fenêtre glissante est activée et l'algorithme sélectionne les variables de la même manière qu'avec la régression logistique.
Le modèle choisi pour calculer les prédictions est un perceptron multicouche avec une couche cachée. L'implémentation dans R se fait à l'aide de la fonction *train* du paquet *caret* combinée à la méthode *nnet*. Le réseau est caractérisé par :

- Une couche cachée.
- Une fonction d'activation pour la couche de sortie adaptée aux problèmes de classification binaire : la fonction logistique.
- Une foneffction d'entropie croisée comme fonction de coût.
- Un algorithme de rétropropagation combiné à l'algorithme de BFGS comme méthode d'optimisation pour ajuster les poids de connexion. Les poids de connexion sont choisis aléatoirement lors de l'initialisation puis l'algorithme ajuste les poids de connexion par itération jusqu'à ce qu'il converge ou atteigne la limite d'itérations. Dans ce cas-ci, les itérations seront limités à 500.
- Le nombre de neurones dans la couche cachée comme premier hyperparamètre.
- La modération des poids de connexion comme deuxième hyperparamètre.

Les hyperparamètres sont des paramètres qui ne peuvent pas être estimés directement lors du processus standard d'apprentissage. La technique appliquée pour déterminer le

nombre optimal de neurones dans la couche cachée et la modération des poids est la recherche par quadrillage (*grid search* en anglais). Elle consiste à énumérer l'ensemble de valeurs que les hyperparamètres pourront prendre pour ensuite les associer par combinaison, optimiser un modèle pour chaque combinaison et ne garder que la combinaison qui offre les meilleures performances. Les ensembles de valeurs à tester sont $\{1; 3; 5; 7; 10\}$ pour le nombre de neurones de la couche cachée et $\{0; 0{,}0001;\ 0{,}01\ ; 0{,}1\}$ pour la modération des poids.

A chaque nouvelle fenêtre, les observations sont réparties en un *training set,* un *validation set* et un *test set* dont les dimensions ont été précisées au point 6.1.1. L'algorithme commence par ajuster un réseau pour chaque combinaison d'hyperparamètres sur base du *training set*. Le *validation set* est ensuite utilisé pour sélectionner le réseau le plus performant sur base du taux de bonne classification. Finalement, l'algorithme effectue les prédictions sur le *test set*.

Dans un second temps, l'« effet de surprise » des publications macroéconomiques est ajouté comme variable explicative. Un perceptron multicouche est à nouveau ajusté pour chaque couple d'hyperparamètres et le plus performant, en termes de classification, est gardé pour générer les prédictions sur le *test set*.

Le processus sera généré pour l'euro-dollar puis pour le dollar-yen avec un total de 4 séries de prédictions comme pour la régression logistique.

Parce que les poids de connexion lors de la phase d'initialisation de l'algorithme de rétropopagation sont générés aléatoirement et qu'il est possible que la fonction d'erreur ait plusieurs minima locaux, deux réseaux de neurones avec les mêmes caractéristiques

peuvent avoir des résultats différents. Pour pallier à ce problème et avoir des performances plus stables, les perceptrons sont réappris à 5 reprises et les prédictions conservées lors de chaque répétition. La prédiction finale, pour chaque observation, n'est autre que la prédiction majoritaire parmi les 5 valeurs disponibles.

# 7 Mesures d'évaluation des performances

Les performances sont mesurées à l'aide, d'une part, du taux de bonne classification accompagné d'une matrice de confusion et, d'autre part, de la racine de l'erreur quadratique moyenne (notée RMSE pour *Root Mean Squared Error*) et de l'erreur moyenne absolue (notée MAE pour *Mean Absolute Error*), qui sont parmi les mesures les plus connues pour évaluer la précision de modèles régressifs.

Les mesures remplissent des objectifs distincts. Le taux de bonne classification est la mesure centrale permettant la comparaison entre les différents modèles, qu'ils soient basés sur la régression ou sur la classification afin de répondre aux objectifs fixés par le mémoire, tandis que le RMSE et le MAE poursuivront un but informationnel sur la qualité des modèles autorégressifs.

## 7.1 Taux de bonne classification

Pour mesurer et comparer les performances de tous les modèles, le taux de bonne classification de la direction des taux de change accompagné d'une matrice de confusion, sera utilisé :

*Tableau 2 - Matrice de confusion*

| | Prédiction: Direction Positive | Prédiction: Direction Négative |
|---|---|---|
| Réalité: Direction Positive | Vrais Positifs (TP) | Faux Négatifs (FN) |
| Réalité: Direction Négative | Faux Positifs (FP) | Vrais Négatifs(TN) |

$$Taux\ de\ bonne\ classification = \frac{Vrais\ Positifs + Faux\ Négatifs}{Population\ totale}$$

Bien que les *outputs* des modèles autorégressifs soient continus et ceux des algorithmes d'apprentissage soient binaires, il est possible, après transformation, de comparer directement les performances de tous les modèles en utilisant le taux de bonne classification. En effet, les modèles autorégressifs calculent, pour chaque prévision, un taux de change. Il suffira de prendre uniquement la direction de ce taux pour avoir le même *output* que les algorithmes d'apprentissage et d'avoir à disposition, pour tous les modèles, l'information suffisante pour calculer les différents éléments de la matrice de confusion et le taux de bonne classification.

Un autre avantage du taux de bonne classification est qu'il permet une interprétation rapide de la capacité d'un modèle à prévoir la direction des taux de change. Il contient donc une information utile à la poursuite des objectifs de ce mémoire.

Les informations contenues dans la matrice de confusion permettront également d'en savoir plus sur la qualité des prédictions. Les matrices font, par exemple, apparaître les modèles qui arrivent à bien répartir les prédictions entre les différentes classes ou, au contraire, les modèles qui ont tendance à surreprésenter les prédictions dans une classe.

## 7.2 RMSE et MAE

En complémentarité avec le taux de bonne classification, les performances des méthodes autorégressives seront analysées à l'aide des mesures de précision RMSE et du MAE.

Il n'est pas possible d'appliquer ces mesures aux résultats des régressions logistiques et des réseaux de neurones artificiels étant donné qu'elles sont basées sur les erreurs de prévision[12], alors que l'*output* des méthodes de classificationne donne pas cette information.

[12] L'erreur de prévision est calculée comme l'écart entre la valeur réelle et la valeur estimée d'une série temporelle.

Etant donné $y_i$ et $\hat{y}_i$, respectivement la i$^{\text{ème}}$ observation et la i$^{\text{ème}}$ prévision d'une série temporelle. L'erreur de prévision est égale à $e_i = y_i - \hat{y}_i$ et les formules du RMSE et du MAE sont :

$$MAE = mean(|e_i|)$$

$$RMSE = \sqrt{mean(e_i^2)}$$

Plusieurs éléments ressortent de l'analyse des formules de ces deux indicateurs. Les modèles les plus performants seront ceux dont les valeurs de ces deux indicateurs seront les plus petites. Etant donné que leurs formules dépendent de l'échelle des données, le RMSE et le MAE ne pourront pas être utilisés pour comparer les performances de séries temporelles dont les données ne sont pas sur la même échelle. Comparativement au MAE, le RMSE donne plus de poids aux erreurs larges. Il sera donc plus pertinent lorsque de grandes erreurs de prévision sont indésirables.

# 8 Evaluation des modèles

Dans ce chapitre, on va évaluer la capacité des différents modèles à prédire la direction des taux de change. Plusieurs éléments rendent cette évaluation particulière :

- L'évaluation se fait à la fois pour des méthodes autorégressives et des méthodes de classification.
- Les prédictions ont été réalisées à l'aide d'une approche par fenêtres glissantes.
- Ce mémoire porte un intérêt particulier aux prédictions qui suivent des annonces macroéconomiques.

Dans un premier temps, on va s'intéresser à la qualité des méthodes autorégressives, des régressions logistiques puis des perceptrons multicouches. Ensuite, on va analyser et comparer les capacités prédictives des différentes méthodes sur l'ensemble des observations. La troisième et dernière partie étudie l'impact de l'« effet de surprise » des publications sur les prédictions, en comparant notamment les performances des modèles ajustés avec et sans cette variable

## 8.1 Evaluation des modèles ARMA

### 8.1.1 Diagnostic des résidus

L'évaluation des processus autorégressifs commence par un diagnostic des résidus. Chaque résidu représente la composante imprévisible associée à une observation et il se calcule comme la différence entre les valeurs effectives et les valeurs ajustées des observations : $e_i = y_i - \hat{y}_i$ [5]. Pour savoir si le modèle a bien réussi à prendre en compte l'information autorégressive contenue dans les données, le diagnostic s'appuie sur l'analyse du corrélogramme des résidus et un test de Ljung-Box défini ci-après. Si, à la suite du diagnostic, les résidus semblent se comporter comme un processus de bruit blanc alors on peut considérer que le modèle a correctement capturé l'information autorégressive disponible.

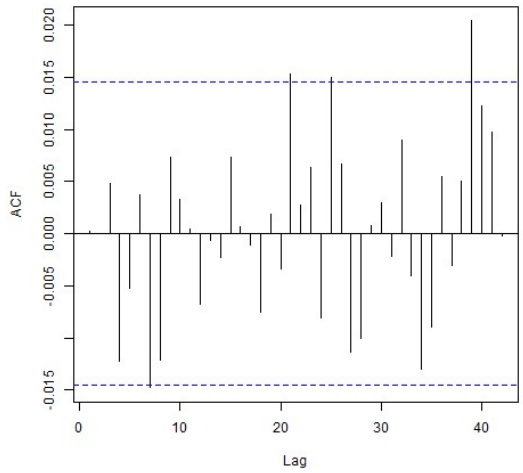
Le test de Ljung-Box est basé sur la statistique suivante :

$$Q^* = T(T+2)\sum_{k=1}^{h}(T-k)^{-1}r_k^2$$

où $T$ est la longueur de la série temporelle, $r_k$ est le $k$ième coefficient d'autocorrélation des résidus et $h$ est le nombre de décalage. Athanasopoulos, G., & Hyndman, R. (2013) expliquent que de grandes valeurs pour la statistique $Q^*$ appliquée à des séries de résidus montrent qu'il y a des autocorrélations significatives. Ils proposent de tester la statistique par rapport à une distribution $\chi^2$ avec $h-K$ degrés de liberté où $K$ est le nombre de paramètres estimés dans le modèle. Pour le choix de $h$, ils conseillent de travailler avec $h = 10$. Les tests de Box-Ljung sont interprétés en comparant la p-valeur de la statistique par rapport à un seuil prédéfini, on choisit généralement 0,05 comme seuil. Une p-valeur inférieure au seuil choisi indique la possibilité d'autocorrélation non-nulle entre les $h$ premiers décalages.

Etant donné que les fonctions d'autocorrélation et le test de Box-Ljung sont analysés pour 4 processus basés sur des modèles ARMA (p,q) distincts (selon la paire de devises étudiée et selon que l'ordre $p$ soit compris entre 0 et 5 ou 2 et 5) et qu'il y a une ACF pour chaque fenêtre glissante, il n'est pas possible d'analyser chaque ACF. Sauf quelques exceptions, la majorité des ACF sur les résidus semblent montrer que les résidus sont proches d'un bruit blanc avec la plupart des barres situées à l'intérieur des seuils critiques et quelques barres légèrement inférieures ou supérieures aux seuils critiques comme par exemple pour l'ACF ci-dessous.

*Figure 13 - Exemple d'ACF sur les résidus*




Cette analyse est cohérente par rapport aux résultats du test de Box-Ljung où le nombre de p-valeurs inférieures à 0,05 est relativement faible : sur 171 fenêtres glissantes, 27 ont une p-valeur inférieure à 0,05 pour le modèle ARMA appliqué à l'euro-dollar et avec l'ordre $p$ compris entre 0 et 5, 26 pour celui appliqué à l'euro-dollar mais avec l'ordre $p$ compris entre 2 et 5, 21 pour le modèle appliqué au dollar-yen avec $p$ entre 0 et 5 et enfin 20 pour le deuxième modèle ARMA appliqué au dollar-yen avec $p$ compris entre 2 et 5.

### 8.1.2 Analyse des ordres $p$ et $q$

Le diagnostic des résidus jugé satisfaisant, on passe à l'analyse des caractéristiques principales des modèles ARMA (p,q) : la taille des ordres $p$ et $q$. Pour rappel, la combinaison finale des ordres $p$ et $q$, choisie lors de chaque fenêtre glissante est celle qui correspond au modèle ARMA (p,q) avec le plus petit AIC.

En observant la répartition des combinaisons (p,q) finales pour les modèles ARMA où $0 \leq p \leq 5$ disponible dans l'annexe **g**, on constate qu'il y a 20 fenêtres glissantes avec l'euro-dollar et 2 fenêtres glissantes avec le dollar-yen pour lesquelles la combinaison

(p,q) qui minimise l'AIC est (0,0). En d'autres mots, le processus de génération des données est, selon l'AIC, le mieux représenté par des bruits blancs pour ces 22 fenêtres glissantes. Lorsque le processus choisi est un modèle ARMA (0,0), les prévisions sont calculées de la même manière qu'avec le *benchmark* modèle « moyenne ».

L'analyse des combinaisons finales avec $0 \leq p \leq 5$ montre aussi que, pour la grande majorité des fenêtres glissantes, les meilleurs modèles selon l'AIC ont au moins l'ordre $p$ ou l'ordre $q$ différent de 0. Dans la plupart des cas, prendre en compte l'autorégression des données et/ou des résidus permettrait de réduire la perte d'information.

### 8.1.3 Analyse des performances

On commence par analyser la relation entre le MAE, le RMSE et le taux de bonne classification à l'aide des 4 matrices de diagrammes de dispersion ci-dessous (deux pour l'euro-dollar avec $0 \leq p \leq 5$ et $2 \leq p \leq 5$, deux pour le dollar-yen avec $0 \leq p \leq 5$ et $2 \leq p \leq 5$). Les mesures sont calculées pour les *test sets* et chaque diagramme contient 171 points correspondant aux 171 fenêtres glissantes. Les valeurs des mesures d'erreur RMSE et MAE sont fortement corrélées avec des corrélations supérieures à 0,99, tandis que la relation entre les mesures d'erreur et le taux de bonne classification est légèrement négative.

*Figure 14 - (EUR/USD) Diagrammes de dispersion entre RMSE, MAE et taux de bonne classification pour ARMA 0≤p≤5*

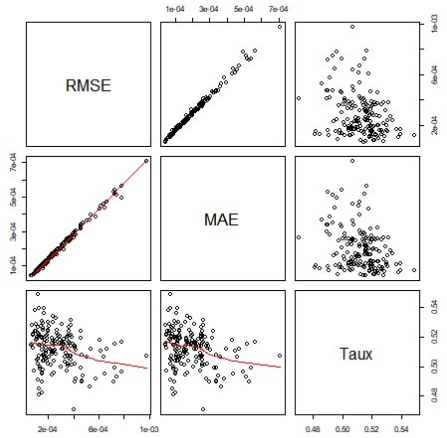



*Figure 15 - (EUR/USD) Diagrammes de dispersion entre RMSE, MAE et taux de bonne classification pour ARMA 2≤p≤5*

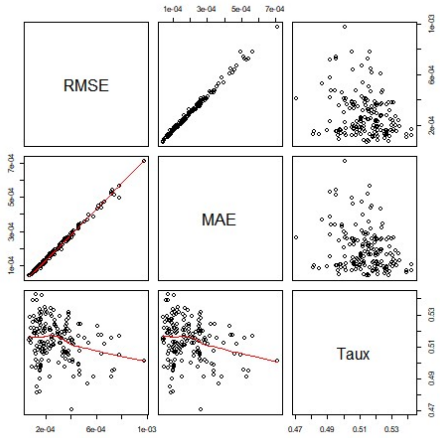



*Figure 16 -(USD/JPY) Diagrammes de dispersion entre RMSE, MAE et taux de bonne classification pour ARMA 0≤p≤5*

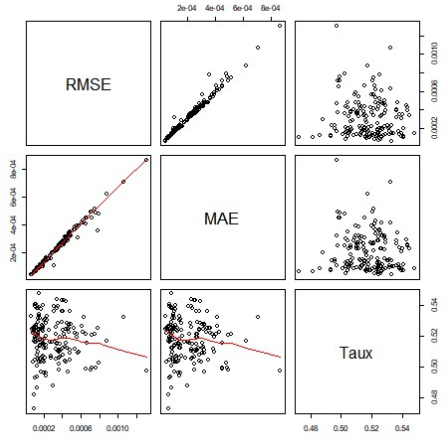
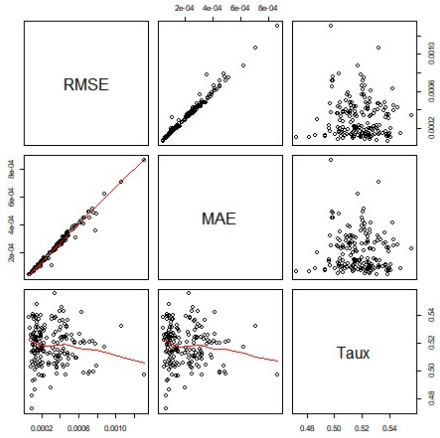



*Figure 17 - (USD/JPY) Diagrammes de dispersion entre RMSE, MAE et taux de bonne classification pour ARMA 2≤p≤5*




A l'aide du RMSE et du taux de bonne classification, nous allons évaluer et comparer la qualité des modèles ARMA et du modèle « moyenne » d'un point de vue économétrique et d'un point de vue « classification ». Pour chaque paire de devises, l'évaluation s'appuie sur deux diagrammes en boîte (également appelés boîtes à moustaches ou *boxplots*) qui représentent schématiquement la distribution du RMSE et du taux de

bonne classification des *test sets* sur l'ensemble des fenêtres glissantes pour les 3 modèles autorégressifs (ARMA (p,q) où $0 \leq p \leq 5$, ARMA (p,q) où $2 \leq p \leq 5$ et modèle « moyenne ») ainsi qu'un graphique qui compare les RMSE des 3 modèles par fenêtre glissante.

*Figure 18 - (EUR/USD) Diagrammes en boîte représentant la distribution du RMSE pour les modèles autorégressifs*

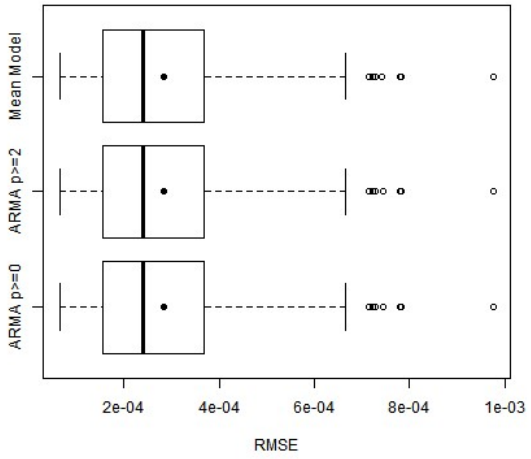



*Figure 19 - (EUR/USD) Diagrammes en boîte représentant la distribution du taux de bonne classification pour les modèles autorégressifs*

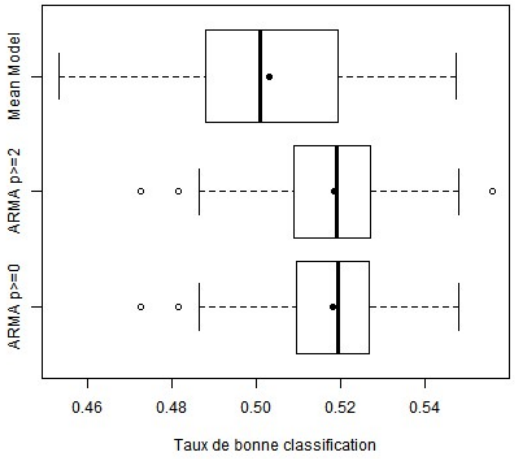



*Figure 20 - (USD/JPY) Diagrammes en boîte représentant la distribution du RMSE pour les modèles autorégressifs*

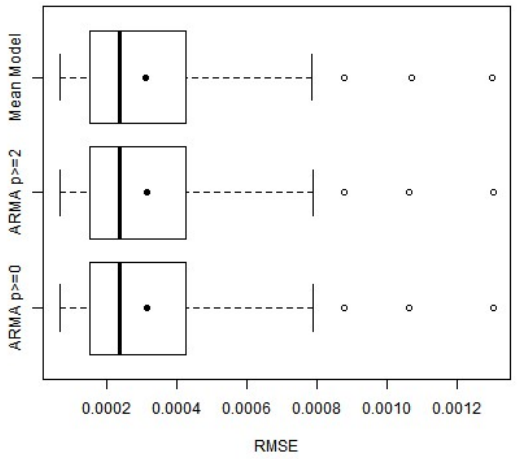



*Figure 21 - (USD/JPY) Diagrammes en boîte représentant la distribution du taux de bonne classification pour les modèles autorégressifs*

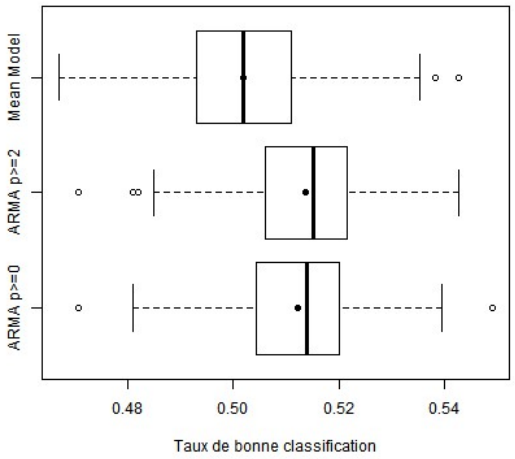


*Figure 22 - (EUR/USD) Comparaison du RMSE des 3 modèles autorégressifs par fenêtre glissante*

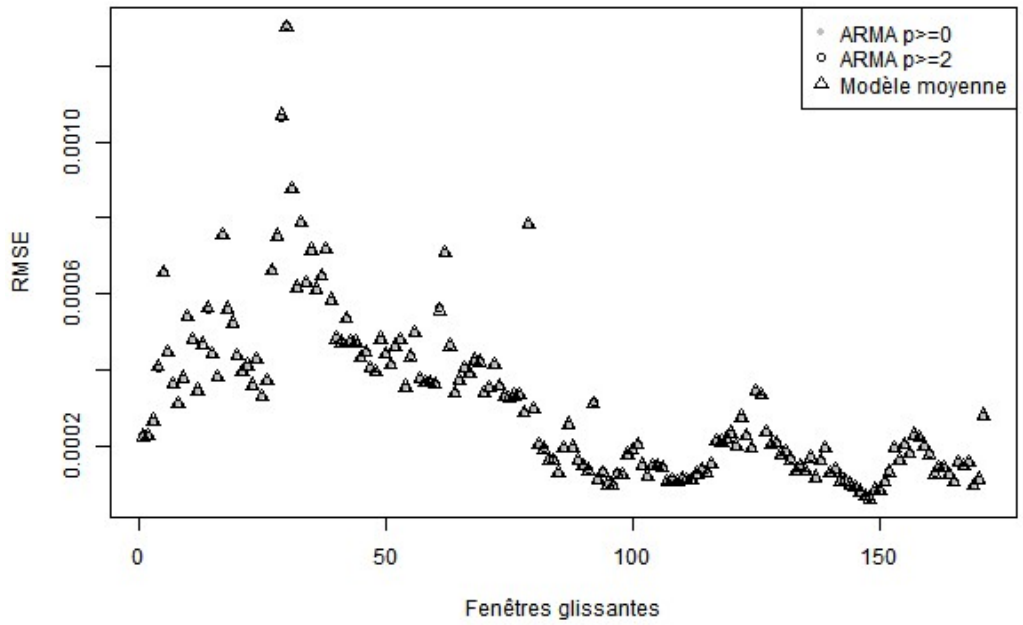
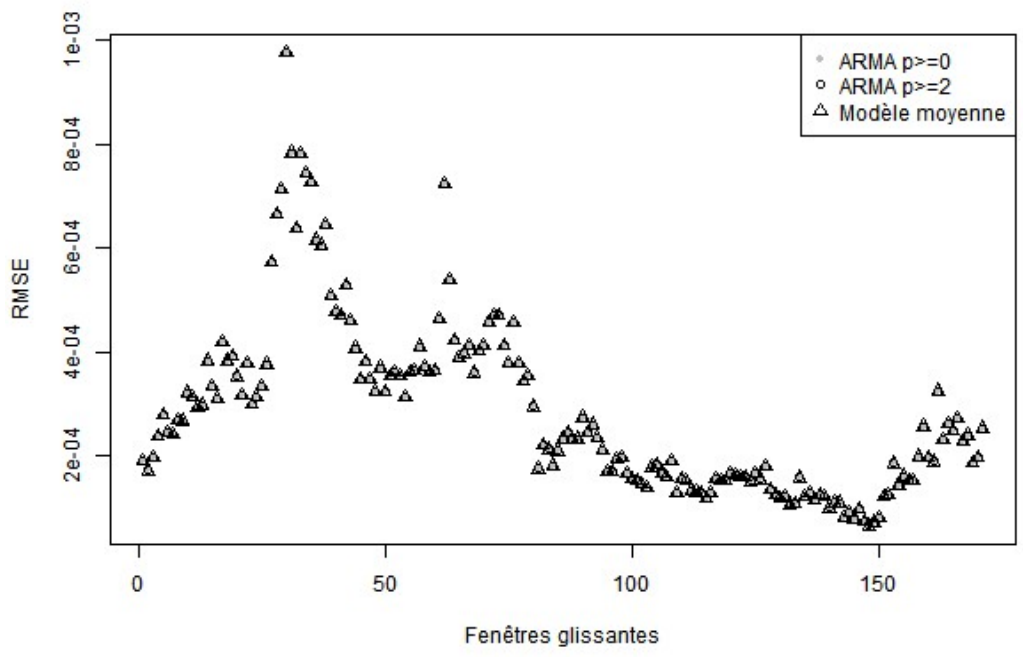



*Figure 23 - (USD/JPY) Comparaison du RMSE des 3 modèles autorégressifs par fenêtre glissante*




En regardant d'abord les *boxplots* des RMSE, on remarque que la distribution du RMSE pour les trois modèles est quasiment identique. Cette observation est confirmée

par le graphique représentant les RMSE en fonction des fenêtres glissantes, dans lequel les RMSE des 3 modèles sont quasi égaux tant pour les modèles ARMA que pour le modèle « moyenne ». La situation est différente pour les *boxplots* des taux de bonne classification : qu'il s'agisse de l'euro-dollar ou du dollar-yen, la moyenne et la médiane des modèles ARMA sont supérieures à celles du modèle « moyenne » et l'écart inter-quartile est plus important pour le taux de bonne classification du modèle « moyenne ». La dispersion, la médiane et la moyenne pour les modèles ARMA (p,q) où $0 \leq p \leq 5$ et $2 \leq p \leq 5$ sont pratiquement équivalentes tant pour le dollar-yen que l'euro-dollar.

Dans la suite du chapitre, nous n'utiliserons plus que les prédictions d'un des deux modèles ARMA (p,q)( celui où l'ordre $p$ est compris entre 2 et 5).

## 8.2 Contrôle du surapprentissage

Pour contrôler le surapprentissage des régressions logistiques et des réseaux de neurones artificiels, nous allons comparer leurs performances en apprentissage par rapport aux performances en test.

Etant donné leurs complexités, le problème d'*overfitting* concerne particulièrement les perceptrons multicouches. Nous espérons que le problème de surapprentissage ait été correctement contenu par les différents moyens mis en place pour le combattre :

- Une étape de validation dans laquelle le nombre de neurones dans la couche cachée et la modération de poids sont choisis.
- Un nombre maximum d'itérations
- Une prédiction « finale » égale à la prédiction majoritaire des réseaux de neurones artificiels réappris plusieurs fois.

Pour évaluer le surapprentissage, on va commencer par calculer la différence entre les taux de bonne classification sur les *training sets* et sur les *test sets* pour chaque fenêtre glissante. On les regroupe ensuite sous forme de boîtes à moustaches classées selon la paire de devises et selon la technique de classification. Les 4 boîtes à moustache sont représentées dans les deux cadres ci-dessous : un cadre pour l'euro-dollar et un cadre pour le dollar-yen.

*Figure 24 - (EUR/USD) Diagrammes en boîte représentant la distribution des différences entre les taux de bonne classification du training set et du test set*

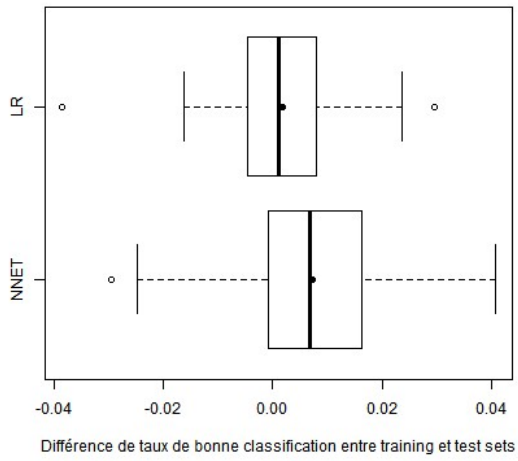
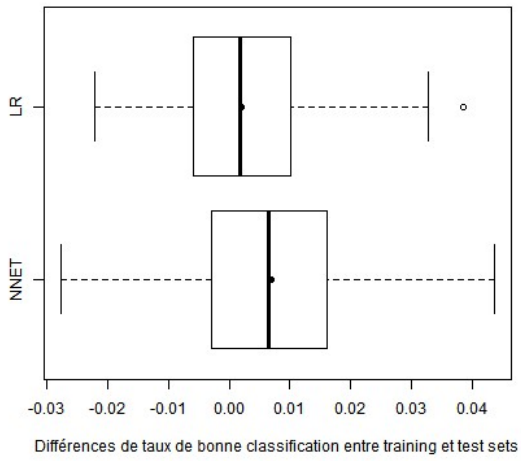



*Figure 25 - (USD/JPY) Diagrammes en boîte représentant la distribution des différences entre les taux de bonne classification du training set et du test set*




Comme attendu, la différence entre les performances en apprentissage et celles en test est moins grande en moyenne pour la régression logistique : elle est d'environ 0,2% pour l'euro-dollar et pour le dollar-yen. Avec 0,72% pour l'euro-dollar et 0,6% pour le dollar-yen, la différence moyenne est déjà plus importante pour les perceptrons multi-couches mais elle reste à un niveau limité. De plus, la dispersion des valeurs pour les perceptrons est assez confinée avec 50% des valeurs comprises dans un intervalle plus petit que 2% pour les deux paires de devises et des valeurs extrêmes situées à -3% et un

peu plus de 4%. Si les réseaux de neurones développés dans ce mémoire sont plus sensibles que les régressions logistiques au surapprentissage, la différence de performances en apprentissage et en test reste toutefois tolérable.

## 8.3 Corrélation des taux de bonne classification

Les deux matrices ci-dessous représentent, pour les deux paires de devises, la relation entre les taux de bonne classification des 4 types de modèles à l'aide de diagrammes de dispersion. Chaque point représente un couple de taux de bonne classification de 2 modèles obtenus lors d'une fenêtre glissante spécifique.

*Figure 26 - (EUR/USD) Matrice de diagrammes de dispersion présentant les taux de classification des modèles « moyenne », ARMA, régressions logistiques et réseaux de neurones par fenêtre glissante.*

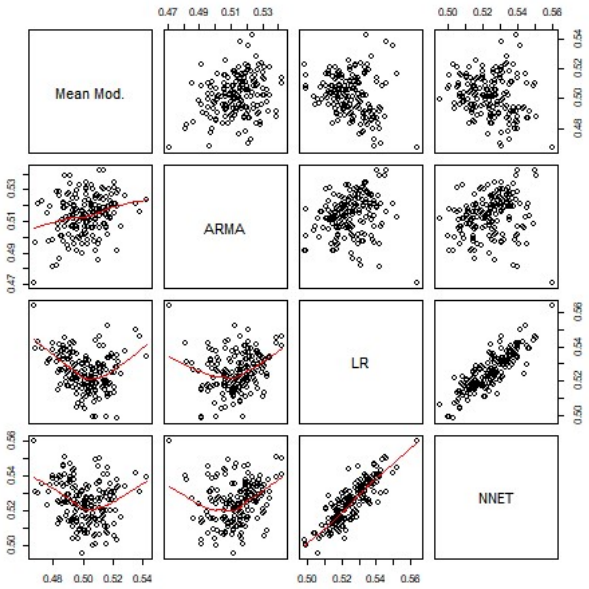
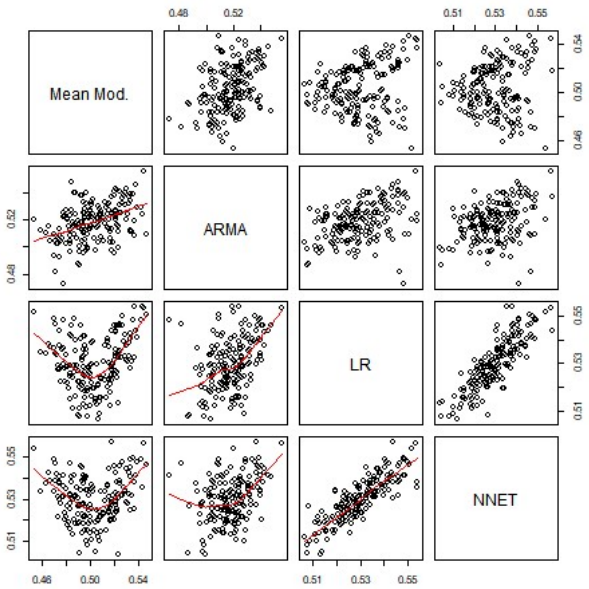


*Figure 27 - (EUR/USD) Matrice de diagrammes de dispersion présentant les taux de classification des modèles « moyenne », ARMA, régressions logistiques et réseaux de neurones par fenêtre glissante.*




En observant les diagrammes de dispersion entre la régression logistique et les perceptrons multicouches, on remarque une corrélation positive forte. Si on calcule les coefficients de corrélation entre les taux, cela donne 0,856 pour l'euro-dollar et 0,857 pour le dollar-yen. Si les performances des régressions logistiques augmentent alors les performances des réseaux de neurones auront aussi tendance à monter. Pour le reste, les relations entre taux sont moins évidentes. On peut toutefois noter une légère corrélation entre les modèles « moyenne » et ARMA avec des coefficients de corrélation égaux à 0,26 pour l'euro-dollar et 0,42 pour le dollar-yen. On note également une légère corrélation des modèles ARMA avec les deux méthodes de classification classiques.

## 8.4 Comparaison des performances sur l'ensemble des prédictions

Dans cette partie, on compare les capacités des différents processus à prédire la direction des taux de change sur l'ensemble des prédictions. Pour nous aider, nous avons à disposition les matrices de confusion et les taux de bonne classification. Pour rappel, on ne s'intéresse pour le moment qu'aux performances des modèles sans la variable exogène qui représente l'« effet de surprise » des publications macroéconomiques.

*Tableau 3 -(EUR/USD) Matrices de confusion sur l'ensemble des prédictions*

| *Modèles "moyenne"* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 173922 | 131060 |
| *Réalité: Direction -* | 177703 | 137190 |

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 167087 | 137895 |
| *Réalité: Direction -* | 163551 | 151342 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 111887 | 193095 |
| *Réalité: Direction -* | 101217 | 213676 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 128884 | 176098 |
| *Réalité: Direction -* | 118566 | 196327 |

*Tableau 4 - (USD/JPY) Matrices de confusion sur l'ensemble des prédictions*

| *Modèles "moyenne"* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 170205 | 131072 |
| *Réalité: Direction -* | 177795 | 141703 |

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 163990 | 136387 |
| *Réalité: Direction -* | 162223 | 157275 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 84044 | 216333 |
| *Réalité: Direction -* | 75058 | 244440 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 96553 | 203824 |
| *Réalité: Direction -* | 87614 | 231884 |

Commençons par analyser les informations contenues dans les matrices d'informations. Nous constatons que tous les modèles n'ont, à peu de chose près, pas plus de difficulté à prédire la bonne direction qu'il s'agisse des prédictions à la baisse ou à la hausse. On

observe aussi que les processus basés sur la régression logistique et les réseaux de neurones artificiels ont tendance à prédire beaucoup plus de directions à la baisse qu'à la hausse, tant pour l'euro-dollar que le dollar-yen, alors que les modèles ARMA et « moyenne » enregistrent un nombre plus élevé de prédictions à la hausse qu'à la baisse mais de manière modérée. En analysant les matrices de confusion des régressions logistiques et des perceptrons multicouches réparties par fenêtre glissante dans l'annexe **h**, on remarque que les modèles ont tendance à surreprésenter nettement une direction par rapport à l'autre.

*Tableau 5 - Taux de bonne classification des modèles prédictifs sur l'ensemble des prédictions*

| | *EUR/USD* | *USD/JPY* |
|---|---|---|
| *Modèles "moyenne"* | 0,5019 | 0,5032 |
| *Modèles ARMA* | 0,5137 | 0,5183 |
| *Régressions logistiques* | 0,5252 | 0,5299 |
| *Perceptrons multicouches* | 0,5246 | 0,5299 |

Malgré ce problème de surreprésentation, les régressions logistiques et les réseaux de neurones artificiels enregistrent de meilleures performances que les modèles autorégressifs à moyennes mobiles et les modèles « moyenne ». Le taux de bonne classification pour l'euro-dollar et le dollar-yen est à peu près équivalent entre les régressions logistiques (52,52 et 53%) et les perceptrons multicouches (52,5% et 53%). Par contre, la différence par rapport aux modèles ARMA devient significative avec plus de 1% de bonnes prédictions supplémentaires. Le potentiel prédictif de ces modèles est encore plus marquant quand on le compare au *benchmark*, le modèle « moyenne » dont les performances approchent les 50% de bonnes classifications.

Il est difficile de déterminer si la régression logistique et les réseaux de neurones offrent de meilleurs résultats grâce à leur capacité à modéliser des relations non linéaires entre variables explicatives et variables à expliquer ou simplement parce qu'ils sont construits spécifiquement pour classer les observations en direction haussière ou baissière, alors que les modèles autorégressifs font des prévisions de la variation des taux de change, à partir desquelles on a extrait la direction. Une autre information qui ressort des résultats est que, malgré une plus grande complexité, les perceptrons multicouches ne montrent pas d'amélioration des capacités prédictives par rapport à la régression logistique. Avec un taux de bonne classification significativement supérieur à celui du *benchmark*, les méthodes de classification montrent que, rien qu'avec les variations historiques des prix, il est possible de trouver de l'information pour mieux prédire la direction des taux de change.

## 8.5 Impact de l'« effet de surprise » des publications macroéconomiques

### 8.5.1 Interprétation des coefficients

Pour rappel, la variable « effet de surprise » des publications macroéconomiques a été construite sur base des travaux d'Andersen et *al.* (année). La variable rassemble les observations de 11 indicateurs économiques qui sont censés avoir un impact significatif sur le prix du dollar. Selon la manière dont nous l'avons construite, la variable devrait avoir un impact positif sur la valeur du dollar lorsqu'elle est égale à 1 et un impact négatif lorsqu'elle est égale à -1.

Il est possible d'interpréter les coefficients et en particulier le signe des coefficients des régressions logistiques et des modèles ARMA. Nous allons utiliser cette information pour vérifier la cohérence des résultats avec la construction théorique de la variable « effet de surprise » des publications macroéconomiques.

L'annexe **i** répertorie les coefficients de la variable exogène obtenus après ajustement pour les modèles ARMA et pour les régressions logistiques de chaque fenêtre glissante. Sur 171 fenêtres glissantes, les modèles ARMA et les régressions logistiques basés sur le dollar-yen ont respectivement 164 et 169 fenêtres dans lesquelles le coefficient de la variable exogène est positif. Comme le dollar-yen exprime la valeur d'un dollar en yen, ce résultat montre bien, dans la majorité des cas, l'impact de la variable est positif sur la valeur du dollar. Pour l'euro-dollar, on s'attend à une majorité de fenêtres glissantes avec un coefficient négatif, ce qui est le cas pour 132 fenêtres avec les modèles ARMA et 136 fenêtres avec les régressions logistiques.

### 8.5.2 Performances sur l'ensemble des prédictions

Avant d'examiner si l'ajout de la variable exogène aux variables explicatives permet de mieux prédire la direction des taux de change suite à des annonces macroéconomiques, on va étudier l'impact de son intégration sur les performances globales des modèles. Dans cette optique, on compare les matrices de confusion et les taux de bonne classification, sur la totalité des prédictions *out-of-sample* disponibles, des modèles sans la variable exogène et des modèles avec la variable exogène.

*Tableau 6 - Taux de bonne classification des modèles prédictifs avec la variable "effet de surprise" des publications macroéconomiques sur l'ensemble des prédictions.*

| | *EUR/USD* | *USD/JPY* |
|---|---|---|
| *Modèles ARMA* | 0,5136 | 0,5187 |
| *Régressions logistiques* | 0,5255 | 0,5303 |
| *Perceptrons multic.* | 0,5252 | 0,5297 |

Les taux de bonne classification et la répartition entre prédictions à la baisse ou à la hausse sont quasi similaires pour les modèles avec ou sans variable exogène. La variable exogène n'a pas une grande incidence sur la manière dont les modèles assimilent des données autorégressives.

### 8.5.3 Comparaison des performances post-annonces macroéconomiques

Dans la dernière partie de ce chapitre, nous allons concentrer l'analyse sur les prédictions faites uniquement à la suite de publication d'annonces macroéconomiques des 11 indicateurs qui composent la variable « effet de surprise ». En d'autres termes, on examine uniquement les prédictions sur les observations où la valeur de la variable exogène est différente de 0. Les observations en question sont au nombre de 932.

A l'aide des matrices de confusion réunies dans l'annexe **k**, nous pouvons examiner la répartition des observations par rapport à la valeur effective de la direction des taux de change. Pour le dollar-yen, 463 observations enregistrent une direction haussière et 469 une direction baissière. Par contre, pour l'euro-dollar, la répartition est moins équitable avec 431 observations dont la direction est haussière et 501 avec une direction baissière.

Il est intéressant de noter que, comme sur l'ensemble des prédictions, les prédictions haussières sont sous-représentées avec les régressions logistiques et les réseaux de neurones artificiels qui ne prennent pas en compte la variable « effet de surprise ». Cependant, le ratio entre prédictions haussières et baissières s'équilibre dès lors qu'on ajoute la variable exogène aux autres variables explicatives. A titre d'exemple, on passe de 35,6% de prédictions à la hausse à 50,7% pour les régressions logistiques avec l'euro-dollar. Du côté des modèles ARMA, le pourcentage de directions haussières reste toujours proche des 50%. Enfin, peu importe le modèle et peu importe la paire de devises, les prédictions haussières et baissières correctement classées augmentent significativement une fois la variable « effet de surprise » intégrée.

*Tableau 7 - Taux de bonne classification des modèles prédictifs SANS la variable "effet de surprise" sur les prédictions post-annonces*

| | *EUR/USD* | *USD/JPY* |
|---|---|---|
| *Modèles ARMA* | 0,5139 | 0,5032 |
| *Régressions logistiques* | 0,5118 | 0,5086 |
| *Perceptrons multic.* | 0,5107 | 0,516 |

*Tableau 8 - Taux de bonne classification des modèles prédictifs AVEC la variable "effet de surprise" sur les prédictions post-annonces*

| | *EUR/USD* | *USD/JPY* |
|---|---|---|
| *Modèles ARMA* | 0,6352 | 0,7221 |
| *Régressions logistiques* | 0,633 | 0,736 |
| *Perceptrons multic.* | 0,617 | 0,7189 |

Le tableau ci-dessus regroupe les taux de bonne classification des différents modèles avec et sans la variable « effet de surprise » pour les 932 observations post-annonces macroéconomiques. Comme nous l'avions déjà expliqué lors de l'analyse des matrices de confusion, il y a une nette amélioration des taux de bonne classification une fois la variable exogène ajoutée, que ce soit pour les processus qui fonctionnent à l'aide de modèles ARMA, des réseaux de neurones artificiels ou des régressions logistiques. Quand on analyse les performances des modèles avec la nouvelle variable explicative par paire de devises, les taux de bonne classification sont relativement proches avec des différences de l'ordre de 1 ou 2 pourcents. Avec 61,7% pour l'euro-dollar et 71,9% pour le dollar-yen, les processus qui fonctionnent à l'aide des perceptrons multicouches offrent les moins bons résultats. Les modèles ARMA viennent en seconde position avec 63,5% de bonne classification pour l'euro-dollar et 72,2% pour le dollar-yen. Les régressions logistiques présentent donc les meilleurs résultats avec des taux de bonne classification de 63,3% pour l'euro-dollar et 73,6% pour le dollar-yen.

# 9 Profitabilité des modèles et efficience des marchés

Comme expliqué dans l'introduction, un des objectifs du mémoire est de confronter nos résultats aux hypothèses de marché efficient. Nous avons notamment vu que Jensen (1978) définit un marché efficient, dans sa forme semi-forte, comme un marché où il n'est pas possible de générer un profit économique sur la base des informations publiquement disponibles. Si l'hypothèse d'efficience semi-forte est avérée, alors les prédictions de nos modèles qui sont ajustés à l'aide d'informations publiques ne devraient pas permettre de générer des profits.

Pour calculer la profitabilité des modèles, nous allons simuler l'évolution d'un capital de départ de 1.000 euros qui sera utilisé pour acheter et vendre de l'euro-dollar et du dollar-yen selon les prédictions données par les différents modèles au moment des publications macroéconomiques. La simulation est réalisée avec des frais de transaction : nous avions vu que Dukascopy annonçait des frais à hauteur de 10 dollars pour un million de dollars échangés, ce qui revient à des frais de 0,001% par transaction. La simulation fonctionne comme suit : si un modèle prédit à la fermeture de la période $t-1$ que le taux de change va augmenter, alors la simulation passe un ordre d'achat à l'ouverture de la période $t$. Tandis que si le modèle prédit à la fermeture de la période $t-1$ que le taux de change va diminuer, alors la simulation passe un ordre de vente à l'ouverture de la période $t$. Les achats se font toujours par rapport au cours de l'*ask* et les ventes au cours du *bid*. Pour rappel, le *bid* et l'*ask* représentent respectivement le prix auquel le marché, ici représenté par le courtier, est prêt à acheter ou vendre la devise de base.

*Figure 28 - (EUR/USD) Simulation de l'évolution d'un capital de 1000 euros sur base des prédictions des différents modèles*

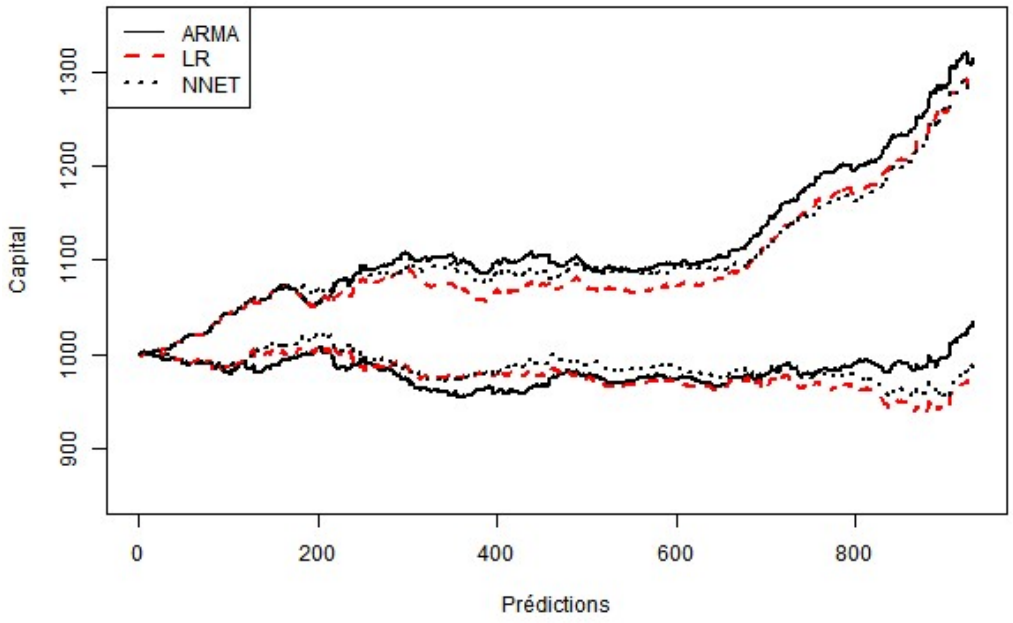
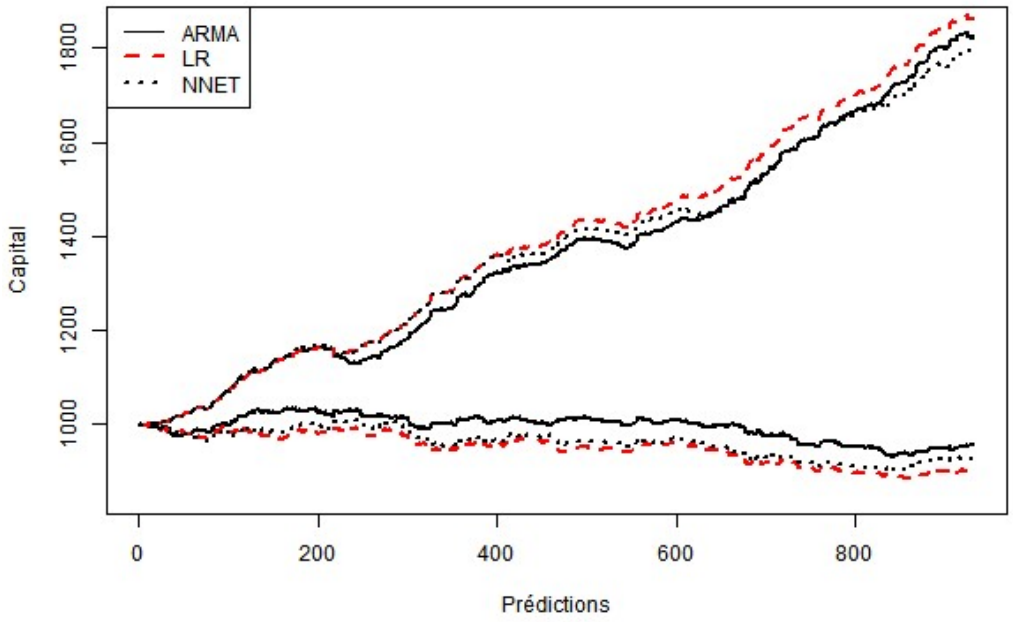



*Figure 29 - (USD/JPY) Simulation de l'évolution d'un capital de 1000 euros sur base des prédictions des différents modèles*



Les graphiques ci-dessus représentent l'évolution du capital par modèle prédictif. Par paires de devises, on retrouve 6 courbes : deux pour les modèles ARMA avec et sans la variable « effet de surprise » des publications macroéconomiques, deux pour les régressions logistiques et deux pour les réseaux de neurones artificiels. Sans surprise, le profit des processus prédictifs, qui intègrent la variable « effet de surprise » des publications macroéconomiques, est nettement supérieur à celui de ceux qui ne la prennent pas en compte. A l'exception des modèles ARMA qui permettent de dégager un petit profit sur l'euro-dollar, les modèles sans la variable exogène entraînent systématiquement une réduction du capital de départ : les profits réalisés grâce aux bonnes prédictions ne permettent pas de compenser les pertes faites sur le reste des prédictions auxquelles on ajoute le coût du *spread* et les frais de transaction. Pour les modèles qui prennent en compte « l'effet de surprise » des publications macroéconomiques, les profits les plus importants sont obtenus lorsque les taux de bonne classification sont les plus élevés : le processus basé sur des modèles ARMA obtient les meilleurs résultats pour l'euro-dollar avec 30% de profit et le processus basé sur des régressions logistiques, qui donnait le meilleur taux de bonne classification parmi tous les modèles, donne encore de meilleurs résultats pour le dollar-yen avec 85% de profit.

Ces résultats sont très encourageants en termes de profitabilité et semblent contester l'hypothèse de marché en forme semi-forte pour les deux taux de change étudiés. Cependant, plusieurs éléments pourraient réduire les profits réels par rapport à ceux enregistrés avec les simulations que nous avons réalisées :

- La vitesse d'exécution réelle des ordres d'achat et de vente pourrait être plus faible que celle que nous avons prévue.
- Certaines données que nous avons récoltées pourraient être erronées. Par exemple, il faudrait vérifier que les annonces macroéconomiques sont précisément disponibles à l'heure fournie par Bloomberg.

- La vitesse de transmission de l'information entre le moment où un modèle fournit une prédiction pour la direction des taux de change et le moment où cette information est reçue.

# 10 Conclusion

Les marchés financiers sont formés par les interactions des différents acteurs financiers dont les décisions sont prises en fonction de l'information qu'ils ont à leur disposition.

Partant de ce constat, nous avons voulu, dans ce mémoire, développer des méthodes de prédiction à court terme des marchés financiers sur base d'une partie de l'information disponible. Les prédictions ont été réalisées sur le marché des changes et en particulier sur l'euro-dollar et le dollar-yen avec, comme données, les rendements sur 5 minutes des taux de changes et une variable « effet de surprise » des publications macroéconomiques. Quant aux méthodes de prédiction, nous avons choisi de travailler avec un système de fenêtres glissantes combiné aux capacités d'apprentissage des modèles ARMA, des régressions logistiques et des réseaux de neurones artificiels.

Dans un premier temps, nous avons expliqué le fonctionnement du marché des changes, les hypothèses liées à la prévisibilité des marchés financiers et le concept d'« effet de surprise » des publications macroéconomiques.

Ensuite, nous avons parcouru la théorie des différents modèles. On s'est d'abord concentré sur les modèles économétriques : le modèle « moyenne » et le modèle ARMA. Puis, nous avons vu la théorie des deux méthodes d'apprentissage automatique : la régression logistique et le perceptron multicouche.

Les deux prochaines étapes furent de préparer les séries de données utilisées pour l'apprentissage des modèles et de concevoir les processus de prédiction de la direction des taux de change sur base des différents modèles.

Une fois l'ensemble des prédictions calculées, nous avons, à partir des résultats, évalué la qualité des différents modèles et comparé leurs performances.

En analysant les performances des différents modèles sur l'ensemble des prédictions, nous avons été surpris par la différence entre les taux de bonne classification du *benchmark* et les taux des autres modèles. Les résultats ont montré qu'il existe de l'information dans les taux de change passés qui permet de mieux prédire la direction des taux de change qu'un système de décision aléatoire et nous avons constaté que les modèles ARMA, les régressions logistiques et les réseaux de neurones artificiels sont capables, du moins en partie, d'interpréter cette information.

Nous avons également vu que, pour les prédictions faites à la suite des annonces macroéconomiques, les taux de bonne classification des modèles ajustés avec la variable « effet de surprise » étaient nettement supérieurs à ceux des modèles qui ne prenaient pas en compte cette variable. Ces résultats sont venus conforter l'hypothèse selon laquelle les marchés financiers ne s'ajustent pas instantanément à la suite des publications macroéconomiques et qu'il était donc possible d'utiliser l'« effet de surprise » pour augmenter la précision des modèles prédictifs.

Alors que nous souhaitions connaître la meilleure méthode de classification, les performances des régressions logistiques et des perceptrons multicouches sont restées quasiment similaires, peu importe la configuration étudiée. Au terme de ce mémoire, si nous devions n'en choisir qu'une, nous opterions certainement pour la régression logistique qui demande beaucoup moins de temps de calcul que les réseaux de neurones artificiels.

En fin de mémoire, nous avons réalisé une simulation du profit qu'on pourrait dégager des prédictions des modèles lors d'annonces macroéconomiques dans des conditions

aussi proches que possible de la réalité. Les modèles avec la variable « effet de surprise » arrivent à dégager d'importants profits tant sur l'euro-dollar que sur le dollar-yen. Cependant, nous avons conseillé de considérer ces résultats avec précaution et nous avons identifié plusieurs facteurs qui pourraient venir réduire les profits lors de la mise en pratique.

Les résultats de ce mémoire nous incitent à encourager la poursuite des recherches sur ce sujet. Il serait notamment intéressant d'ajouter d'autres *inputs* aux modèles, tels que des données relatives aux flux d'ordres ou à la volatilité, ou de permettre aux modèles d'apprendre avec plus de rendements des valeurs passées si la puissance de calcul le permet. Nous conseillons également de travailler sur d'autres échelles de temps. Il paraît aussi judicieux d'essayer d'autres modèles pour prédire la direction des taux de change comme, par exemple, les machines à vecteurs de support.

# Bibliographie

[1] N. K. Ahmeda, A. F. Atiyab, N. Gayarc, and H. El-Shishinyd, "An Empirical Comparison of Machine Learning Models for Time Series Forecasting," *Econometric Reviews*, vol. 29, no. 5, pp. 594-621, 2010.

[2] E. Alpaydin, *Introduction to Machine Learning*, 2nd ed. Cambridge: The MIT Press, 2009.

[3] T.G. Andersen, T. Bollerslev, F. X. Diebold, and C. Vega, "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," vol. 93, no. 1, pp. 38-62, 2003.

[4] T. G. Andersen, T. Bollerslev, F. X. Diebold, and C. Vega, "Real-Time Price Discovery in Stock, Bond and Foreign Exchange Markets," *Journal of International Economics*, vol. 73, pp. 251-277, 2007.

[5] G. Athanasopoulos and R. J. Hyndman, *Forecasting: principles and practice*.: Otexts, 2013. [Online]. https://www.otexts.org/fpp/2/2

[6] Bank For International Settlements, "Triennial Central Bank Survey - Foreign Exchange turnover in April 2013: preliminary global results," 2013.

[7] A. Bellier and W. Sayeh, *Neural networks versus logistic regression: the best accuracy in*, University of Cergy-Pontoise, Ed., 2014.

[8] M. Boyd and I. Kaastra, "Designing a neural network for forecasting financial and economic time series," *Neurocomputing*, vol. 10, pp. 215-236, 1996.

[9] S. M. Burney, T. A. Jilani, and C. Ardil, "A Comparison of First and Second Order Training," *International Journal of Computer, Electronical Automation, Control and Information Engineering*, vol. 1, no. 1, 2007.

[10] C. Chesneau. (2017) Modèles de régression. [Online]. http://www.math.unicaen.fr/~chesneau/Reg-M2.pdf

[11] J. P. Davim, *Machining of Hard Materials*. London: Springer, 2011.

[12] Dukascopy. (2015a) F.A.Q. [Online]. https://www.dukascopy.com/swiss/french/about/faq/#a662

[13] Dukascopy. (2015b) Fee Schedule. [Online]. https://www.dukascopy.com/europe/english/about/fee-schedule/

[14] W. Enders, *Applied Econometric Time Series*. New York: Wiley, 2015.

[15] R. Finberg. (2015) With strong second half of the year, CLS Bank volumes grow in 2014. [Online]. http://www.financemagnates.com/institutional-forex/execution/strong-second-half-year-cls-bank-volumes-grow-2014/

[16] A. Fischer, "Deux méthodes d'apprentissage non supervisé : synthèse sur la méthode des centres mobiles et présentation des courbes principales," *Journal de la Société Franc¸aise de Statistique*, vol. 155, no. 2, pp. 2-35, 2014.

[17] A. Gramfort. (2014) (Quasi-) Newton Methods. [Online]. http://www.lix.polytechnique.fr/bigdata/mathbigdata/wp-content/uploads/2014/09/notes_quasi_newton.pdf

[18] P. Gupta, *Principles of Business - Financial Accounting*. Bloomington: AuthorHouse, 2012.

[19] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York: Springer Series in Statistics, 2009.

[20] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Hoboken : John Wiley & Sons, 2013.

[21] T. Imam, "Intelligent Computing and Foregin Echange Rate Prediction: What we know and we don't," *Progress in Intelligent Computing and Applications*, vol. 1, pp. 1-15, 2012.

[22] M. C. Jensen, "Some Anomalous Evidence Regarding Market Efficiency," *Journal Of Financial Economics*, vol. 6, no. 2/3, pp. 95-101, 1978.

[23] I. Kaastra and M. Boyd, "Desigining a neural network for forecasting financial and economic time series," *Neurocomputing*, vol. 10, no. 3, pp. 215-236, 1996.

[24] MathWorks. (2017) Rolling-window Analysis of Time-Series Models. [Online]. http://nl.mathworks.com/help/econ/rolling-window-estimation-of-state-space-models.html

[25] R. A. Meese and K. Rogoff, "Empire Exchange Rate Models Of The Seventies - Do they fit out of sample?," *Journal of International Economics*, vol. 14, pp. 3-24, 1983.

[26] M. Moffatt. (2003) A Beginner's Guide To Economic Indicators. [Online]. http://economics.about.com/cs/businesscycles/a/economic_ind.htm

[27] R. Nau, *Notes on the random walk model.*, 2014.

[28] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer, 1999.

[29] R.W. Parks and E. Zivot. (2006) Financial Market Efficiency and its Implications. [Online]. http://faculty.washington.edu/ezivot/econ422/Market%20Efficiency%20EZ.pdf

[30] L. Rouvière. (2017) Régression logistique avec R. [Online]. https://perso.univ-rennes2.fr/system/files/users/rouviere_l/poly_logistique_web.pdf

[31] J. Saettele and C. Vecchio. (2011) DailyFX. [Online]. https://www.dailyfx.com/forex/technical/article/forex_strategy_corner/2011/11/04/Trading_Nonfarm_Payrolls_Here_is_an_Alternative_EURUSD_Strategy.html

[32] Techopedia. (2010) Artificial Neural Network (ANN). [Online]. https://www.techopedia.com/definition/5967/artificial-neural-network-ann

[33] R. S. Tsay, *Analysis of Financial Time Series, Second Edition*. New York: Wiley, 2005.

[34] J. Wang and E. Zivot, *Modeling Financial Time Series With S-Plus*, 2nd ed. New York: Springer, 2006.

[35] Wikipedia. (2017) Akaike information criterion. [Online]. https://en.wikipedia.org/wiki/Akaike_information_criterion

[36] Wikipedia. (2017) R (langage). [Online]. https://fr.wikipedia.org/wiki/R_(langage)

[37] Wikipedia. (2017) Rétropropagation du gradient. [Online]. https://fr.wikipedia.org/wiki/R%C3%A9tropropagation_du_gradient

# Annexes

## a. Algorithme BFGS

Nommé d'après les 4 chercheurs qui l'ont découvert, l'algorithme de Broyden, Fletcher, Goldfarb et Shanno est une méthode d'optimisation numérique itérative de la famille des méthodes de quasi-Newton.

Typiquement, l'algorithme BFGS cherche à trouver, par itération, le vecteur $x^*$ qui minimise une fonction $f(x)$ avec $f$: $\mathbb{R}^n \rightarrow \mathbb{R}^n$ [17].

Etant donné le vecteur de départ $x_0$, l'algorithme BFGS consiste à itérer :

$$x_{k+1} = x_k + \alpha_k d_k$$

avec

$d_k = -\mathrm{H}_k \nabla f(x_k)$, la direction de recherche en $x_k$

$\mathrm{H}_k$, une matrice définie positive qui approxime le hessien en $x_k$

$\nabla f(x_k)$, le gradient en $x_k$

$\alpha_k$, la longueur de pas.

L'idée centrale des méthodes de quasi-Newton est de considérer que les deux vecteurs successifs $x_k$ et $x_{k+1}$, conjointement avec les gradients $\nabla f(x_k)$ et $\nabla f(x_{k+1})$, contiennent suffisamment d'informations sur la courbe, donc sur le hessien. On peut donc supposer comme valide l'approximation suivante :

$$\nabla f(x_{k+1}) - \nabla f(x_k) \approx \nabla^2 f(x_k)(x_{k+1} - x_k)$$

où $\nabla^2 f(x_k)$ est la matrice hessienne de la fonction en $x_k$.

Dès lors, l'objectif sera de trouver à chaque itération $\mathrm{H}_{k+1}$ qui satisfait l'équation :

$$y = \mathrm{H}_{k+1} s$$

où $y = \nabla f(x_{k+1}) - \nabla f(x_k)$ et $s = x_{k+1} - x_k$

L'équation ci-dessus est appelée équation sécante ou condition de quasi-Newton.

La formule de BFGS permet de mettre à jour $\mathrm{H}_k$, l'approximation de la matrice hessienne, tout en respectant l'équation sécante et en forçant $\mathrm{H}_k$ à rester définie, positive et symétrique si $\mathrm{H}_{k-1}$ l'est aussi. Ces conditions permettent de s'assurer que $d_k$ soit une direction de descente permettant ainsi de trouver $\alpha_k$ par recherche linéaire [9]. La formule de mise à jour donne :

$$\mathrm{H}_{k+1} = \mathrm{H}_k + \frac{yy^T}{y^T s} - \frac{\mathrm{H}_k s s^T \mathrm{H}_k}{s^T \mathrm{H}_k s}$$

La direction de recherche $d_k = -\mathrm{H}_k \nabla f(x_k)$ connue, il ne reste plus qu'à trouver la longueur de pas $\alpha_k$ grâce à une recherche linéaire. Généralement, un algorithme de recherche linéaire se passe en deux phases : une première phase où l'intervalle des longueurs de pas désirables est délimité et une seconde phase où l'algorithme cherche, parmi les candidats, la longueur de pas qui remplit certaines conditions [28].
Parmi les conditions les plus populaires, on retrouve **les conditions de Wolfe** qui sont souvent associées à l'algorithme BFGS. La première condition est appelée **condition d'Armijo** ou condition de décroissance linéaire :

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k \beta_1 \nabla f(x_k)^T d_k$$

avec $0 < \beta_1 < 1$.

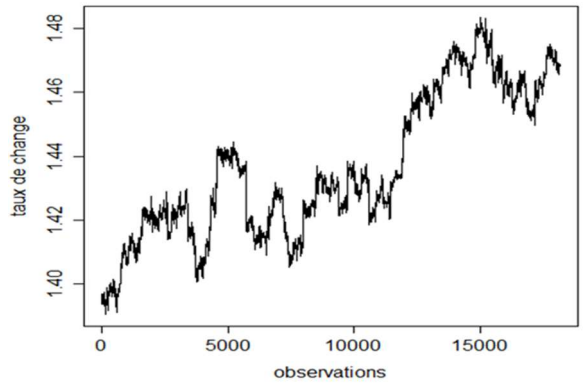
La deuxième condition, appelée condition de courbure, force le pas à être de longueur suffisante. Elle est représentée par l'inéquation suivante :

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geq \beta_2 \nabla f(x_k)^T d_k$$

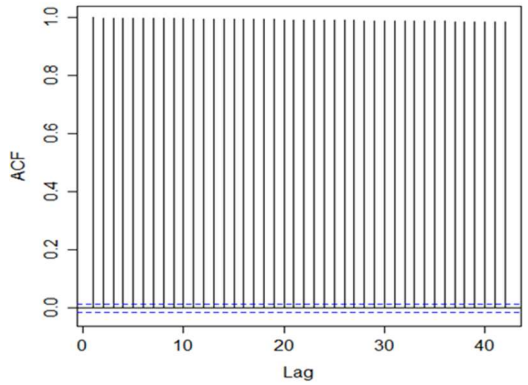
avec $0 < \beta_2 < 1$.

**b. Graphiques et corrélogrammes de sous-échantillons de la série de taux de change euro-dollar**

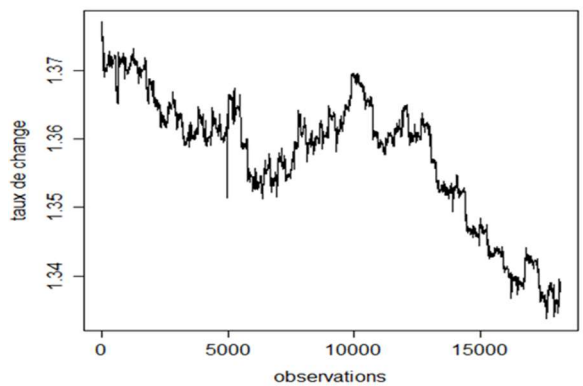
*Evolution d'un sous-échantillon de la série euro-dollar*



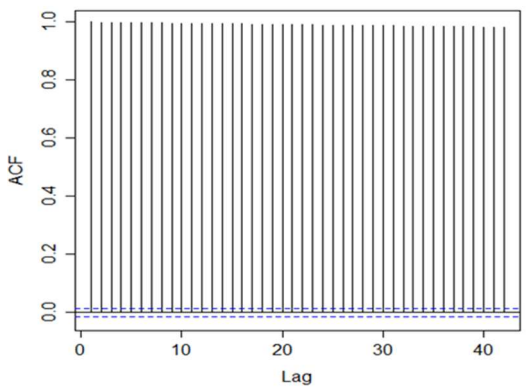

*ACF d'un sous-échantillon de la série euro-dollar*




*Evolution d'un sous-échantillon de la série euro-dollar*



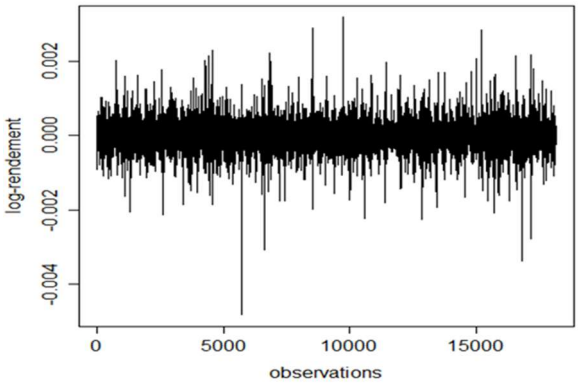
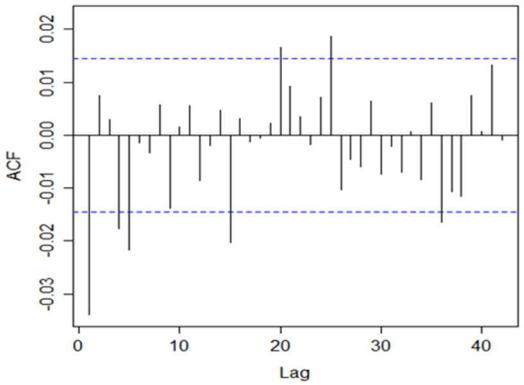
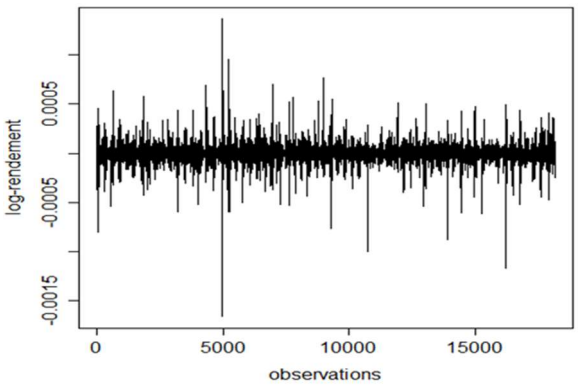
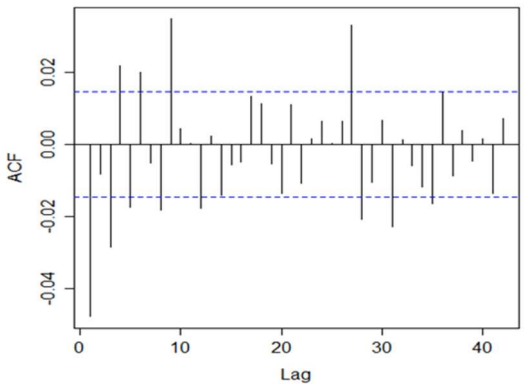

*ACF d'un sous-échantillon de la série euro-dollar*



c. Résultats du test de Dickey-Fuller sur des sous-échantillons de la série de taux de change euro-dollar

| | |
|---|---|
| Augmented Dickey-Fuller Test<br><br>data: trainPrice<br>Dickey-Fuller = -2.9346, Lag order = 3, **p-value = 0.1821**<br>alternative hypothesis: stationary | Augmented Dickey-Fuller Test<br><br>data: trainPrice<br>Dickey-Fuller = -2.3956, Lag order = 3, **p-value = 0.4107**<br>alternative hypothesis: stationary |

d. Graphiques et corrélogrammes de sous-échantillons de la série log-rendements de l'euro-dollar

| | |
|---|---|
| *Evolution d'un sous-échantillon de la série log-rendement* | *ACF d'un sous-échantillon de la série log-rendement* |
|  |  |
| *Evolution d'un sous-échantillon de la série log-rendement* | *ACF d'un sous-échantillon de la série log-rendement* |
|  |  |

e. Résultats du test de Dickey-Fuller sur des sous-échantillons de la série log-rendements de l'euro-dollar

| | |
|---|---|
| Augmented Dickey-Fuller Test<br><br>data: train<br>Dickey-Fuller = -68.709, Lag order = 3, **p-value = 0.01**<br>alternative hypothesis: stationary | Augmented Dickey-Fuller Test<br><br>data: train<br>Dickey-Fuller = -68.63, Lag order = 3, **p-value = 0.01**<br>alternative hypothesis: stationary |

f. Critère d'information d'Akaike

Le critère d'information d'Akaike est une mesure de la qualité relative de modèles statistiques pour un certain ensemble de données. Il présente une estimation relative de la perte d'information lorsqu'un certain modèle est utilisé pour représenter le processus qui génère les données [35].

Le critère d'information d'Akaike est calculé comme suit :

$$AIC = 2k - 2\ln(L)$$

où $k$ est le nombre de paramètres estimés et $L$ est la valeur maximale de la fonction de vraisemblance du modèle.

Etant donné un ensemble de modèles candidats, le modèle sélectionné sera celui dont l'AIC est le plus petit.

g. Répartition des combinaisons (p,q) finales pour les modèles ARMA 0 ≤p≤5 et ARMA 2≤p≤5

| Combinaisons (p,q) | ARMA 0≤p≤5 (EUR/USD) | ARMA 0≤p≤5 (USD/JPY) | ARMA 2≤p≤5 (EUR/USD) | ARMA 2≤p≤5 (USD/JPY) |
|---|---|---|---|---|
| (0,0) | 20 | 2 | 0 | 0 |
| (0,1) | 22 | 18 | 0 | 0 |
| (0,2) | 9 | 8 | 0 | 0 |
| (0,3) | 10 | 7 | 0 | 0 |
| (0,4) | 6 | 6 | 0 | 0 |
| (0,5) | 8 | 4 | 0 | 0 |
| (1,0) | 6 | 9 | 0 | 0 |
| (1,1) | 8 | 5 | 0 | 0 |
| (1,2) | 0 | 0 | 0 | 0 |
| (1,3) | 0 | 1 | 0 | 0 |
| (1,4) | 0 | 1 | 0 | 0 |
| (1,5) | 0 | 0 | 0 | 0 |
| (2,0) | 4 | 21 | 52 | 51 |
| (2,1) | 0 | 0 | 1 | 2 |
| (2,2) | 2 | 6 | 5 | 6 |
| (2,3) | 3 | 2 | 3 | 3 |
| (2,4) | 0 | 0 | 0 | 0 |
| (2,5) | 3 | 1 | 3 | 1 |
| (3,0) | 7 | 8 | 24 | 20 |
| (3,1) | 0 | 0 | 0 | 1 |
| (3,2) | 1 | 1 | 2 | 1 |
| (3,3) | 10 | 5 | 10 | 5 |
| (3,4) | 2 | 4 | 2 | 4 |
| (3,5) | 2 | 2 | 2 | 3 |
| (4,0) | 6 | 9 | 16 | 16 |
| (4,1) | 0 | 0 | 0 | 0 |
| (4,2) | 0 | 2 | 0 | 2 |
| (4,3) | 1 | 2 | 1 | 2 |
| (4,4) | 5 | 3 | 5 | 4 |
| (4,5) | 6 | 5 | 6 | 5 |
| (5,0) | 5 | 6 | 13 | 12 |
| (5,1) | 2 | 2 | 2 | 2 |
| (5,2) | 2 | 0 | 2 | 0 |
| (5,3) | 3 | 2 | 3 | 2 |
| (5,4) | 4 | 9 | 4 | 9 |
| (5,5) | 14 | 20 | 15 | 20 |

## h. Matrices de confusion des régressions logistiques et des perceptrons multi-couches par fenêtre glissante

| Régressions logistiques (EUR/USD) | | | | Réseaux de neurones artificiels (EUR/USD) | | | |
|---|---|---|---|---|---|---|---|
| Vraie + | | Vraie - | | Vraie + | | Vraie - | |
| Prédiction + | Prédiction - | Prédiction + | Prédiction - | Prédiction + | Prédiction - | Prédiction + | Prédiction - |
| 212 | 1473 | 159 | 1781 | 275 | 1410 | 250 | 1690 |
| 123 | 1641 | 111 | 1750 | 126 | 1638 | 111 | 1750 |
| 178 | 1595 | 112 | 1740 | 279 | 1494 | 184 | 1668 |
| 439 | 1297 | 394 | 1495 | 546 | 1190 | 509 | 1380 |
| 521 | 1191 | 480 | 1433 | 597 | 1115 | 569 | 1344 |
| 499 | 1272 | 450 | 1404 | 488 | 1283 | 475 | 1379 |
| 515 | 1260 | 434 | 1416 | 532 | 1243 | 463 | 1387 |
| 529 | 1129 | 559 | 1408 | 523 | 1135 | 533 | 1434 |
| 326 | 1465 | 260 | 1574 | 428 | 1363 | 365 | 1469 |
| 474 | 1334 | 396 | 1421 | 614 | 1194 | 534 | 1283 |
| 577 | 1171 | 515 | 1362 | 602 | 1146 | 538 | 1339 |
| 487 | 1249 | 474 | 1415 | 531 | 1205 | 507 | 1382 |
| 392 | 1362 | 413 | 1458 | 467 | 1287 | 484 | 1387 |
| 629 | 1143 | 588 | 1265 | 711 | 1061 | 673 | 1180 |
| 482 | 1239 | 486 | 1418 | 598 | 1123 | 629 | 1275 |
| 378 | 1389 | 366 | 1492 | 383 | 1384 | 368 | 1490 |
| 429 | 1354 | 462 | 1380 | 353 | 1430 | 379 | 1463 |
| 426 | 1326 | 396 | 1477 | 455 | 1297 | 421 | 1452 |
| 410 | 1368 | 352 | 1495 | 459 | 1319 | 394 | 1453 |
| 409 | 1357 | 337 | 1522 | 512 | 1254 | 461 | 1398 |
| 448 | 1340 | 415 | 1422 | 518 | 1270 | 489 | 1348 |
| 554 | 1201 | 544 | 1326 | 600 | 1155 | 628 | 1242 |
| 296 | 1527 | 262 | 1540 | 195 | 1628 | 170 | 1632 |
| 446 | 1368 | 424 | 1387 | 308 | 1506 | 323 | 1488 |
| 494 | 1315 | 484 | 1332 | 312 | 1497 | 288 | 1528 |
| 663 | 1092 | 659 | 1211 | 613 | 1142 | 597 | 1273 |
| 655 | 1074 | 746 | 1150 | 684 | 1045 | 756 | 1140 |
| 655 | 1092 | 654 | 1224 | 102 | 1645 | 95 | 1783 |
| 578 | 1204 | 556 | 1287 | 592 | 1190 | 566 | 1277 |
| 539 | 1248 | 568 | 1270 | 573 | 1214 | 597 | 1241 |
| 127 | 1626 | 141 | 1731 | 372 | 1381 | 406 | 1466 |
| 278 | 1510 | 235 | 1602 | 175 | 1613 | 140 | 1697 |
| 384 | 1442 | 325 | 1474 | 412 | 1414 | 341 | 1458 |
| 630 | 1195 | 568 | 1232 | 681 | 1144 | 642 | 1158 |
| 778 | 1037 | 710 | 1100 | 829 | 986 | 763 | 1047 |
| 801 | 972 | 760 | 1092 | 804 | 969 | 729 | 1123 |

| 805 | 943 | 824 | 1053 | 718 | 1030 | 678 | 1199 |
|---|---|---|---|---|---|---|---|
| 730 | 1101 | 647 | 1147 | 796 | 1035 | 706 | 1088 |
| 741 | 1047 | 670 | 1167 | 596 | 1192 | 553 | 1284 |
| 672 | 1117 | 585 | 1251 | 795 | 994 | 697 | 1139 |
| 602 | 1173 | 578 | 1272 | 738 | 1037 | 706 | 1144 |
| 655 | 1100 | 616 | 1254 | 772 | 983 | 772 | 1098 |
| 739 | 1049 | 686 | 1151 | 814 | 974 | 783 | 1054 |
| 514 | 1179 | 401 | 1531 | 606 | 1087 | 506 | 1426 |
| 365 | 1503 | 314 | 1443 | 536 | 1332 | 476 | 1281 |
| 500 | 1387 | 378 | 1360 | 732 | 1155 | 636 | 1102 |
| 814 | 1038 | 627 | 1146 | 1029 | 823 | 856 | 917 |
| 972 | 869 | 915 | 869 | 831 | 1010 | 781 | 1003 |
| 1113 | 750 | 962 | 800 | 1135 | 728 | 1010 | 752 |
| 1386 | 489 | 1259 | 491 | 1288 | 587 | 1129 | 621 |
| 1383 | 442 | 1309 | 491 | 1261 | 564 | 1190 | 610 |
| 1269 | 582 | 1198 | 576 | 1203 | 648 | 1142 | 632 |
| 1478 | 355 | 1400 | 392 | 1412 | 421 | 1309 | 483 |
| 1371 | 474 | 1271 | 509 | 1299 | 546 | 1223 | 557 |
| 1309 | 537 | 1175 | 604 | 1230 | 616 | 1066 | 713 |
| 1237 | 581 | 1158 | 649 | 1229 | 589 | 1132 | 675 |
| 1140 | 688 | 1063 | 734 | 1175 | 653 | 1114 | 683 |
| 1135 | 699 | 1041 | 750 | 1141 | 693 | 1061 | 730 |
| 1151 | 667 | 1058 | 749 | 1289 | 529 | 1188 | 619 |
| 1058 | 740 | 1055 | 772 | 1170 | 628 | 1157 | 670 |
| 1007 | 780 | 953 | 885 | 1054 | 733 | 1065 | 773 |
| 929 | 869 | 871 | 956 | 883 | 915 | 822 | 1005 |
| 854 | 988 | 787 | 996 | 925 | 917 | 817 | 966 |
| 857 | 970 | 796 | 1002 | 1023 | 804 | 973 | 825 |
| 895 | 937 | 800 | 993 | 951 | 881 | 831 | 962 |
| 991 | 896 | 859 | 879 | 965 | 922 | 844 | 894 |
| 1271 | 550 | 1184 | 620 | 1231 | 590 | 1150 | 654 |
| 1426 | 394 | 1302 | 503 | 1440 | 380 | 1331 | 474 |
| 1189 | 677 | 1083 | 676 | 1130 | 736 | 1035 | 724 |
| 1303 | 513 | 1236 | 573 | 1244 | 572 | 1208 | 601 |
| 1280 | 523 | 1229 | 593 | 1173 | 630 | 1147 | 675 |
| 1106 | 671 | 1081 | 767 | 1070 | 707 | 1053 | 795 |
| 1033 | 754 | 966 | 872 | 1069 | 718 | 1040 | 798 |
| 923 | 838 | 880 | 984 | 1097 | 664 | 1061 | 803 |
| 617 | 1220 | 495 | 1293 | 1024 | 813 | 931 | 857 |
| 709 | 1118 | 616 | 1182 | 974 | 853 | 848 | 950 |
| 787 | 992 | 703 | 1143 | 1031 | 748 | 957 | 889 |
| 737 | 1086 | 658 | 1144 | 920 | 903 | 821 | 981 |
| 865 | 923 | 756 | 1081 | 964 | 824 | 900 | 937 |
| 965 | 834 | 849 | 977 | 1025 | 774 | 948 | 878 |
| 792 | 997 | 661 | 1175 | 1022 | 767 | 887 | 949 |

| 667 | 1092 | 656 | 1210 | 895 | 864 | 895 | 971 |
|---|---|---|---|---|---|---|---|
| 594 | 1218 | 540 | 1273 | 814 | 998 | 736 | 1077 |
| 623 | 1209 | 521 | 1272 | 678 | 1154 | 560 | 1233 |
| 758 | 1027 | 711 | 1129 | 773 | 1012 | 748 | 1092 |
| 786 | 980 | 710 | 1149 | 848 | 918 | 771 | 1088 |
| 742 | 1056 | 666 | 1161 | 784 | 1014 | 694 | 1133 |
| 798 | 994 | 715 | 1118 | 743 | 1049 | 703 | 1130 |
| 711 | 1088 | 669 | 1157 | 823 | 976 | 789 | 1037 |
| 711 | 1070 | 650 | 1194 | 857 | 924 | 794 | 1050 |
| 601 | 1198 | 565 | 1261 | 780 | 1019 | 706 | 1120 |
| 722 | 1114 | 622 | 1167 | 723 | 1113 | 660 | 1129 |
| 818 | 972 | 755 | 1080 | 862 | 928 | 789 | 1046 |
| 775 | 990 | 719 | 1141 | 720 | 1045 | 696 | 1164 |
| 652 | 1154 | 544 | 1275 | 780 | 1026 | 661 | 1158 |
| 731 | 1082 | 623 | 1189 | 1014 | 799 | 897 | 915 |
| 799 | 1006 | 682 | 1138 | 978 | 827 | 886 | 934 |
| 802 | 1004 | 691 | 1128 | 888 | 918 | 800 | 1019 |
| 812 | 942 | 682 | 1189 | 970 | 784 | 866 | 1005 |
| 812 | 1010 | 695 | 1108 | 923 | 899 | 795 | 1008 |
| 806 | 974 | 718 | 1127 | 715 | 1065 | 611 | 1234 |
| 752 | 1039 | 642 | 1192 | 722 | 1069 | 607 | 1227 |
| 687 | 1066 | 694 | 1178 | 747 | 1006 | 762 | 1110 |
| 685 | 1076 | 664 | 1200 | 759 | 1002 | 771 | 1093 |
| 640 | 1170 | 602 | 1213 | 742 | 1068 | 714 | 1101 |
| 581 | 1182 | 483 | 1379 | 703 | 1060 | 605 | 1257 |
| 502 | 1270 | 443 | 1410 | 671 | 1101 | 619 | 1234 |
| 433 | 1374 | 344 | 1474 | 603 | 1204 | 528 | 1290 |
| 461 | 1314 | 390 | 1460 | 530 | 1245 | 470 | 1380 |
| 559 | 1250 | 470 | 1346 | 682 | 1127 | 599 | 1217 |
| 632 | 1116 | 584 | 1293 | 705 | 1043 | 632 | 1245 |
| 616 | 1142 | 535 | 1332 | 743 | 1015 | 653 | 1214 |
| 607 | 1132 | 527 | 1359 | 692 | 1047 | 620 | 1266 |
| 472 | 1337 | 366 | 1450 | 617 | 1192 | 509 | 1307 |
| 527 | 1266 | 462 | 1370 | 520 | 1273 | 453 | 1379 |
| 513 | 1329 | 417 | 1366 | 558 | 1284 | 434 | 1349 |
| 779 | 1004 | 677 | 1165 | 718 | 1065 | 616 | 1226 |
| 810 | 973 | 699 | 1143 | 853 | 930 | 710 | 1132 |
| 929 | 861 | 800 | 1035 | 913 | 877 | 787 | 1048 |
| 820 | 947 | 724 | 1134 | 897 | 870 | 795 | 1063 |
| 705 | 1079 | 676 | 1165 | 878 | 906 | 851 | 990 |
| 635 | 1163 | 535 | 1292 | 817 | 981 | 708 | 1119 |
| 621 | 1153 | 523 | 1328 | 838 | 936 | 748 | 1103 |
| 630 | 1147 | 560 | 1288 | 861 | 916 | 730 | 1118 |
| 620 | 1163 | 492 | 1350 | 845 | 938 | 694 | 1148 |
| 641 | 1164 | 537 | 1283 | 826 | 979 | 729 | 1091 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 654 | 1119 | 614 | 1238 | 741 | 1032 | 688 | 1164 |
| 578 | 1160 | 514 | 1373 | 724 | 1014 | 658 | 1229 |
| 494 | 1269 | 407 | 1455 | 686 | 1077 | 607 | 1255 |
| 428 | 1270 | 403 | 1524 | 622 | 1076 | 625 | 1302 |
| 342 | 1410 | 282 | 1591 | 556 | 1196 | 469 | 1404 |
| 337 | 1428 | 238 | 1622 | 598 | 1167 | 481 | 1379 |
| 389 | 1368 | 322 | 1546 | 601 | 1156 | 511 | 1357 |
| 535 | 1219 | 444 | 1427 | 698 | 1056 | 665 | 1206 |
| 398 | 1393 | 333 | 1501 | 702 | 1089 | 564 | 1270 |
| 539 | 1244 | 415 | 1427 | 729 | 1054 | 614 | 1228 |
| 633 | 1099 | 596 | 1297 | 764 | 968 | 717 | 1176 |
| 542 | 1251 | 456 | 1376 | 750 | 1043 | 645 | 1187 |
| 599 | 1175 | 472 | 1379 | 785 | 989 | 642 | 1209 |
| 611 | 1159 | 489 | 1366 | 776 | 994 | 634 | 1221 |
| 571 | 1196 | 519 | 1339 | 735 | 1032 | 691 | 1167 |
| 536 | 1228 | 444 | 1417 | 816 | 948 | 697 | 1164 |
| 479 | 1225 | 433 | 1488 | 749 | 955 | 749 | 1172 |
| 343 | 1388 | 310 | 1584 | 665 | 1066 | 633 | 1261 |
| 273 | 1454 | 250 | 1648 | 624 | 1103 | 625 | 1273 |
| 215 | 1542 | 197 | 1671 | 514 | 1243 | 461 | 1407 |
| 235 | 1495 | 190 | 1705 | 424 | 1306 | 369 | 1526 |
| 205 | 1550 | 167 | 1703 | 411 | 1344 | 368 | 1502 |
| 221 | 1453 | 193 | 1758 | 495 | 1179 | 521 | 1430 |
| 274 | 1473 | 217 | 1661 | 502 | 1245 | 437 | 1441 |
| 382 | 1329 | 349 | 1565 | 559 | 1152 | 538 | 1376 |
| 426 | 1309 | 402 | 1488 | 613 | 1122 | 616 | 1274 |
| 481 | 1300 | 435 | 1409 | 567 | 1214 | 545 | 1299 |
| 285 | 1459 | 224 | 1657 | 491 | 1253 | 463 | 1418 |
| 443 | 1285 | 416 | 1481 | 551 | 1177 | 513 | 1384 |
| 340 | 1430 | 293 | 1562 | 467 | 1303 | 411 | 1444 |
| 415 | 1349 | 358 | 1503 | 558 | 1206 | 498 | 1363 |
| 506 | 1243 | 448 | 1428 | 697 | 1052 | 646 | 1230 |
| 595 | 1197 | 514 | 1319 | 762 | 1030 | 684 | 1149 |
| 459 | 1271 | 439 | 1456 | 644 | 1086 | 605 | 1290 |
| 418 | 1325 | 399 | 1483 | 728 | 1015 | 685 | 1197 |
| 486 | 1329 | 453 | 1357 | 747 | 1068 | 701 | 1109 |
| 405 | 1359 | 381 | 1480 | 681 | 1083 | 680 | 1181 |
| 435 | 1359 | 408 | 1423 | 776 | 1018 | 745 | 1086 |
| 387 | 1394 | 369 | 1475 | 738 | 1043 | 747 | 1097 |
| 438 | 1368 | 390 | 1429 | 836 | 970 | 785 | 1034 |
| 611 | 1187 | 571 | 1256 | 956 | 842 | 918 | 909 |
| 665 | 1154 | 620 | 1186 | 785 | 1034 | 701 | 1105 |
| 720 | 1052 | 691 | 1162 | 1051 | 721 | 1029 | 824 |
| 730 | 1075 | 657 | 1163 | 787 | 1018 | 762 | 1058 |
| 816 | 958 | 765 | 1086 | 956 | 818 | 905 | 946 |

| Régressions logistiques (USD/JPY) | | | | Réseaux de neurones artificiels (USD/JPY) | | | |
|---|---|---|---|---|---|---|---|
| Vraie + | | Vraie - | | Vraie + | | Vraie - | |
| Prédiction + | Prédiction - | Prédiction + | Prédiction - | Prédiction + | Prédiction - | Prédiction + | Prédiction - |
| 184 | 1473 | 166 | 1802 | 216 | 1441 | 193 | 1775 |
| 264 | 1379 | 246 | 1736 | 295 | 1348 | 268 | 1714 |
| 242 | 1421 | 265 | 1697 | 343 | 1320 | 370 | 1592 |
| 369 | 1309 | 386 | 1561 | 487 | 1191 | 529 | 1418 |
| 359 | 1375 | 314 | 1577 | 462 | 1272 | 442 | 1449 |
| 207 | 1485 | 173 | 1760 | 265 | 1427 | 239 | 1694 |
| 143 | 1584 | 110 | 1788 | 132 | 1595 | 108 | 1790 |
| 100 | 1607 | 86 | 1832 | 138 | 1569 | 123 | 1795 |
| 225 | 1502 | 185 | 1713 | 253 | 1474 | 221 | 1677 |
| 430 | 1323 | 412 | 1460 | 613 | 1140 | 579 | 1293 |
| 417 | 1331 | 375 | 1502 | 484 | 1264 | 450 | 1427 |
| 313 | 1396 | 275 | 1641 | 360 | 1349 | 334 | 1582 |
| 377 | 1319 | 374 | 1555 | 368 | 1328 | 367 | 1562 |
| 462 | 1273 | 410 | 1480 | 448 | 1287 | 407 | 1483 |
| 331 | 1405 | 252 | 1637 | 315 | 1421 | 238 | 1651 |
| 346 | 1342 | 358 | 1579 | 391 | 1297 | 402 | 1535 |
| 446 | 1280 | 484 | 1415 | 514 | 1212 | 589 | 1310 |
| 222 | 1566 | 170 | 1667 | 382 | 1406 | 335 | 1502 |
| 248 | 1472 | 223 | 1682 | 299 | 1421 | 277 | 1628 |
| 144 | 1589 | 137 | 1755 | 233 | 1500 | 222 | 1670 |
| 129 | 1541 | 124 | 1831 | 242 | 1428 | 211 | 1744 |
| 67 | 1654 | 60 | 1844 | 131 | 1590 | 114 | 1790 |
| 228 | 1487 | 220 | 1690 | 286 | 1429 | 271 | 1639 |
| 192 | 1502 | 176 | 1755 | 257 | 1437 | 235 | 1696 |
| 141 | 1568 | 121 | 1795 | 283 | 1426 | 246 | 1670 |
| 218 | 1463 | 226 | 1718 | 301 | 1380 | 298 | 1646 |
| 356 | 1390 | 320 | 1559 | 471 | 1275 | 441 | 1438 |
| 429 | 1363 | 403 | 1430 | 586 | 1206 | 556 | 1277 |
| 562 | 1230 | 491 | 1342 | 734 | 1058 | 655 | 1178 |
| 634 | 1162 | 599 | 1230 | 666 | 1130 | 665 | 1164 |
| 536 | 1232 | 445 | 1412 | 536 | 1232 | 458 | 1399 |
| 506 | 1251 | 477 | 1391 | 461 | 1296 | 430 | 1438 |
| 491 | 1261 | 486 | 1387 | 408 | 1344 | 396 | 1477 |
| 389 | 1342 | 305 | 1589 | 338 | 1393 | 281 | 1613 |
| 382 | 1388 | 312 | 1543 | 478 | 1292 | 436 | 1419 |
| 386 | 1370 | 308 | 1561 | 404 | 1352 | 299 | 1570 |
| 450 | 1356 | 429 | 1390 | 513 | 1293 | 490 | 1329 |
| 533 | 1272 | 470 | 1350 | 587 | 1218 | 470 | 1350 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 608 | 1157 | 551 | 1309 | 530 | 1235 | 466 | 1394 |
| 625 | 1099 | 633 | 1268 | 549 | 1175 | 537 | 1364 |
| 590 | 1147 | 564 | 1324 | 455 | 1282 | 447 | 1441 |
| 468 | 1268 | 464 | 1425 | 377 | 1359 | 381 | 1508 |
| 437 | 1287 | 358 | 1543 | 350 | 1374 | 299 | 1602 |
| 349 | 1375 | 309 | 1592 | 354 | 1370 | 307 | 1594 |
| 303 | 1458 | 282 | 1582 | 373 | 1388 | 331 | 1533 |
| 314 | 1376 | 290 | 1645 | 345 | 1345 | 291 | 1644 |
| 295 | 1418 | 228 | 1684 | 404 | 1309 | 390 | 1522 |
| 299 | 1417 | 214 | 1695 | 442 | 1274 | 366 | 1543 |
| 374 | 1317 | 395 | 1539 | 498 | 1193 | 530 | 1404 |
| 305 | 1476 | 251 | 1593 | 474 | 1307 | 383 | 1461 |
| 300 | 1436 | 279 | 1610 | 513 | 1223 | 502 | 1387 |
| 324 | 1329 | 288 | 1684 | 481 | 1172 | 512 | 1460 |
| 243 | 1513 | 216 | 1653 | 420 | 1336 | 364 | 1505 |
| 126 | 1515 | 115 | 1869 | 257 | 1384 | 261 | 1723 |
| 232 | 1520 | 203 | 1670 | 593 | 1159 | 536 | 1337 |
| 198 | 1522 | 154 | 1751 | 408 | 1312 | 353 | 1552 |
| 122 | 1596 | 118 | 1789 | 251 | 1467 | 268 | 1639 |
| 172 | 1557 | 126 | 1770 | 496 | 1233 | 468 | 1428 |
| 192 | 1555 | 159 | 1719 | 464 | 1283 | 427 | 1451 |
| 267 | 1445 | 246 | 1667 | 475 | 1237 | 423 | 1490 |
| 347 | 1358 | 277 | 1643 | 494 | 1211 | 423 | 1497 |
| 392 | 1409 | 361 | 1463 | 494 | 1307 | 469 | 1355 |
| 292 | 1462 | 246 | 1625 | 392 | 1362 | 350 | 1521 |
| 268 | 1421 | 253 | 1683 | 294 | 1395 | 264 | 1672 |
| 249 | 1454 | 244 | 1678 | 333 | 1370 | 340 | 1582 |
| 205 | 1494 | 206 | 1720 | 325 | 1374 | 338 | 1588 |
| 184 | 1519 | 157 | 1765 | 239 | 1464 | 188 | 1734 |
| 157 | 1537 | 136 | 1795 | 232 | 1462 | 197 | 1734 |
| 158 | 1516 | 139 | 1812 | 227 | 1447 | 216 | 1735 |
| 122 | 1530 | 90 | 1883 | 226 | 1426 | 181 | 1792 |
| 175 | 1516 | 151 | 1783 | 296 | 1395 | 253 | 1681 |
| 249 | 1560 | 193 | 1623 | 353 | 1456 | 318 | 1498 |
| 227 | 1556 | 225 | 1617 | 331 | 1452 | 324 | 1518 |
| 321 | 1490 | 277 | 1537 | 459 | 1352 | 383 | 1431 |
| 426 | 1307 | 411 | 1481 | 512 | 1221 | 504 | 1388 |
| 544 | 1250 | 413 | 1418 | 680 | 1114 | 548 | 1283 |
| 677 | 1119 | 543 | 1286 | 739 | 1057 | 612 | 1217 |
| 646 | 1101 | 628 | 1250 | 676 | 1071 | 669 | 1209 |
| 654 | 1142 | 616 | 1213 | 578 | 1218 | 526 | 1303 |
| 382 | 1465 | 317 | 1461 | 542 | 1305 | 411 | 1367 |
| 647 | 1146 | 564 | 1268 | 953 | 840 | 871 | 961 |
| 604 | 1150 | 574 | 1297 | 566 | 1188 | 532 | 1339 |
| 381 | 1363 | 343 | 1538 | 479 | 1265 | 432 | 1449 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 363 | 1417 | 322 | 1523 | 468 | 1312 | 405 | 1440 |
| 484 | 1270 | 488 | 1383 | 417 | 1337 | 419 | 1452 |
| 426 | 1308 | 385 | 1506 | 586 | 1148 | 562 | 1329 |
| 525 | 1223 | 498 | 1379 | 648 | 1100 | 604 | 1273 |
| 324 | 1415 | 284 | 1602 | 428 | 1311 | 428 | 1458 |
| 313 | 1441 | 263 | 1608 | 383 | 1371 | 357 | 1514 |
| 284 | 1484 | 246 | 1611 | 436 | 1332 | 375 | 1482 |
| 254 | 1509 | 246 | 1616 | 348 | 1415 | 344 | 1518 |
| 322 | 1426 | 229 | 1648 | 565 | 1183 | 421 | 1456 |
| 204 | 1518 | 193 | 1710 | 586 | 1136 | 541 | 1362 |
| 220 | 1563 | 209 | 1633 | 567 | 1216 | 529 | 1313 |
| 255 | 1479 | 160 | 1731 | 456 | 1278 | 386 | 1505 |
| 206 | 1502 | 142 | 1775 | 614 | 1094 | 548 | 1369 |
| 231 | 1473 | 207 | 1714 | 538 | 1166 | 525 | 1396 |
| 336 | 1449 | 312 | 1528 | 514 | 1271 | 455 | 1385 |
| 523 | 1285 | 432 | 1385 | 721 | 1087 | 660 | 1157 |
| 565 | 1277 | 479 | 1304 | 618 | 1224 | 512 | 1271 |
| 626 | 1172 | 574 | 1253 | 683 | 1115 | 620 | 1207 |
| 654 | 1144 | 532 | 1295 | 506 | 1292 | 374 | 1453 |
| 869 | 889 | 773 | 1094 | 604 | 1154 | 521 | 1346 |
| 811 | 948 | 698 | 1168 | 696 | 1063 | 603 | 1263 |
| 745 | 1057 | 614 | 1209 | 773 | 1029 | 656 | 1167 |
| 693 | 1067 | 604 | 1261 | 730 | 1030 | 650 | 1215 |
| 687 | 1036 | 610 | 1292 | 626 | 1097 | 549 | 1353 |
| 580 | 1160 | 486 | 1399 | 765 | 975 | 672 | 1213 |
| 501 | 1236 | 419 | 1469 | 680 | 1057 | 645 | 1243 |
| 435 | 1317 | 402 | 1471 | 460 | 1292 | 407 | 1466 |
| 395 | 1284 | 333 | 1613 | 539 | 1140 | 534 | 1412 |
| 366 | 1405 | 317 | 1537 | 532 | 1239 | 491 | 1363 |
| 405 | 1363 | 321 | 1536 | 564 | 1204 | 496 | 1361 |
| 495 | 1311 | 434 | 1385 | 568 | 1238 | 520 | 1299 |
| 474 | 1308 | 477 | 1366 | 552 | 1230 | 568 | 1275 |
| 530 | 1307 | 482 | 1306 | 525 | 1312 | 457 | 1331 |
| 781 | 1093 | 674 | 1077 | 774 | 1100 | 641 | 1110 |
| 939 | 898 | 839 | 949 | 875 | 962 | 777 | 1011 |
| 1099 | 695 | 997 | 834 | 993 | 801 | 875 | 956 |
| 1103 | 763 | 904 | 855 | 975 | 891 | 791 | 968 |
| 1200 | 613 | 1112 | 700 | 1118 | 695 | 1048 | 764 |
| 1089 | 737 | 1005 | 794 | 1036 | 790 | 968 | 831 |
| 1144 | 647 | 1048 | 786 | 966 | 825 | 825 | 1009 |
| 1017 | 777 | 906 | 925 | 791 | 1003 | 667 | 1164 |
| 982 | 785 | 933 | 925 | 945 | 822 | 918 | 940 |
| 844 | 1003 | 722 | 1056 | 856 | 991 | 745 | 1033 |
| 841 | 975 | 698 | 1111 | 797 | 1019 | 660 | 1149 |
| 852 | 966 | 739 | 1068 | 890 | 928 | 759 | 1048 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 854 | 908 | 865 | 998 | 878 | 884 | 881 | 982 |
| 753 | 1054 | 679 | 1139 | 962 | 845 | 875 | 943 |
| 926 | 889 | 819 | 991 | 988 | 827 | 888 | 922 |
| 894 | 890 | 777 | 1064 | 962 | 822 | 867 | 974 |
| 760 | 971 | 735 | 1159 | 774 | 957 | 724 | 1170 |
| 566 | 1235 | 453 | 1371 | 625 | 1176 | 513 | 1311 |
| 717 | 1083 | 599 | 1226 | 696 | 1104 | 562 | 1263 |
| 753 | 1036 | 663 | 1173 | 816 | 973 | 699 | 1137 |
| 581 | 1241 | 500 | 1303 | 800 | 1022 | 670 | 1133 |
| 782 | 1029 | 681 | 1133 | 964 | 847 | 877 | 937 |
| 873 | 921 | 813 | 1018 | 954 | 840 | 895 | 936 |
| 856 | 913 | 755 | 1101 | 967 | 802 | 897 | 959 |
| 844 | 941 | 705 | 1135 | 1019 | 766 | 881 | 959 |
| 768 | 997 | 639 | 1221 | 1029 | 736 | 927 | 933 |
| 623 | 1119 | 590 | 1293 | 785 | 957 | 778 | 1105 |
| 504 | 1230 | 450 | 1441 | 725 | 1009 | 692 | 1199 |
| 490 | 1282 | 426 | 1427 | 712 | 1060 | 624 | 1229 |
| 484 | 1258 | 433 | 1450 | 640 | 1102 | 585 | 1298 |
| 380 | 1325 | 336 | 1584 | 502 | 1203 | 485 | 1435 |
| 246 | 1435 | 204 | 1740 | 427 | 1254 | 378 | 1566 |
| 328 | 1420 | 290 | 1587 | 468 | 1280 | 430 | 1447 |
| 351 | 1464 | 248 | 1562 | 546 | 1269 | 428 | 1382 |
| 472 | 1307 | 408 | 1438 | 600 | 1179 | 523 | 1323 |
| 566 | 1222 | 521 | 1316 | 648 | 1140 | 622 | 1215 |
| 756 | 995 | 747 | 1127 | 829 | 922 | 823 | 1051 |
| 655 | 1180 | 526 | 1264 | 804 | 1031 | 692 | 1098 |
| 811 | 1076 | 645 | 1093 | 760 | 1127 | 603 | 1135 |
| 908 | 938 | 817 | 962 | 779 | 1067 | 694 | 1085 |
| 1047 | 746 | 1020 | 812 | 888 | 905 | 841 | 991 |
| 1055 | 724 | 1007 | 839 | 849 | 930 | 816 | 1030 |
| 1089 | 687 | 1045 | 804 | 945 | 831 | 900 | 949 |
| 1006 | 792 | 931 | 896 | 924 | 874 | 843 | 984 |
| 735 | 1018 | 642 | 1230 | 713 | 1040 | 602 | 1270 |
| 596 | 1123 | 532 | 1374 | 653 | 1066 | 608 | 1298 |
| 441 | 1303 | 430 | 1451 | 500 | 1244 | 505 | 1376 |
| 435 | 1333 | 420 | 1437 | 575 | 1193 | 553 | 1304 |
| 315 | 1468 | 263 | 1579 | 407 | 1376 | 338 | 1504 |
| 424 | 1413 | 322 | 1466 | 582 | 1255 | 486 | 1302 |
| 491 | 1308 | 447 | 1379 | 578 | 1221 | 537 | 1289 |
| 638 | 1158 | 559 | 1270 | 745 | 1051 | 682 | 1147 |
| 687 | 1128 | 591 | 1219 | 717 | 1098 | 664 | 1146 |
| 836 | 927 | 799 | 1063 | 651 | 1112 | 594 | 1268 |
| 848 | 940 | 798 | 1039 | 846 | 942 | 788 | 1049 |

i. Valeurs des coefficients de la variable « effet de surprise » obtenus après ajustement des modèles ARMA et des régressions logistiques de chaque fenêtre glissante

| EUR/USD | | USD/JPY | |
|---|---|---|---|
| Modèles ARMA | Régressions Logistiques | Modèles ARMA | Régressions Logistiques |
| -0,00042902 | -0,69910544 | 0,00066757 | 1,39934447 |
| -0,00040056 | -0,44936033 | 0,00051896 | 1,06330597 |
| -0,00036564 | -0,50719311 | 0,00038453 | 1,12219994 |
| -0,00040795 | -0,65580668 | 0,0004416 | 1,38301718 |
| -0,00031009 | -0,46136952 | 0,00035134 | 1,00613342 |
| -0,00025925 | -0,39190745 | 0,00028832 | 0,9347888 |
| -0,00043961 | -0,63783784 | 0,00064828 | 1,16791642 |
| -0,00042647 | -0,40786503 | 0,00071515 | 0,85392651 |
| -0,00056231 | -0,4503215 | 0,00083844 | 0,59402348 |
| -0,00050803 | -0,64962901 | 0,00070332 | 0,66201547 |
| -0,00054415 | -0,92507715 | 0,00105954 | 0,92970782 |
| -0,00040882 | -0,96469105 | 0,00075026 | 0,80697066 |
| -0,00043946 | -1,14120813 | 0,00069753 | 0,80451401 |
| -0,00044818 | -1,33543502 | 0,00084599 | 0,79700129 |
| -0,00051196 | -1,40900194 | 0,00107024 | 0,87537882 |
| -0,00072608 | -1,46588075 | 0,00119839 | 0,7584664 |
| -0,00079617 | -1,36273948 | 0,00139905 | 0,93872206 |
| -0,00068769 | -0,95251807 | 0,00170086 | 1,09428007 |
| -0,00072876 | -1,14042574 | 0,00203724 | 1,52202458 |
| -0,00067403 | -0,80305021 | 0,00192288 | 1,36839682 |
| -0,0005839 | -0,44905233 | 0,00169805 | 1,16818038 |
| -0,00065334 | -0,57102153 | 0,00183201 | 1,25985512 |
| -0,0005469 | -0,6884352 | 0,00142446 | 1,08002198 |
| -0,00033483 | -0,37557748 | 0,00073407 | 0,74065119 |
| -0,00044288 | -0,78417977 | 0,00087529 | 0,99903837 |
| -0,0005134 | -1,13640774 | 0,00098369 | 1,60998702 |
| -0,0004842 | -1,14186537 | 0,00077377 | 1,92282631 |
| -0,00065697 | -1,22112425 | 0,00106926 | 1,69508846 |
| -0,00061489 | -1,13002627 | 0,00115181 | 1,83802278 |
| -0,00039338 | -0,61246728 | 0,00106974 | 1,23554082 |
| -1,17E-05 | -0,23870653 | 0,0009049 | 0,77507538 |
| 0,00030114 | 0,18623633 | 0,00093868 | 0,51734595 |
| 0,0002729 | 0,21746047 | 0,00094358 | 0,73713848 |
| 0,00024828 | 0,4137496 | 0,00070689 | 0,5769438 |
| 0,00025037 | 0,20509547 | 0,00082207 | 0,6882844 |
| 0,00031834 | 0,29010308 | 0,00070693 | 0,76614744 |
| 0,00017819 | -0,0006187 | 0,00039005 | 0,72740483 |

| | | | |
|---|---|---|---|
| 0,00060757 | 0,54834735 | 0,00035245 | 0,5374184 |
| 0,00093034 | 0,50955571 | -7,43E-05 | 0,6494436 |
| 0,00107204 | 0,97887989 | -0,00023031 | 0,78841723 |
| 0,00064072 | 0,45757943 | -9,30E-05 | 0,7926487 |
| 0,00089768 | 0,77156858 | 7,51E-05 | 0,79395436 |
| 0,00089201 | 0,86370308 | 0,00038721 | 1,0498243 |
| 0,00057912 | 0,61920614 | 0,00077331 | 1,12220769 |
| 0,00037624 | 0,53685806 | 0,00081189 | 1,15827839 |
| 0,0005925 | 0,82384349 | 0,00076297 | 1,19190408 |
| 0,00036091 | 0,65434425 | 0,00105424 | 1,3975 |
| 5,94E-05 | 0,2321988 | 0,00078984 | 1,08013728 |
| 0,00011761 | 0,43979478 | 0,00106532 | 1,33096415 |
| 0,00041375 | 0,67624296 | 0,00095399 | 0,86704913 |
| 0,00037173 | 0,59056386 | 0,00098327 | 0,93236896 |
| 0,00059579 | 0,93892259 | 0,00109412 | 1,1278609 |
| 0,00051949 | 0,57448654 | 0,00106515 | 1,42041696 |
| 0,00037134 | 0,35752222 | 0,00124848 | 1,15743184 |
| 4,22E-05 | -0,10540662 | 0,00143623 | 1,91736452 |
| -2,38E-05 | 0,07792575 | 0,00163494 | 1,7523511 |
| -0,00052707 | -0,33285472 | 0,00153359 | 1,48032751 |
| -0,00035034 | -0,06133675 | 0,00178519 | 1,89285788 |
| -0,00035217 | -0,11236822 | 0,00185088 | 2,39718601 |
| -0,00028761 | -0,26542418 | 0,00168413 | 2,03489251 |
| -0,00025675 | -0,39140426 | 0,00141619 | 2,11004286 |
| -7,79E-05 | -0,43540438 | 0,00112739 | 1,78333225 |
| -0,00027955 | -0,26710844 | 0,0009817 | 1,50670351 |
| 0,00016356 | 0,12480122 | 0,00088999 | 1,25368699 |
| 0,00036411 | 0,63472497 | 0,00107406 | 1,51585561 |
| 0,00025521 | 0,42936556 | 0,00092352 | 1,16710946 |
| 0,00019501 | 0,36433272 | 0,0011202 | 1,54673206 |
| 7,04E-05 | -0,14983374 | 0,00143202 | 1,58430093 |
| -0,00021361 | -0,4206223 | 0,00136748 | 1,89453522 |
| -0,00039772 | -0,87268051 | 0,00123225 | 1,90158343 |
| -0,00059602 | -0,92492326 | 0,0015231 | 2,61925207 |
| -0,0005275 | -0,6924958 | 0,00124042 | 1,43768676 |
| -0,00046056 | -0,48654974 | 0,00100863 | 1,22662565 |
| -0,00041726 | -0,33472059 | 0,00087539 | 0,67094402 |
| -0,00040224 | -0,33041953 | 0,00073142 | 0,40749575 |
| -0,0002157 | -0,40052152 | 0,00065881 | 0,32217413 |
| -0,00031072 | -0,66738057 | 0,00074703 | 0,39685836 |
| -0,00034106 | -0,66213599 | 0,0005696 | 0,40731667 |
| -0,00048552 | -0,96444548 | 0,00036849 | 0,84442025 |
| -0,00030431 | -1,11151256 | 0,00038037 | 0,92635483 |
| -0,000354 | -0,99315193 | 0,00052182 | 1,01679211 |
| -0,00019368 | -0,51630214 | 0,00047037 | 1,11759369 |

| | | | |
|---|---|---|---|
| 6,62E-05 | -0,57422919 | 0,00052267 | 1,38683813 |
| 8,52E-05 | -0,63352437 | 0,00048608 | 1,20524578 |
| -7,61E-05 | -0,54501515 | 0,00057827 | 1,25293777 |
| -9,68E-06 | -0,59305964 | 0,00046531 | 1,5284205 |
| -0,00011654 | -1,10256389 | 0,00057993 | 1,55558391 |
| -0,00028122 | -1,4795624 | 0,00061678 | 1,46286515 |
| -0,00018519 | -0,86442623 | 0,00058047 | 1,46813628 |
| -0,00011207 | -0,57126334 | 0,00049901 | 1,13713501 |
| -9,54E-06 | -0,03822176 | 0,00042756 | 0,82118753 |
| 0,00012783 | 0,30129843 | 0,00026874 | 0,91489541 |
| 0,00025525 | 1,10147961 | 0,00011659 | 0,42622922 |
| 0,00016363 | 0,71455699 | 5,97E-05 | 0,10098466 |
| 0,00016928 | 0,66311417 | -1,09E-05 | 0,09349847 |
| 9,34E-05 | 0,27432196 | -4,25E-05 | -0,02715152 |
| 3,76E-05 | 0,23763889 | -5,05E-05 | -0,02282267 |
| 4,89E-05 | 0,17655619 | -8,17E-06 | 0,08142684 |
| 7,52E-05 | 0,30497215 | 8,33E-05 | 0,52804522 |
| -4,26E-05 | -0,11232029 | 0,00023492 | 1,02815333 |
| -1,91E-05 | -0,06511563 | 0,00032252 | 1,22374383 |
| -3,71E-05 | -0,15273287 | 0,00058067 | 1,646598 |
| -9,21E-05 | -0,3650074 | 0,00058244 | 2,11040002 |
| -0,00012387 | -0,48693056 | 0,00054787 | 1,6440505 |
| -0,00013318 | -0,41363961 | 0,00049886 | 1,41492917 |
| -0,00015403 | -0,59759696 | 0,00048192 | 1,19018143 |
| -0,00012211 | -0,55984286 | 0,00028671 | 0,95947741 |
| -0,00014331 | -0,59468555 | 0,00026672 | 0,83773237 |
| -0,0001105 | -0,55903146 | 0,00037903 | 0,92654078 |
| -9,94E-05 | -0,54882781 | 0,00033786 | 0,91163483 |
| -0,00011439 | -0,63144736 | 0,00034748 | 1,10822108 |
| -0,0001012 | -0,57817656 | 0,00034647 | 1,33595894 |
| -3,54E-05 | -0,56957996 | 0,00025683 | 0,95566744 |
| -8,37E-05 | -0,7699327 | 0,00020423 | 0,84792512 |
| -5,36E-05 | -0,56578447 | 0,00011667 | 0,38848538 |
| -6,43E-05 | -0,60730625 | 0,00015562 | 0,27111487 |
| -9,01E-05 | -0,60808223 | 0,00011936 | 0,12710174 |
| -0,00017261 | -0,71738323 | 0,00031695 | 0,52303154 |
| -0,00014353 | -0,68203018 | 0,00038579 | 0,84776147 |
| -0,00012841 | -0,64045236 | 0,00042254 | 1,18050286 |
| -0,00021575 | -0,57923762 | 0,00051761 | 1,34257598 |
| -0,00019604 | -0,48512695 | 0,00049697 | 1,21466604 |
| -0,00018445 | -0,34762921 | 0,00041984 | 1,10010051 |
| -0,00022711 | -0,37928976 | 0,00051087 | 0,97508577 |
| -0,0003344 | -0,82703258 | 0,0006335 | 1,12150408 |
| -0,00025831 | -0,84296601 | 0,00045042 | 0,90164432 |
| -0,00034518 | -1,14718406 | 0,00059202 | 1,184346 |

| | | | |
|---|---|---|---|
| -0,00043156 | -1,29891123 | 0,00071979 | 1,16379266 |
| -0,0004895 | -1,62199387 | 0,00062023 | 1,38603317 |
| -0,00043303 | -1,27182366 | 0,00054064 | 1,09902015 |
| -0,0005078 | -1,43555774 | 0,00069098 | 1,77535653 |
| -0,00052864 | -1,38655213 | 0,00083428 | 2,06237666 |
| -0,00039736 | -1,30742322 | 0,00062366 | 1,87185555 |
| -0,00037882 | -0,62377658 | 0,00057653 | 1,20384487 |
| -0,00039767 | -0,66239074 | 0,00064614 | 1,40493673 |
| -0,00037203 | -0,60163894 | 0,00054098 | 1,13919254 |
| -0,00030285 | -0,60743572 | 0,00035894 | 1,00719409 |
| -0,00031582 | -0,78230881 | 0,00033505 | 0,99165425 |
| -0,0003888 | -1,11691597 | 0,00034818 | 1,01255541 |
| -0,00040206 | -1,34390032 | 0,00046839 | 0,95012075 |
| -0,00032958 | -1,12227544 | 0,00045369 | 0,770436 |
| -0,00027969 | -1,06684572 | 0,00049919 | 0,8767087 |
| -0,0002987 | -0,52004869 | 0,00051162 | 0,58611423 |
| -0,00023643 | -0,91452968 | 0,00052032 | 1,18198637 |
| -0,0001462 | -0,50616363 | 0,00031162 | 1,22515529 |
| -0,00014613 | -0,78862526 | 0,00034726 | 1,82314304 |
| -5,45E-05 | -0,74219515 | 0,00032171 | 1,55564802 |
| -0,00011225 | -1,27884103 | 0,00042593 | 2,58094039 |
| -5,77E-05 | -0,88350255 | 0,00035369 | 1,8020788 |
| -6,41E-05 | -1,09431326 | 0,00029935 | 1,54685977 |
| -8,07E-05 | -0,8717539 | 0,00023537 | 1,05425307 |
| -0,00021368 | -1,17434595 | 0,00025463 | 1,21193017 |
| -0,00018285 | -1,10490099 | 0,00020453 | 1,14470539 |
| -0,00034831 | -1,4746522 | 0,00038547 | 1,54662017 |
| -0,00040615 | -1,64600275 | 0,00043914 | 1,65548742 |
| -0,0004984 | -1,6286086 | 0,00060477 | 1,63462213 |
| -0,00047417 | -1,33674261 | 0,00057828 | 1,61020532 |
| -0,00039603 | -0,99100462 | 0,00051409 | 1,43784475 |
| -0,00031402 | -0,86145395 | 0,00043839 | 1,25448613 |
| -0,00029832 | -0,97215063 | 0,00044662 | 1,33236634 |
| -0,00040516 | -1,0566799 | 0,00055452 | 1,73874547 |
| -0,00039084 | -0,82018457 | 0,00055324 | 1,40083877 |
| -0,00050363 | -1,17771303 | 0,00059342 | 1,34681276 |
| -0,00079763 | -1,04100225 | 0,00077656 | 1,85730933 |
| -0,00070662 | -0,77192166 | 0,00069001 | 1,63113484 |
| -0,00052849 | -0,91322513 | 0,00045961 | 1,4993115 |
| -0,00068255 | -0,97159966 | 0,00052137 | 1,73626848 |
| -0,00060918 | -0,54726779 | 0,000497 | 1,70928331 |
| -0,00044522 | -0,5339028 | 0,00038335 | 1,31243019 |
| -0,00054458 | -0,85080424 | 0,00038146 | 1,24010915 |
| -0,00032817 | -0,37181425 | 0,00026076 | 0,71227456 |

j. Matrices de confusion de l'ensemble des prédictions des différents modèles réajustés avec la variable « effet de surprise »

*(EUR/USD)*

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 167294 | 137688 |
| *Réalité: Direction -* | 163831 | 151062 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 112043 | 192939 |
| *Réalité: Direction -* | 101210 | 213683 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 129643 | 175339 |
| *Réalité: Direction -* | 118955 | 195938 |

*(USD/JPY)*

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 163738 | 136639 |
| *Réalité: Direction -* | 161707 | 157791 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 84382 | 215995 |
| *Réalité: Direction -* | 75185 | 244313 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 96513 | 203864 |
| *Réalité: Direction -* | 87644 | 231854 |

## k. Matrices de confusion des prédictions post-annonces macroéconomiques des différents modèles avec et sans la variable « effet de surprise »

*Modèles SANS la variable « effet de surprise » (EUR/USD)*

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 228 | 203 |
| *Réalité: Direction -* | 250 | 251 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 154 | 277 |
| *Réalité: Direction -* | 178 | 323 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 170 | 261 |
| *Réalité: Direction -* | 195 | 306 |

*Modèles SANS la variable « effet de surprise » (USD/JPY)*

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 229 | 234 |
| *Réalité: Direction -* | 229 | 240 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 125 | 338 |
| *Réalité: Direction -* | 120 | 349 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 147 | 316 |
| *Réalité: Direction -* | 135 | 334 |

*Modèles AVEC la variable « effet de surprise » (EUR/USD)*

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 281 | 150 |
| *Réalité: Direction -* | 190 | 311 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 281 | 150 |
| *Réalité: Direction -* | 192 | 309 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 257 | 174 |
| *Réalité: Direction -* | 183 | 318 |

*Modèles AVEC la variable « effet de surprise » (USD/JPY)*

| *Modèles ARMA* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 335 | 128 |
| *Réalité: Direction -* | 131 | 338 |

| *Régressions logistiques* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 341 | 122 |
| *Réalité: Direction -* | 124 | 345 |

| *Perceptrons multicouches* | *Prédiction: Direction +* | *Prédiction: Direction -* |
|---|---|---|
| *Réalité: Direction +* | 326 | 137 |
| *Réalité: Direction -* | 125 | 344 |

Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve, Belgique **www.uclouvain.be/lsm**