Table des matières

Introduction au rapport

Le présent rapport se veut être l'analyse de différentes séries chronologiques. Plus précisément, ce rapport a pour but de trouver le modèle qui s'ajuste le plus exactement à nos échantillons de données. Le travail en bref sera alors de séparer nos bases de données en un échantillons *entraînement* qui servira à trouver un modèle et en un échantillon *test* qui servira à valider la précision du modèle retenu. Le travail est fait à partir de deux bases de données pour un total de cinq variables à modéliser. La première base de données, traitée au numéro 1, contient les données mensuelles du taux de change du dollar américain par rapport à la monnaie Euro de janvier 1999 à décembre 2016. La seconde base de données contient une séries de statistiques de la *SAAQ*, soit le nombre d'accidents automobiles avec dommages corporels, le nombre de personnes accidentés, le nombre de demandes d'indemnités et le coût total de lindemnisation (en dollar constants 2015) pour les années 1978 à 2015 inclusivement. Ces données seront traîtées au numéro 2. Afin d'alléger la lecture du présent rapport, les figures et tableaux furent placés en annexe.

Taux de change US/Euro

Analyse primaire du jeu de données

Tout d'abord, nous observons, à la *Figure 1* en annexe, la série chronologique du taux de change américain/européen depuis janvier 1999 jusqu'à décembre 2016. Les données ont été collectées de façon quotidienne pour ensuite être transformées en taux mensuelles.

Grâce à une première analyse, on remarque sur le tableau de la fonction d'autocorrélation échantillonale qu'il y a une forte autocorrélation dans notre jeu de données et que celle-ci diminue lentement plus le lag augmente (voir *Figure 2*). On soupçonne donc fortement ce processus d'être non-stationnaire. Il n'est pas nécessaire de tracer le graphique de la fonction d'autocorrélation partielle puisque on doit d'abord stationnarisé notre processus avant de réellement débuter son analyse.

Test de stationnarité

Le test de stationnarité de Dickey-Fuller est alors effectué sur notre base de données. On remarque que pour un ordre $k = 2$ de processus auto-autoregressif tel que proposé par la fonction *ar* de *R*, notre processus est non stationnaire à un niveau de confiance de 5% selon ce test avec une *p-value* très forte de 91.33%. On vérifie alors si une transformée Box-Cox est apppropriée à notre modèle avant de tester la stationnarité d'une première différentiation.

Transformation des données

Étant donné que les données sont positives, il est possible d'utiliser la transformée de Box-Cox afin de stabiliser notre processus. On rappelle que la famille des fonctions de puissance est définie de la façon suivante:

$$g(x) = \frac{x^\lambda - 1}{\lambda} \times 1_{\{\lambda \neq 0\}} + \ln(x) \times 1_{\{\lambda = 0\}}$$

λ est donc déterminé en maximisant la fonction de log-vraisemblance de nos données fournie à la *Figure 3*.

On constate que $\lambda = 0.1$ semble être l'estimé MLE situé au centre de l'intervalle de confiance 95%, soit $]-0.5, 0.7[$. Puisque $\lambda = 0$ est dans notre intervale de confiance, cette valeur du paramètre est également à considérer. Dans le cas ou le processus serait à différencier une seule fois, la transformation logarithmique est particuliairement attirante puisqu'elle permet de modéliser non pas le «prix» du taux de change, mais sont rendement comme il est usage de la faire dans le cas d'outils financiers. En effet, soit Y_t le prix d'un outil financier au temps t,le modèle logarithmique modélise le rendement de la façon suivante.[1] SI ON GARDE PLUS LOIN LE MODELE ARIMA(0,1,1) CELA REVIENS AU MODELE LOGNORMALE (A MENTIONNER)

$$Y_t = Y_{t-1} e^X$$

Où X, le rendement continu mensuelle, est la variable aléatoire à modéliser. Ainsi, suite à une première différentiation de notre modèle logarithmique, on obtient direction ce rendement. Soit,

$$log\left(\frac{Y_t}{Y_{t-1}}\right) = X$$

On conservera alors cette transformation à condition que la première différentiation de notre processus soit stationnaire. Le graphique de cette transformation est affiché à la *Figure 4*.

1. WEISHAUS, Abraham, SOA Exam MFE Models for Financial Economics 10th Edition.

On effectue alors à nouveau , suite à une première différentiation, le test augmenté de Dickey-Fuller. On remarque cette fois qu'avec un processus autorégressif d'ordre 1 tel que suggéré par la fonction *ar*, on ne peut rejetter l'hypothèse de stationnarité avec un p-value inférieur à 1%. Ainsi, tel que mentionnée précédement, la transformation logarithmique sera concervée. Le graphique de la première différentiation du logarithme de la série chronologique à l'étude se trouve à la *Figure 5*.

Identification du modèle

Maintenant que la série est stationnaire, on peut en identifier le modèle. La première différentiation du logarithme du taux de change US/Euro sera dorénevent notre modèle de base. On fera ainsi référence à ce modèle par défaut.

On observe les fonctions d'autocorrélation et d'autocorrélation affichées aux *Figure 6* et *Figure 7* de la série à l'étude. On remarque que la fonction d'autocorrélation suggère fortement un modèle à moyenne mobile d'ordre 1, soit IMA(1,1) pour notre modèle tansformé alors que la fonction d'autocorrélation partielle suggère un modèle autorégressif d'ordre 1, soit ARI(1,1). On testera alors également le modèle ARIMA(1,1,1) qui est suggérer par la combinaison de ces graphiques.

On observe ensuite la fonction d'autocorrélation étendue afin de voir si ce test à un autre modèle à proposer. Le tableau de la fonction EACF est à la *Table I*. On observe d'abord de ce tableau que le modèle IMA(1,1) est suggéré pour la différentiation de notre modèle transformé initial. On peut également chercher à savoir si la valeur du o en ARMA(0,1) et du x en ARMA(1,1) sont significativement différent de leur valeurs inverses (x pour le o et vice versa), sans quoi le modèle ARIMA(1,1,1) serait également suggeré par le tableau EACF. Cependant, comme il fut déjà décidé d'évaluer la pertinance du modèle ARIMA(1,1,1) suite à l'analyse des fonction d'autocorrélation et d'autocorrélation partielle, il n'est pas nécessaire de test la pertiance de ce modèle selon le EACF.

Le dernier test effectué pour ajuster un modèle à notre série chronologique est le test du BIC. Le tableau du BIC est tracé à la *Table II* propose un modèle ARI(1,1). L'ordre maximal de moyenne mobile pour ce test a été établi à 1, sans quoi la fonction *armasubsets* permettant de tracer le tableau du BIC retourne des avertissements de forte dépendance linéaire entre les paramètres. On étudiera ainsi les quatre tests proposés par les tests précédents qui peuvent être résumé de la façon suivante;

Résumé	
Tests et tableaux	Modèles suggérés
ACF	IMA(1,1) ARIMA(1,1,1)
PACF	ARI(1,1) ARIMA(1,1,1)
EACF	IMA(1,1)
BIC	ARI(1,1)

On teste alors les modèles ARI(1,1), IMA(1,1) et ARIMA(1,1,1) qui sont développé, dans l'ordre, des façons suivantes.

$$W_t - \mu = \phi(W_{t-1} - \mu) + e_t \quad (1)$$
$$W_t - \mu = e_t - \theta e_{t-1} \quad (2)$$
$$W_t - \mu = \phi(W_{t-1} - \mu) + e_t - \theta e_{t-1} \quad (3)$$

Où

$$W_t = \nabla log(Y_t)$$

Et e_t est un bruit blanc et μ la moyenne de la série chronologique non-centrée. Il est ici important de noter que les paramètres estimés $\hat{\theta}$ et $\hat{\phi}$ ne seront pas les mêmes dans les trois modèles et devront donc être estimés pour chacun d'entre eux.

Estimation des paramètres des différents modèles et test de résidus

On estime alors les valeurs des différents paramètres selon la méthode du maximum de vraisemblance à l'aide de la fonction *arima*. Les résultats se trouvent à la *Table III*.

On peut alors faire l'analyse de nos résidus pour faire notre choix parmi nos trois modèles. Les résidus standardisés de ces trois modèles sont tracés à la *Figure 8*. On remarque d'abord que ces trois courbes se superposent et ne peuvent donc pas être utilisées pour sélecionner un modèle. De plus, on remarque de ce graphique que les valeurs de résidus supérieures à 2 en valeur absolue sont relativement fréquente ce qui pousse à questionner la théorie selon laquelle les résidus suivent une loi normale. En effet, dans le cas des résidus normaux, les valeurs supérieurs à 2 en valeur absolue devrait ce que matérialiser dans environ 2.28% des cas. La fréquence observée est nettement supérieur.

On regarde ensuite les histogrammes et les graphiques QQ aux *Table V* et *Figure 9* des résidus. On remarque que les trois histogrammes semblent tracé avec efficacité la cloche de la loi normal. De plus, on remarque des trois graphiques QQ, que les résidus observés sont en tout point simiaire à leur valeurs théoriques en cas de normalité. Ces deux test semblent donc justifier la distribution normale des résidus de nos trois modèles.

On effectue alors, pour tenter de départager nos modèles, un dernier test, soit le test de Shapiro-Wilk évaluant le dégré de corrélation entre les quantiles des résidus standardisés et la loi normale standard. L'hypothèse nulle étant que la distribution des résidus suit une loi normale, on confirme, avec des p-values respectivent de 0.3247, 0.2846 et 0.3051, que les résidus engendrés par les modèles IMA(1,1), ARI(1,1) et ARIMA(1,1,1) sont tous normalement distribués pour un test à un niveau de confiance α de 5%.

On se trouve ainsi dans l'impossibilité de départager nos modèles à l'aide de l'étude de leurs résidus. On choisit alors de conserver le modèles IMA(1,1) parce qu'il est simple, parce que la valeur maximale de sa fonction de vraisemblance est la plus élevée et la valeur de AIC est la plus faible selon la *Table III.* De plus, ce modèle nous permet de retrouver un modèle de loi lognormale pour modéliser le taux de US/Euro qui est modèle fortement utilisé dans le monde de la mathématique finanière.

f)

Par la suite, il faut estimer les paramètres de chacun des modèles énumérés en (e) par la méthode du maximum de vraisemblance (MLE).

Pour l'ARIMA(0,1,1), le paramètre θ_1 de notre MA(1) est de 0.3118, ce qui nous donne le modèle suivant:

$$logY_t = logY_{t-1} + e_t + 0.3118e_{t-1}$$

Pour l'ARIMA(1,1,0), le paramètre ϕ_1 de notre AR(1) est de 0.2933, ce qui nous donne le modèle suivante:

$$logY_t = 1,2933\ logY_{t-1} - 0.2933\ logY_{t-2} + e_t$$

Pour l'ARIMA(0,1,3), le paramètre ϕ_1 de notre AR(1) est de 0.2933, ce qui nous donne le modèle suivante:

voir chapitre 5 pour écrire le modèle

Annexe

Tableaux et figures

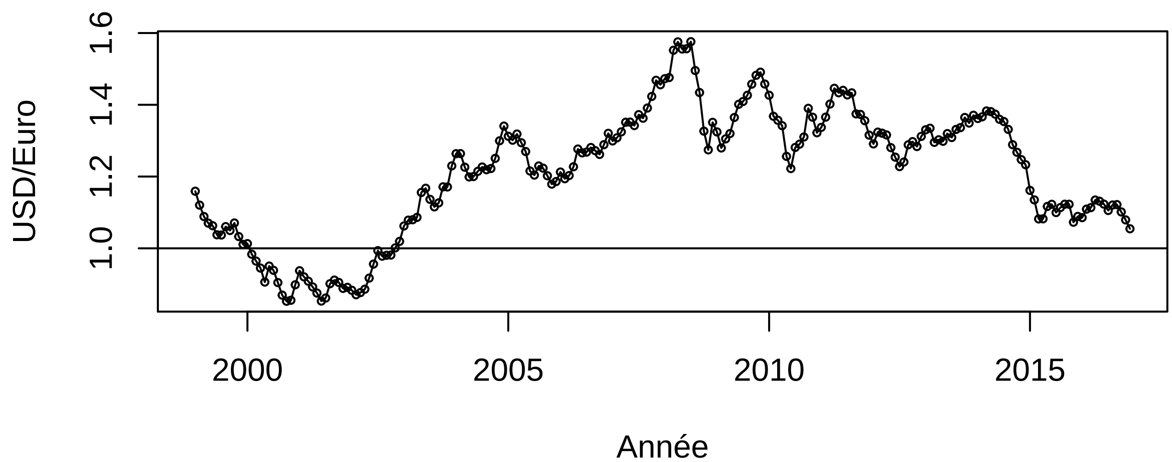

Figure 1. Taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

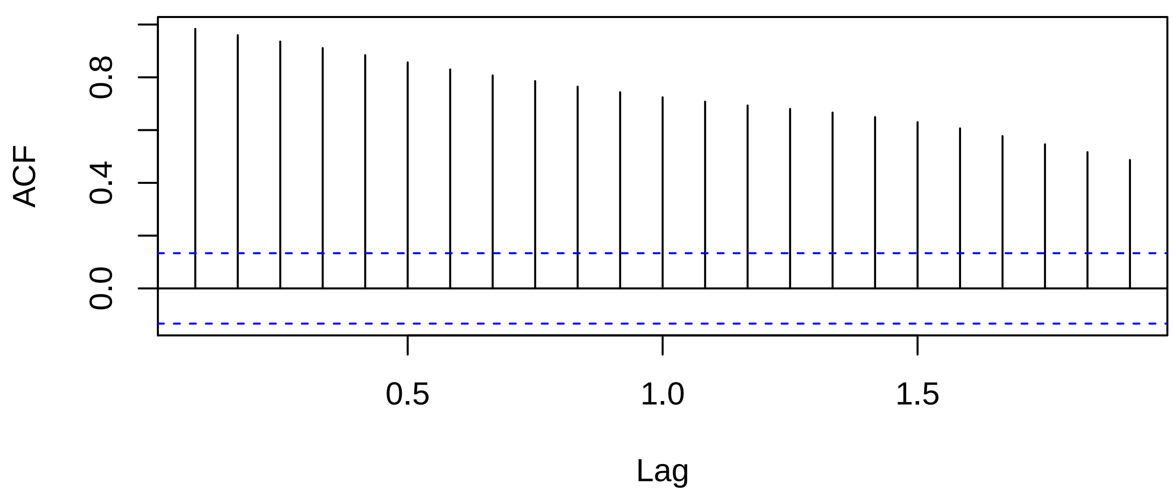

Figure 2. Fonction d'autocorrélation échantillonales du taux de change US/Euro. Le lag étant exprimé en années.

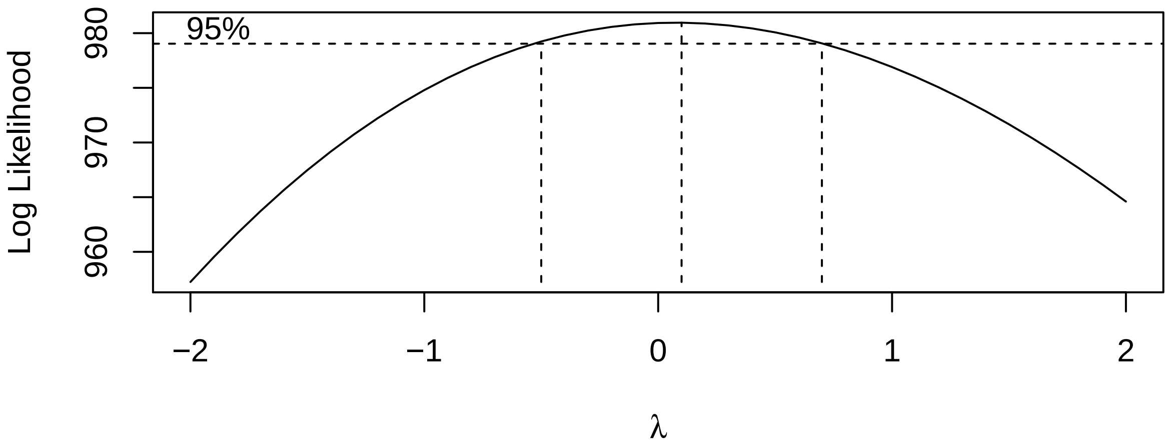

Figure 3. Fonction du logarithme de vraissemblance de la transformée de Box-Cox de la série chronologique du taux de change US/Euro.

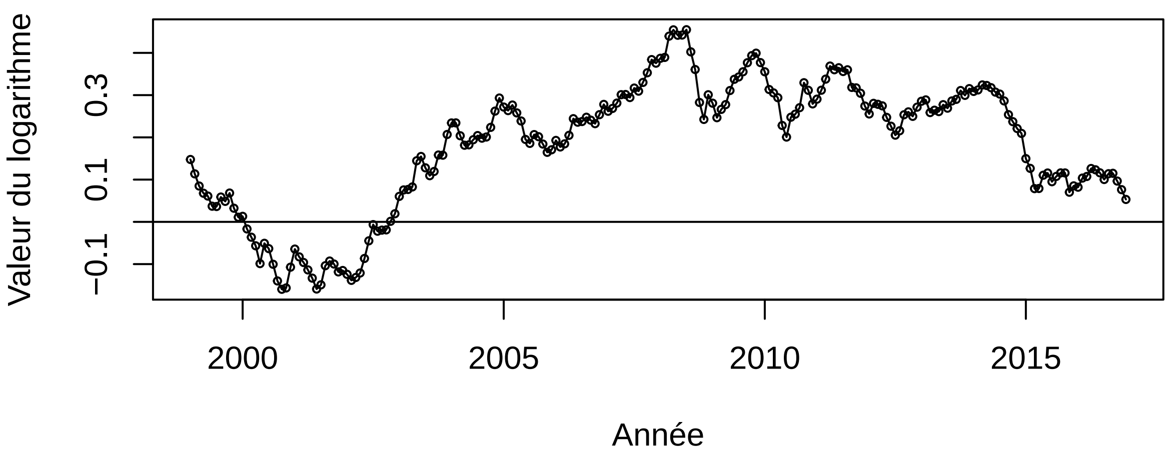

Figure 4. Logarithme du taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

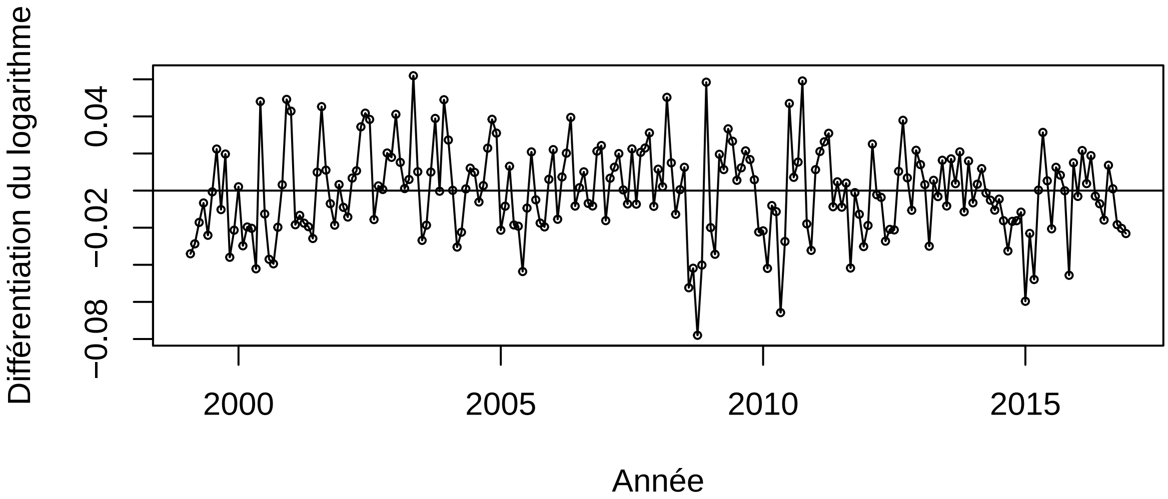

Figure 5. Première différentiation du logarithme du taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

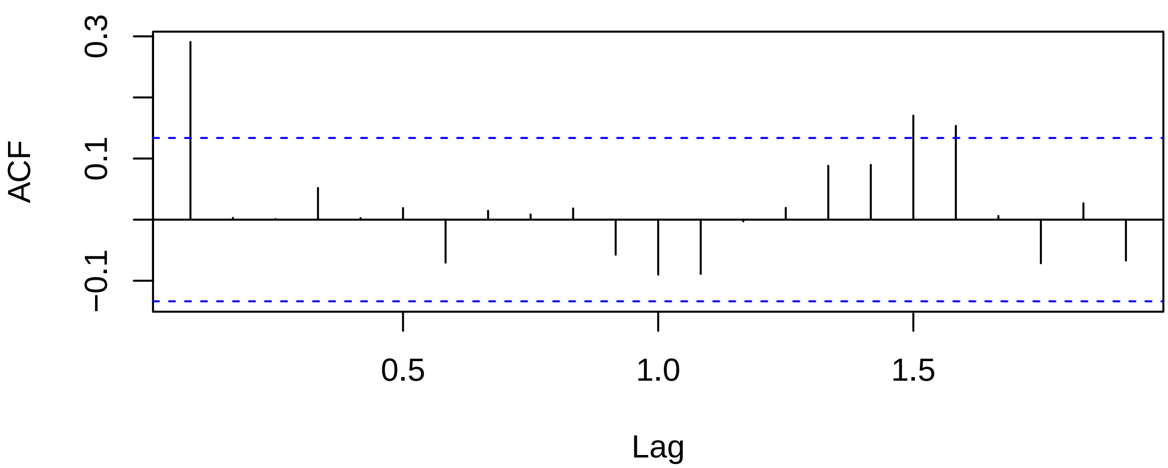

Figure 6. Fonction d'autocorrélation de la première différentiation du logarithme du taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

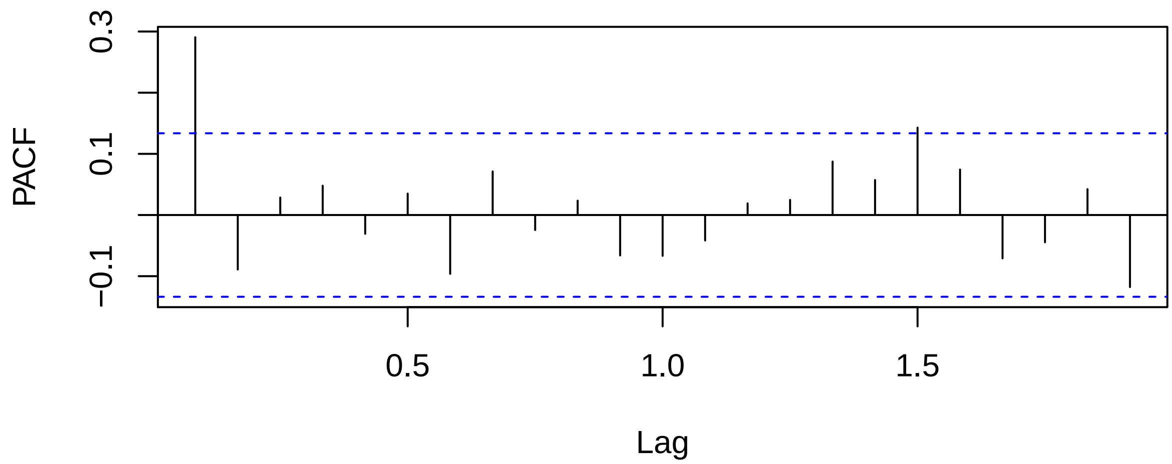

Figure 7. Fonction d'autocorrélation partielle de la première différentiation du logarithme du taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 x o o o o o o o o o o  o  o  o
## 1 x x o o o o o o o o o  o  o  o
## 2 x o o o o o o o o o o  o  o  o
## 3 x x x x o o o o o o o  o  o  o
## 4 x x o x o o o o o o o  o  o  o
## 5 x o x x o o o o o o o  o  o  o
## 6 x x x x o o o o o o o  o  o  o
## 7 x x x x x x o o o o o  o  o  o
```

Table I. Tableau de la fonction EACF de la première différentiation du logarithme du taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

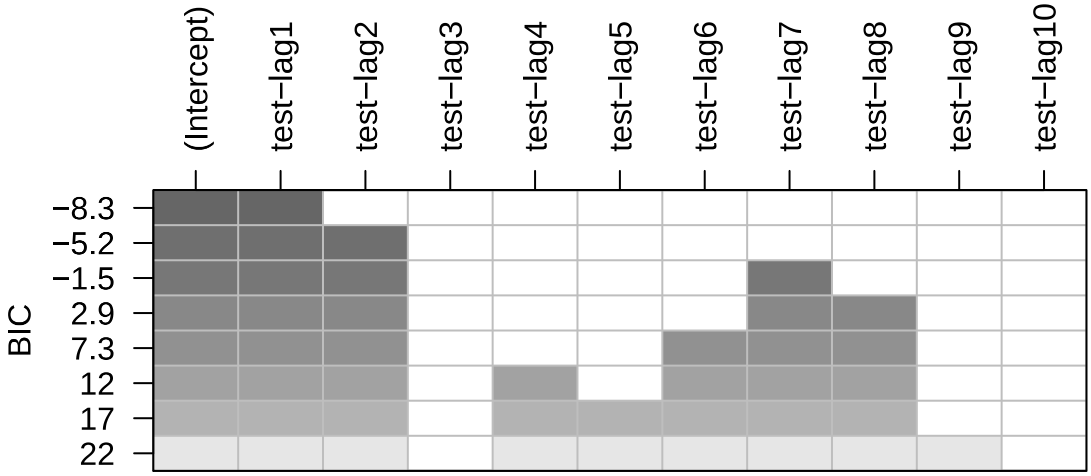

Table II. Tableau de la fonction BIC de la première différentiation du logarithme du taux de change US/Euro mensuelle de janvier 1999 à décembre 2015.

Modèles	Paramètres	Valeurs estimés	AIC	Logarithme du max de vraisemblance
IMA(1,1)	μ	0	-1010	506
	θ	0.3118		
ARI(1,1)	μ	0	-1009	505
	ϕ	0.2933		
ARIMA(1,1,1)	μ	0	-1008	506
	θ	0.2635		
	ϕ	0.0557		

Table III. Tableau des valeurs estimées des différents coefficients obtenus de la fonction *arima* pour les trois modèles à l'étude du logarithme du taux de change US/Euro.

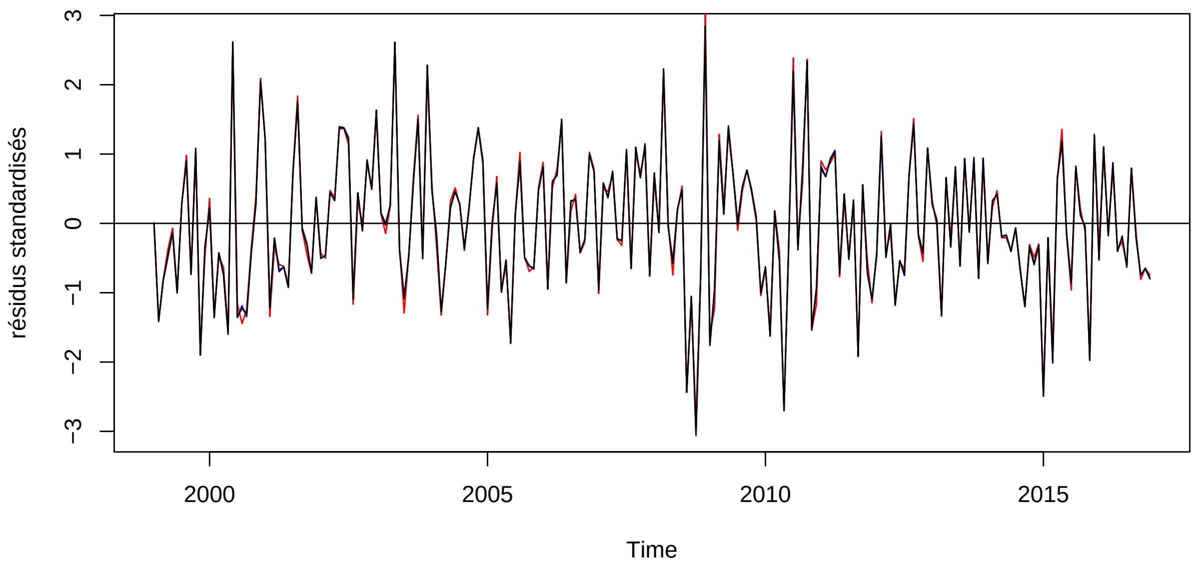

Figure 8. Superposition des graphiques des résidus standardisés des modèles ARI(1,1), IMA(1,1) et ARIMA(1,1,1) du logarithme du taux de change US/Euro.

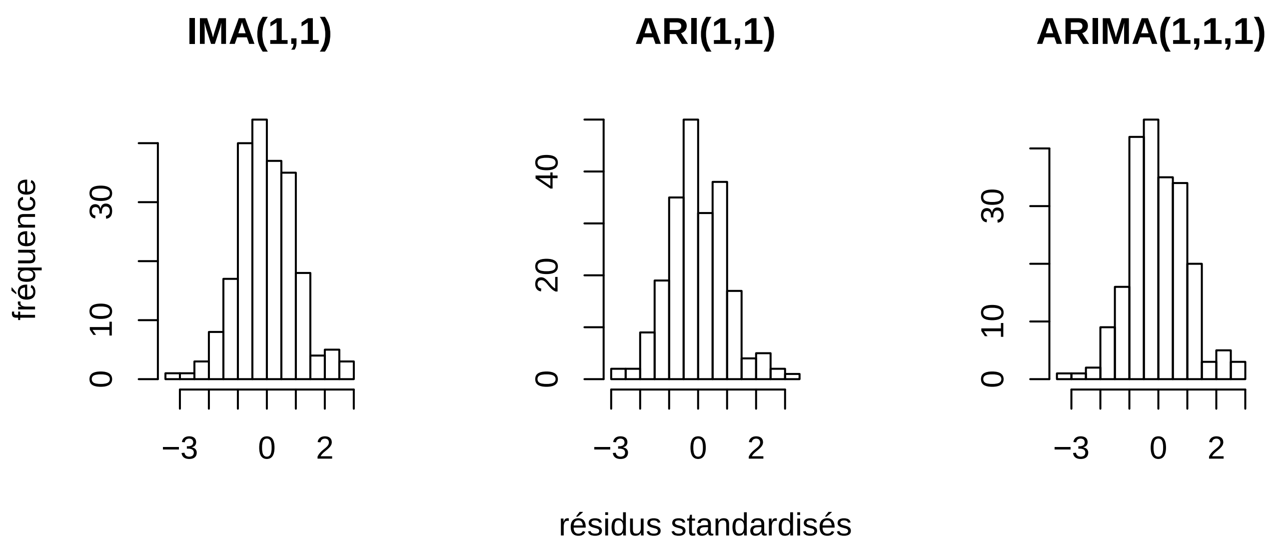

Table IV. Histogramme des résidus standardisés des modèles étuidés du logarithme du taux de change US/Euro.

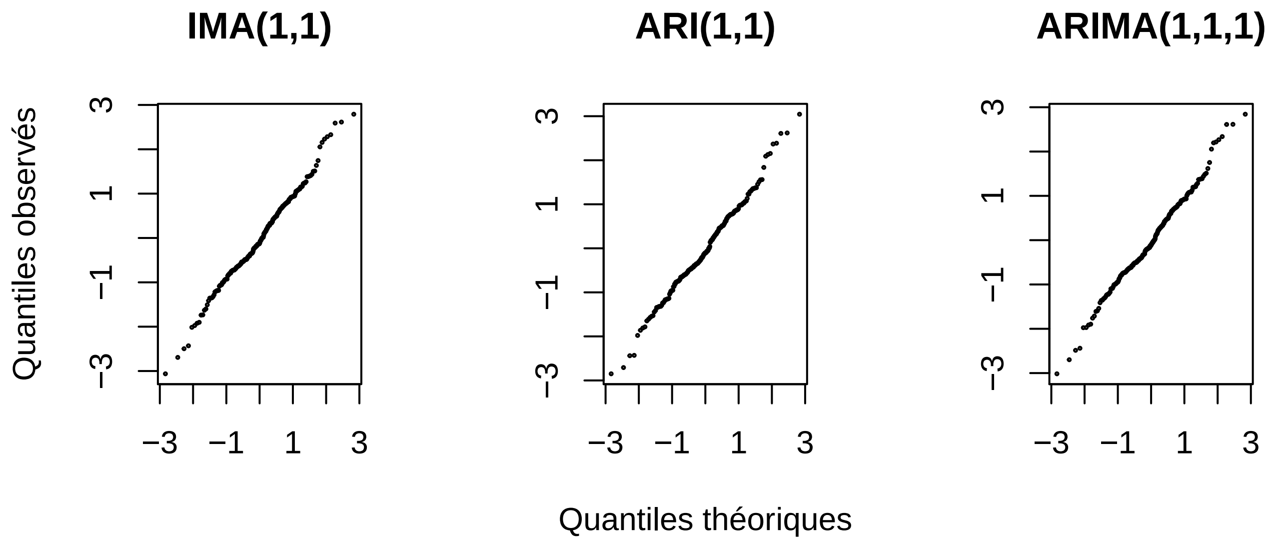

Figure 9. Graphiques QQ des modèles étuidés du logarithme du taux de change US/Euro.

Code informatique

```
library('TSA')
library('tseries')

# Importation des données
taux <- read.csv2("C:/Users/TEMP/Documents/GitHub/TP_ACT2010/Taux_de_change_US_Euro.csv")
rendement<-taux$US.Euro
anne.mois<-taux$Année.mois

# création de la série chronologique
ttaux<-ts(rendement,start=c(1999,1),end=c(2016,12),frequency = 12)
```

```
# Test augmenté de Dickey-Fuller (ADF)
ar(log(ttaux))
adf.test(log(ttaux), k=2)
```

```
ttaux011 <- arima(log(ttaux), order=c(0,1,1), method='ML')
```