
Value at Risk, Projet 1

Ecrit par
Sami AZZOUG

1^{er} novembre 2021

Table des matières

1	Introduction	2
2	Préparation des données	3
3	Analyse du chronogramme	4
4	Tests de Racine Unitaire	5
4.1	Test de Dickey-Fuller (DF)	5
4.2	Dickey-Fuller augmenté	9
4.3	Test de Zivot et Andrews (ZA)	10
4.4	Test de Lee et Strazicich (LS)	13
5	Asymétrie perte/gain	15
5.1	Skewness	15
6	Queues de distribution épaisses	16
6.1	Kurtosis	16
7	Autocorrélations des carrés des rendements	17
7.1	ACF	17
7.2	PACF	18
7.3	Ljung-Box	19
8	ARMA(P,Q)	20
8.1	ARMA(0,1)	21
8.2	Résidus	22
8.3	Standardisation des résidus	23
8.4	BIC	23
9	Clusters de volatilité	24
10	Queues épaisses conditionnelles	25
11	Effet de levier	26
12	La saisonnalité	27
12.1	Effet Week-End	27
12.2	Effet Janvier	28
13	Conclusion	30
14	Annexes	31

1 Introduction

Ce projet est consacré à l'étude d'une série temporelle à l'aide du logiciel R.

Le choix de l'entreprise est Shiseido qui est la plus vieille société de cosmétique au monde.

Shiseido, qui signifie en idéogrammes japonais « La Maison qui contribue à la qualité de la vie », porte en son nom sa vocation depuis 1872.

Shiseido est une société dont les fondements s'appuient sur la créativité. Quand son fondateur Yushin Fukuhara a ouvert les portes de Shiseido en 1872, il a initié le Japon à la première pharmacie de style occidental.

Le jeu de données provient de yahoo.fr, il porte sur les rendements de cette entreprise et il s'étale du 1 avril 2010 au 31 octobre 2021.

L'objectif de cette étude est double, nous aspirons à pouvoir avoir une idée précise des caractéristiques de l'action Shiseido.

Puis, dans un second temps, nous chercherons à déterminer un modèle adéquat du processus générateur des données (PGD) de la série, afin de pouvoir, si désiré, réaliser des prévisions sur les valeurs futures de l'action Shiseido.

2 Préparation des données

La préparation des données requiert de pouvoir bien définir quel est l'échantillon d'estimation (rte) ainsi que celui de test (rt).

Nous veillons à préparer correctement les dates afin de savoir quels sont leurs dates précise de début et de fin prise pour l'étude.

Pour ce faire nous effectuons une vérification au préalable par la commande suivante :

```
dates_rt = dates[-1]
dates_rte = dates_rt[1:1966]
dates_rte[1966]
```

FIGURE 1 – L'horizon de l'étude.

Notre échantillon rte s'étend alors du 4 janvier 2010 au 29 décembre 2017 et il comporte 1966 observations.

Nous veillons également à avoir un horizon de 10 ans afin de pouvoir avoir suffisamment de recul sur la série.

3 Analyse du chronogramme

Nous pouvons donc commencer l'étude de notre série par un chronogramme.

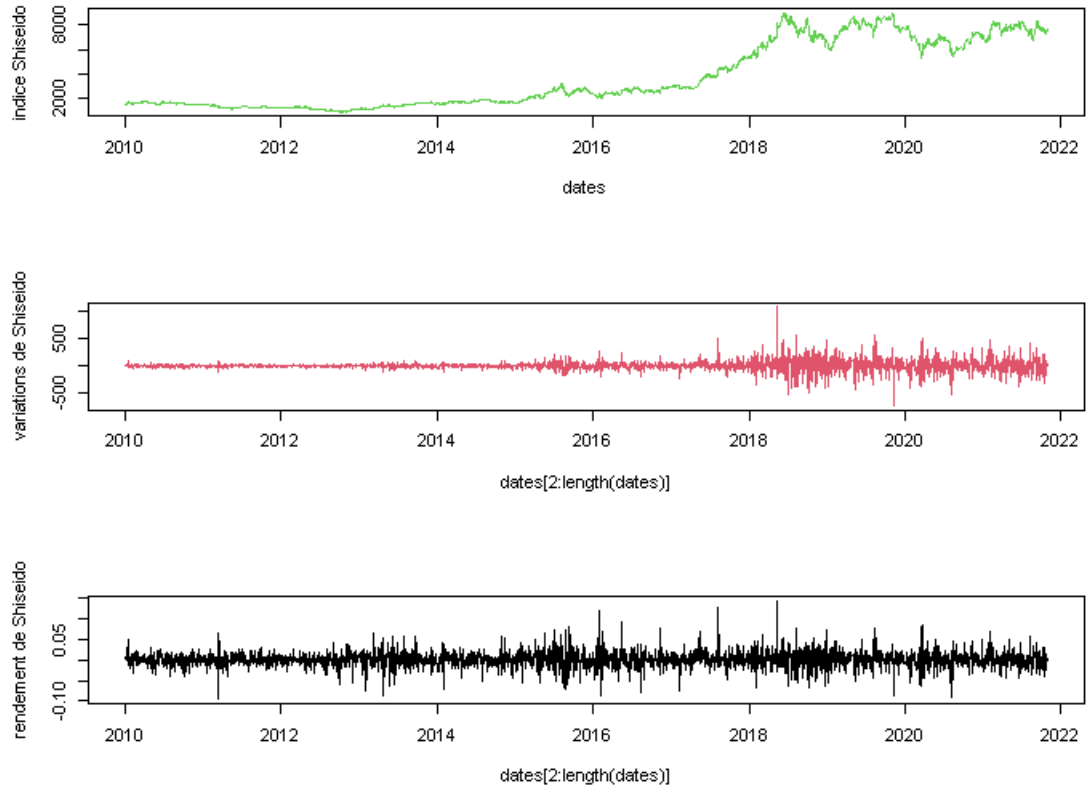


FIGURE 2 – Chronogramme de Shiseido.

Nous observons dès lors une tendance croissante de l'indice de Shiseido alors que son rendement logarithmique et ses variations varient autour de 0.

Nous pouvons néanmoins remarquer que les variations autour de 0 tendent à augmenter avec le temps et plus particulièrement entre 2018 et 2020.

Il semblerait également y avoir des paquets de volatilités entre 2015 et 2019.

4 Tests de Racine Unitaire

4.1 Test de Dickey-Fuller (DF)

Les tests de racine unitaire permettent de déterminer si le PGD est stationnaire, DS ou TS.

Afin de debuter ces tests, nous allons commencer par réaliser le test de Dickey-Fuller en recherchant tout d'abord la bonne spécification puisqu'une seule est correcte parmi les trois, puis en examinant le rejet ou non de l'hypothèse nulle.

- (i) Dans la spécification "trend", la série possède une tendance et une constante. On considère ce modèle :

$$\Delta X_t = (\rho - 1)X_{t-1} + \beta_0 + \beta_1 \text{tendance}_t + \epsilon_t$$

- (ii) Avec la spécification "drift", la série possède une constante mais ne possède pas de tendance. On considère ce modèle :

$$\Delta X_t = (\rho - 1)X_{t-1} + \beta_0 + \epsilon_t$$

- (iii) La spécification "none" ne possède ni tendance, ni constante. On considère ce modèle :

$$\Delta X_t = (\rho - 1)X_{t-1} + \epsilon_t$$

Au terme de chacun des trois modèles, sous l'hypothèse nulle, nous allons conclure à la présence d'une racine unitaire. Nous commençons par effectuer le test avec la spécification "trend". Nous donnons les résultats du test dans la figure 3.

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt)

Residuals:
    Min       1Q   Median       3Q      Max
-0.096308 -0.008663  0.000571  0.008928  0.127085

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.196e-04  8.046e-04  -0.894   0.3712
z.lag.1      -1.006e+00  2.258e-02 -44.546 <2e-16 ***
tt           1.386e-06  7.095e-07   1.954   0.0509 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01782 on 1962 degrees of freedom
Multiple R-squared:  0.5028,    Adjusted R-squared:  0.5023
F-statistic: 992.2 on 2 and 1962 DF,  p-value: < 2.2e-16

Value of test-statistic is: -44.5461 661.4517 992.1769

Critical values for test statistics:
      1pct   5pct 10pct
tau3  -3.96 -3.41 -3.12
phi2   6.09  4.68  4.03
phi3   8.27  6.25  5.34
```

FIGURE 3 – Test Dickey-Fuller avec spécification trend.

Nous observons que le coefficient associé à tt n'est pas significatif car sa p-value n'est pas inférieur à 0,05. Dès lors, nous allons devoir passer à la spécification "drift" et allons présenter les résultats en figure 4.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0006415  0.0004026   1.593   0.111
z.lag.1      -1.0040744  0.0225780 -44.471 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

FIGURE 4 – Test Dickey-Fuller avec spécification drift.

La spécification drift s'avère ne pas être adéquate également car le coefficient associé à B_0 n'est pas significatif. De ce fait, il nous reste à exécuter la spécification "none" en figure 5.

```

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
z.lag.1 -1.00276    0.02257  -44.42  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01784 on 1964 degrees of freedom
Multiple R-squared:  0.5012,    Adjusted R-squared:  0.501
F-statistic: 1974 on 1 and 1964 DF,  p-value: < 2.2e-16

Value of test-statistic is: -44.4254

Critical values for test statistics:
      1pct   5pct 10pct
taul -2.58 -1.95 -1.62

```

FIGURE 5 – Test Dickey-Fuller avec spécification none.

C'est donc la spécification "none" qui s'avère être correcte car le coefficient restant est largement significatif puisque sa p-value est inférieur à 0.05.

Nous pouvons alors interpréter le test de Dickey-Fuller en comparant la valeur de la statistique de test qui est de -44.4254, elle est donc inférieur au critère de 5% qui est de -1.95.

Nous pouvons ainsi regarder notre $\rho - 1$ qui est égale à 0, nous pouvons donc rejeter H_0 pour $\rho - 1$ puisque la statistique t (-44.4254) est supérieur au critère de risque de 5% choisi par les auteurs qui est de -1.95.

On rejette alors H_0 et donc la présence d'une racine unitaire dans la série. Selon le test de Dickey-Fuller, le PGD de la série serait un $AR(1)$ stationnaire avec constante.

Cependant, les résultats de notre test de Dickey-Fuller ne sont valides uniquement en cas de non auto-corrélation dans les résidus. Afin de pouvoir vérifier cela, voici les résultats obtenus en figure 5.

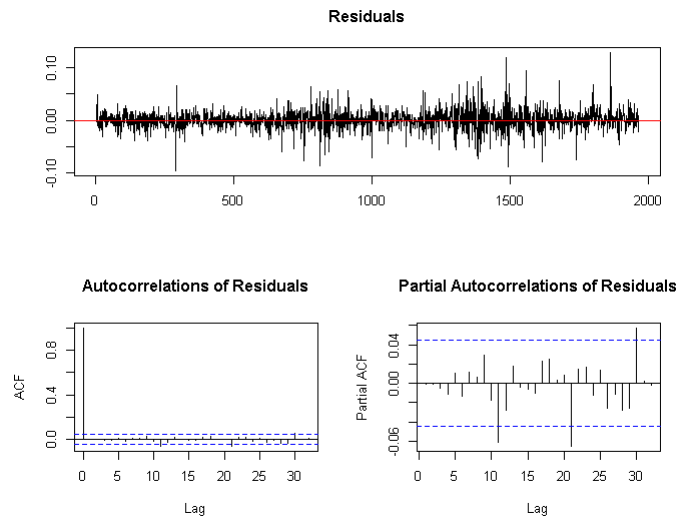


FIGURE 6 – Auto-corrélations des résidus.

En ayant réalisé un test ACF et PACF sur les résidus avec un lag de 0 et en intégrant la spécification "none" nous obtenons une autocorrélation dans les résidus puisque nous avons trois barres qui sortent de notre intervalle de confiance.

Les résultats de notre test Dickey-Fuller sont donc invalides, on ne sait pas si notre série est donc stationnaire.

Il serait alors judicieux de réaliser un test de Dickey-Fuller augmenté afin de prendre en compte cette possible autocorrélation dans les résidus et savoir si notre série est réellement stationnaire.

4.2 Dickey-Fuller augmenté

Afin de pouvoir réaliser un test de Dickey-Fuller augmenté nous devons réaliser le critère de Schwartz afin de savoir combien de γ_{20} nous devons implémenter. Nous obtenons une valeur de pmax égale à 25, cela nous servira pour utiliser le critère MAIC. Selon le critère MAIC nous devons implémenter huit variables explicatives, voir figure 7.

```
Augmented DF test
t-test statistic:      ADF test
p-value:              -1.402991e+01
Max lag of the diff. dependent variable:  9.365014e-30
                                          8.000000e+00
```

FIGURE 7 – Critère MAIC.

Nous pouvons dès lors les implémenter dans le test de Dickey-Fuller afin de savoir si le dernier γ ajouté à une statistique t en valeur absolue supérieur à 1.6, ce qui nous amènerait à dire que le γ est significatif.

Sinon, nous serons amenés à réaliser une analyse top-down jusqu'à ce que le dernier γ ne soit pas significatif, afin d'obtenir des résultats robustes.

En réalisant cela, nous nous apercevons qu'aucun des γ ajoutés n'est significatif, nous n'arrivons donc pas à prendre en compte l'autocorrélation de notre série rte.

Pour l'instant, nous ne pouvons pas comparer la valeur de $\rho-1$ avec celle calculé par les auteurs ρ .

Cependant, dans le cours Varfiche(2)¹, il nous est offert la possibilité de prendre la valeur de pmax donnée par le critère de Schwartz afin de pouvoir calculer le critère MAIC.

Nous allons donc partir de la valeur donnée de Swartz qui est de 25 est réaliser la procédure top-down dessus² en retirant les γ non significatifs.

Nous arrivons à obtenir un γ_{20} qui est significatif.

De ce fait, nous pouvons ensuite conclure sur la stationnarité de notre série en comparant la valeur de la statistique t (-9.9868) avec celle calculée par les auteurs ρ (-1.95).

```
Value of test-statistic is: -9.9868
Critical values for test statistics:
1pct 5pct 10pct
taul -2.58 -1.95 -1.62
```

FIGURE 8 – Critère de décision.

Ici nous rejetons H_0 , $(\rho - 1) = -1.055607$, donc $\rho = -0.055607$ nous pouvons donc dire que $|\rho| = 0.055607 < 1$. Nous pouvons donc dire que selon Dickey-Fuller, avec la procédure top-down, le PGD qui a généré notre série rte est stationnaire.

1. Voir figure 28
2. Voir figure 29

4.3 Test de Zivot et Andrews (ZA)

Malgrès le test de Dickey-Fuller et celui de Dickey-Fuller augmenté, nous n'avons pas pris en compte le fait qu'il puisse y avoir des changements structurels dans notre série.

Dans le test de Zivot et Andrews nous avons deux spécifications possibles :

- (i) La spécification "both". La série présente une rupture dans la tendance et dans la constante. On considère le modèle :

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \delta_1 DU_t(T_B) + \delta_2 DT_t(T_B) + \sum_{j=1}^p \gamma_j \Delta y_{t-j} + \epsilon_t$$

avec t_B la date de rupture endogène (déterminée par le test).

- (ii) La spécification "crash" ou "intercept". La série présente une rupture dans la constante uniquement. On considère le modèle :

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \delta_1 DU_t(T_B) + \sum_{j=1}^p \gamma_j \Delta y_{t-j} + \epsilon_t$$

Nous considérons ces tests-ci :

$$\begin{cases} (H_0) : & \text{Le PGD est DS, présence de racine unitaire et sans changement structurel.} \\ (H_a) : & \text{Le PGD est TS avec changement structurel.} \end{cases}$$

Nous pouvons donc réaliser le test de Zivot et Andrews, nous utiliserons la valeur de pmax utilisé pour la procédure Top-Down car c'est cette dernière qui offre une t statistique supérieur à 1.6 en valeur absolue³.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.715e-03  1.001e-03  -1.714  0.08672 .
y.l1         -2.228e-01  1.128e-01  -1.975  0.04841 *
trend        3.331e-06  1.270e-06  2.623  0.00879 **
y.d11        2.085e-01  1.098e-01  1.898  0.05785 .
y.d12        2.035e-01  1.068e-01  1.906  0.05677 .
y.d13        1.918e-01  1.036e-01  1.851  0.06433 .
y.d14        1.711e-01  1.002e-01  1.707  0.08793 .
y.d15        1.742e-01  9.663e-02  1.803  0.07158 .
y.d16        1.591e-01  9.305e-02  1.710  0.08741 .
y.d17        1.671e-01  8.937e-02  1.870  0.06162 .
y.d18        1.675e-01  8.560e-02  1.957  0.05048 .
y.d19        1.901e-01  8.156e-02  2.331  0.01986 *
y.d110       1.656e-01  7.759e-02  2.135  0.03292 *
y.d111       9.549e-02  7.387e-02  1.293  0.19624
y.d112       6.275e-02  7.006e-02  0.896  0.37059
y.d113       7.522e-02  6.586e-02  1.142  0.25357
y.d114       6.558e-02  6.144e-02  1.067  0.28591
y.d115       5.124e-02  5.662e-02  0.905  0.36564
y.d116       3.417e-02  5.152e-02  0.663  0.50724
y.d117       5.045e-02  4.584e-02  1.101  0.27121
y.d118       6.985e-02  3.959e-02  1.764  0.07782 .
y.d119       6.759e-02  3.225e-02  2.096  0.03619 *
y.d120       7.109e-02  2.269e-02  3.133  0.00175 **
du           -5.469e-03  1.804e-03  -3.032  0.00246 **
dt           9.216e-06  4.503e-06  2.047  0.04082 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01777 on 1920 degrees of freedom
(21 observations effacées parce que manquantes)
Multiple R-squared:  0.0201,    Adjusted R-squared:  0.007855
F-statistic: 1.641 on 24 and 1920 DF,  p-value: 0.02592

Teststatistic: -10.8386
Critical values: 0.01= -5.57 0.05= -5.08 0.1= -4.82

Potential break point at position: 1373

```

FIGURE 9 – Zivot et Andrews lag= 20.

La figure 9 nous apporte donc que les coefficients δ_1 et δ_2 sont significatifs car les p-values sont inférieur à 0.05. Le modèle both est donc la bonne spécification. Nous pouvons dès lors comparer la statistique t de γ_{20} qui est égale à 3.133, donc $|\gamma_{20}| > 1.6$. Il nous faut ensuite regarder la valeur de la statistique de test qui est de -10.8386, elle est bien inférieur au critère de 5% qui lui est de -5.08. Nous pouvons donc rejeter H_0 , la série n'est pas DS sans changement structurel.

Cependant nous voyons que notre β_1 est significatif car sa p-value est inférieur à 0.05, donc nous pouvons affirmer que, selon le test de Zivot et Andrews, le processus qui génère de notre série est TS avec un changement structurel.

Nous avons également une date de rupture à la position 1373 qui correspond au 10 aout 2015, nous l'affichons en figure 10.

```

> Rupture = dates_rte[1373]
> Rupture
[1] "2015-08-10"

```

FIGURE 10 – Date de rupture Zivot et Andrews

3. Voir figure pour le résultat avec la procédure MAIC

Nous avons également affiché en figure 11 le graphique illustrant la date de rupture le 10 aout 2015.

La statistique t qui test $H_0 : \rho = 1$ pour cette date là est bien la plus petite donc nous avons bien la confirmation que c'est une date de rupture.

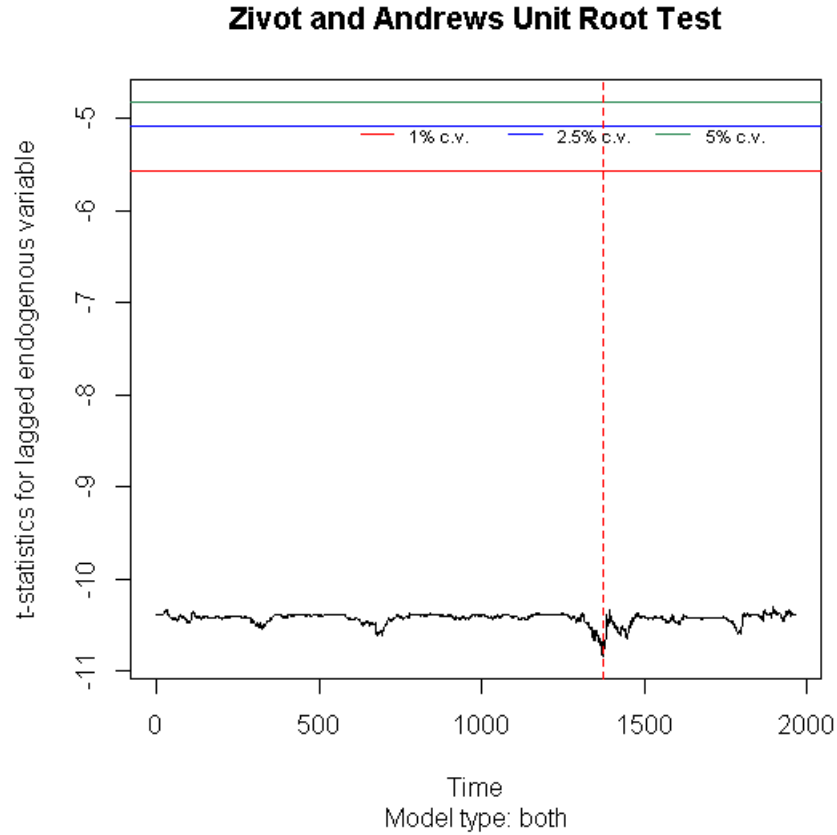


FIGURE 11 – Plot Zivot et Andrews

On pourrait donc être tenté de croire que notre série est TS. Néanmoins, il nous faut prendre en compte le fait que nous puissions avoir une série DS avec changement structurel.

Les conclusions que nous avons tirées pour ce test-ci ne sont donc pas forcément valide, ce problème soulevé fera l'objet de notre prochain test qui est celui de Lee et Strazicich.

4.4 Test de Lee et Strazicich (LS)

Sous ce test-ci nous avons 2 hypothèses à tester qui sont :

$$\begin{cases} (H_0) : y_t = \mu_0 + d_1 B_{1_t} + d_2 B_{2_t} + y_{t-1} + v_{1_t} \\ (H_a) : y_t = \mu_1 + \gamma trend_t + d_1 D_{1_t} + d_2 D_{2_t} + v_{2_t} \end{cases}$$

Autrement dit :

$$\begin{cases} (H_0) : \text{Le PGD est DS, présence de racine unitaire avec changement structurel.} \\ (H_a) : \text{Le PGD est TS avec changement structurel.} \end{cases}$$

Sur le test de Zivot et Andrews nous avons δ_1 et δ_2 qui étaient significatif donc nous allons privilégier le modèle "break" dans le test LS. Afin de pouvoir garder la prise en compte de l'autocorrélation des résidus ,le nombre de lag est le même que celui donné par la procédure Top-Down⁴, les résultats pour une date de rupture s'affichent en figure⁵ 12.

```
> myBreaks <- 1
> myLS_test <- ur.ls(y=rte, model = myModel, breaks = myBreaks, lags = myLags, method = "GTOS",pn = 0.1, print.results =
"print" )
[1] -12.08506
[1] "First possible structural break at position: 1744"
[1] "The location of the first break - lambda_1: 0.9 , with the number of total observations: 1966"
Critical values - Break model:
lambda    1%    5%   10%
[1,]    0.1 -5.11 -4.50 -4.21
[2,]    0.2 -5.07 -4.47 -4.20
[3,]    0.3 -5.15 -4.45 -4.18
[4,]    0.4 -5.05 -4.50 -4.18
[5,]    0.5 -5.11 -4.51 -4.17
[1] "Number of lags determined by general-to-specific lag selection: 20"
Runtime:
Time difference of 0.4718326 mins
```

FIGURE 12 – Lee et Strazicich avec break =1

Avec une date de rupture nous regardons le λ_1 qui est égale à 0.9, nous allons donc le remplacer par la valeur qui s'en rapproche le plus au seuil de 5% avec un λ_5 qui vaut 0.5 , nous remarquons que cette valeur est égale à -4.51.

Nous comparons alors cette valeur avec la statistique t qui est elle-même égale à -12.08506, nous constatons que la statistique t est donc inférieur au seuil de risque de 5 %. Nous rejetons alors H_0 et concluons que notre série n'est pas DS avec 1 changement structurel, le PGD qui génère notre série rte n'a pas de racine unitaire.

4. Voir figure 33 pour savoir les résultats avec le critère MAIC

5. Voir le code R pour les résultats avec 2 dates de ruptures

Nous remarquons également que la date de rupture se situe le 12 février 2017 comme affiché en figure 13.

```
> dates_rte[1744]  
[1] "2017-02-16"
```

FIGURE 13 – Date de rupture

5 Asymétrie perte/gain

5.1 Skewness

Désormais, nous souhaitons tester la nullité de la Skewness qui est égale au moment centré d'ordre 3 normalisé de la distribution.

Nous testons :

(i)

$$H_0 = E \left[\left(\frac{X - E(X)}{\sigma_x} \right)^3 \right] = 3$$

contre

(ii)

$$H_a = E \left[\left(\frac{X - E(X)}{\sigma_x} \right)^3 \right] \neq 3$$

Lorsque nous lançons le test Agostino afin de savoir si notre distribution est symétrique ou non, nous voyons les résultats du test en figure 14.

```
> agostino.test(rte)

D'Agostino skewness test

data:  rte
skew = 0.15616, z = 2.82060, p-value = 0.004793
alternative hypothesis: data have a skewness
```

FIGURE 14 – Skewness

Nous observons que la p-value associé au test est égale à 0,004793, elle est donc inférieur à 0.05. Nous pouvons donc interpréter les résultats du test en affirmant que la skewness est positive et significative.

Cela revient à dire que la probabilité de gains sera inférieur à la probabilité de perte mais que les gains seront plus élevés et les pertes (fréquentes) seront plus faibles.

6 Queues de distribution épaisses

6.1 Kurtosis

Nous savons que les queues de distribution des rendements logarithmiques sont souvent plus épaisses que celles d'une loi normale.

Nous testons :

(i)

$$H_0 = E \left[\left(\frac{X - E(X)}{\sigma_x} \right)^4 \right] = 0$$

contre

(ii)

$$H_a = E \left[\left(\frac{X - E(X)}{\sigma_x} \right)^4 \right] \neq 0$$

Les résultats du test seront affichés en figure 15

```
> anscombe.test(rte)

Anscombe-Glynn kurtosis test

data:  rte
kurt = 8.6998, z = 15.6919, p-value < 2.2e-16
alternative hypothesis: kurtosis is not equal to 3
```

FIGURE 15 – Kurtosis

La kurtosis est significative est supérieur à 3, nous rejetons H_0 donc la distribution est leptokurtique.

Nous pouvons donc dire que la queue de distribution est plus épaisse que celle d'une loi normale, nous aurons plus de valeurs extrêmes que dans le cas d'une loi normale.

7 Autocorrélations des carrés des rendements

7.1 ACF

Afin de mesurer le degré d'autocorrélation d'une série on emploie des coefficients d'autocorrélation totales.

On veut vérifier qu'il existe des faibles autocorrélations des rendements si elles existent ainsi que de fortes autocorrélations des carrés des rendements si elles existent.

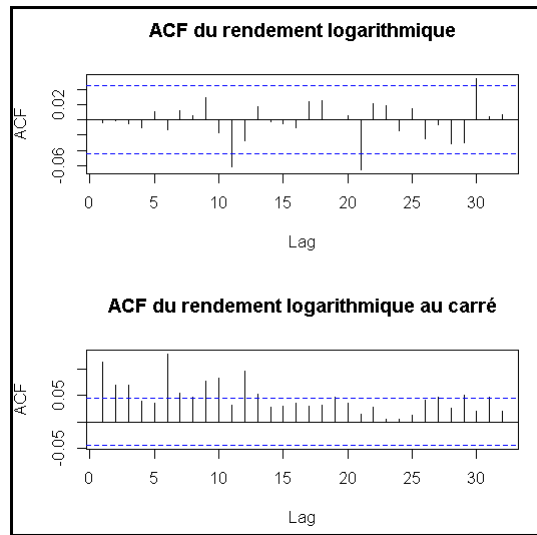


FIGURE 16 – ACF

Nous sommes donc en position de pouvoir affirmer que nous avons la présence d'autocorrélation car nous avons des valeurs qui sortent de notre intervalle de confiance (pointillés bleu).

Cependant le test ACF possède un défaut majeur, si l'autocorrélation à l'ordre 1 est élevé alors celle à l'ordre 2 l'est aussi, ce qui fait que nous ne pouvons pas savoir quel est le véritable degré d'autocorrélation des rendements ainsi que des rendements au carré. Afin de pouvoir être plus précis il nous faut pouvoir lancer le test PACF.

7.2 PACF

Le test PACF nous permet de voir le degré exact d'autocorrélation, il est plus précis que le test ACF.

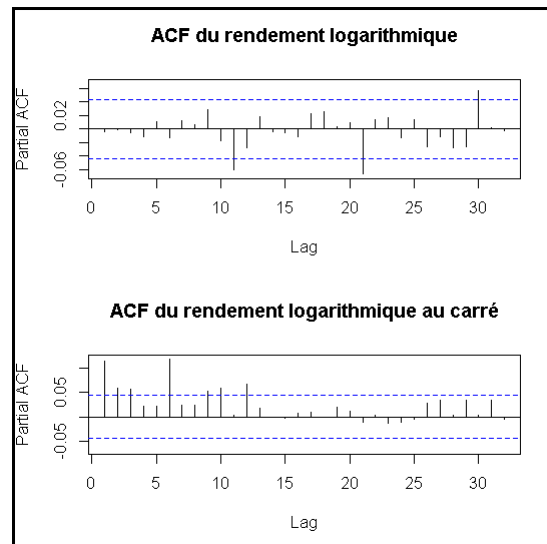


FIGURE 17 – PACF

Nous observons de l'autocorrélation dans les rendements logarithmique à l'ordre 11, 21 et 30.

Également, nous avons la présence d'autocorrélation dans les rendements au carré à l'ordre 1,2,3,6,9,10 et 12.

7.3 Ljung-Box

Nous testons :

(i)

$$H_0 : p(k) = 0$$

pour $k = 1$ jusqu'à K

contre

(ii)

$$H_a : p(k) \neq 0$$

pour une valeur de k comprise entre 1 et K .

La statistique Lundg-Box s'écrit de la manière suivante :

$$Q_K = N(N+2) \sum_{k=1}^K \frac{\hat{\rho}(k)^2}{N-k}$$

Cette statistique suit une loi de Khi-deux.

Nous affichons les résultats de cette statistique sur les rendements au carré en figure ⁶18.

```
> pvaluesrte2
[1] 5.854623e-07 3.277184e-08 1.470724e-09 1.523865e-09 1.812975e-09 1.332268e-15 3.330669e-16 1.110223e-16
[9] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
[17] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
[25] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
[33] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
```

FIGURE 18 – Lundg-box sur les rendements au carré

Nous pouvons donc conclure à la présence d'autocorrélation dans les résidus car les p-value sont toutes inférieure à 0.05. Nous rejetons donc l'hypothèse nulle H_0 d'absence d'autocorrélation dans les résidus.

Nous pouvons donc affirmer avec ces trois tests (ACF,PACF,Lundj-box) que nous possédons bien de l'autocorrélation dans la série rte ainsi que rte au carré.

Nous devons dès lors modéliser cette autocorrélation grâce au modèle ARMA.

6. Voir figure annexe pour les rendements normaux

8 ARMA(P,Q)

Le modèle ARMA nous permet de pouvoir modéliser cette autocorrélation présente dans la série rte.

Pour ce faire, nous devons tout d'abord déterminer le P et le Q du modèle grâce à la procédure `eacf()` présent en figure 19.

```
> eacf(rte)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 0 0 0 0 0 0 0 0 0 0 x 0 0 0
1 x 0 0 0 0 0 0 0 0 0 0 0 x 0 0
2 x x 0 0 0 0 0 0 0 0 0 0 0 0
3 x x x 0 0 0 0 0 0 0 0 0 0 0
4 x x x x 0 0 0 0 0 0 0 0 0 0
5 x x 0 x x 0 0 0 0 0 0 0 0 0
6 x x x 0 x x 0 0 0 0 0 0 0 0
7 x x x x 0 x x 0 0 0 0 0 0
```

FIGURE 19 – EACF

Nous avons fait le choix de partir sur un modèle de type (0,11) car les autres modèles disposaient de coefficients non significatifs ou bien présentaient des valeurs manquantes.⁷

7. Voir en figure les différents modèles estimés.

8.1 ~~ARIMA~~^{MA}(0,11)

Nous partons donc de ce modèle-ci, nous devons donc retirer les ma non significatifs un à un puis, si nécessaire, enlever l'intercept jusqu'à tant obtenir un modèle avec uniquement des coefficients significatifs.

Les résultats de départ sont affichés en figure 20

```
z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ma1      -0.00762895  0.02249752 -0.3391 0.734533
ma2       0.00160890  0.02253363  0.0714 0.943079
ma3      -0.00221158  0.02247485 -0.0984 0.921613
ma4      -0.01217102  0.02249626 -0.5410 0.588491
ma5       0.00952497  0.02265153  0.4205 0.674120
ma6      -0.01122904  0.02295151 -0.4893 0.624664
ma7       0.01690681  0.02257122  0.7490 0.453831
ma8       0.00641503  0.02286148  0.2806 0.779014
ma9       0.02935536  0.02221513  1.3214 0.186364
ma10      -0.02692356  0.02255866 -1.1935 0.232677
ma11      -0.06431685  0.02238380 -2.8734 0.004061 **
intercept  0.00064241  0.00037726  1.7028 0.088601 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

FIGURE 20 – ARIMA (0,11)

Les résultats de fin de procédure sont affichés en figure⁸ 21

```
z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ma11 -0.058083  0.022111 -2.6268 0.008618 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

FIGURE 21 – ARIMA (0,11) après retrait des coefficients non significatifs

Nous arrivons donc à un modèle avec uniquement un seul coefficient significatif MA(11), nous avons donc pris en compte l'autocorrélation présente sur la série rte.

Nous avons également testé plusieurs modèles possible, nous en mettrons quelques uns en annexe. Le reste sera avec le code R correspondant.

8. Se référer au code R afin de voir toutes les étapes de retraits de coefficients

8.2 Résidus

Il nous faut dès lors tester si les ~~résidus~~ ^{calcul} du modèle MA(11) afin de s'assurer qu'ils ne soient pas autocorrélés.

Nous testons alors :

(i)

$$H_0 : E[\epsilon] = 0$$

contre

(ii)

$$H_a : E[\epsilon] \neq 0$$

Les résultats du test sont affichés en figure 21.

```
One Sample t-test
data:  residu
t = 1.7, df = 1965, p-value = 0.08929
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0001048228  0.0014692906
sample estimates:
mean of x
0.0006822339
```

FIGURE 22 – Tests sur les résidus

La p-value est égale à 0.08929, elle est supérieur à 0.05, donc pouvons donc accepter l'hypothèse nulle H_0 , nous devons en conclure que l'espérance des ~~aléas~~ ^{résidus} est nulle.

8.3 Standardisation des résidus

Nous devons ensuite réaliser la standardisation des résidus afin d’obtenir un modèle où les aléas ne sont pas autocorrélés.

Nous testons :

(i) H_0 : Absence d’autocorrélation jusqu’à l’ordre K

contre

(ii) H_a : Présence d’autocorrélation

Nous affichons les résultats en figure 23

```
[1] 0.7598551 0.9505829 0.9850209 0.9846598 0.9900199 0.9916431 0.9918720 0.9963758 0.9675203 0.9558
[12] 0.9495046 0.9533205 0.9711966 0.9811996 0.9867864 0.9837050 0.9757715 0.9843890 0.9893991 0.7024
[23] 0.7345199 0.7562593 0.7805831 0.7596855 0.7987903 0.7527383 0.7223893 0.4692709 0.5191411 0.5688
[34] 0.6411214 0.5200419 0.5625252 0.5850097 0.5824607 0.6250502 0.6118938
```

FIGURE 23 – Tests sur les aléas

Nous observons que toutes les p-values sont supérieures à 0.05, donc les aléas du modèle ARMA(0,11) ne sont pas autocorrélés. Nous pouvons donc retenir le modèle ARMA(0,11).

8.4 BIC

Il nous est également possible de pouvoir intégrer un indicateur BIC afin de tester la robustesse de notre modèle.⁹

Les résultats de BIC du modèle ARMA(0,11) sont donnés en figure 24.

```
> BIC(rgl)
[1] -10245.17
```

FIGURE 24 – BIC ARMA(0,11)

Parmi tous les modèles testés, ARMA(0,11) est celui qui produit les meilleurs résultats.

9. Le BIC sera à comparer avec les autres modèle donnés à partir de la figure 35.

9 Clusters de volatilité

Sur des séries financières nous observons de fortes variations des rendements qui sont généralement suivies de fortes variations.

Ce phénomène amène à un regroupement de cluster de volatilité qui peuvent perturber la validité des tests effectués précédemment.

Désormais il sera alors question de modéliser la volatilité de la série `rte`, nous cherchons à pouvoir détecter la présence d'effet ARCH.

Nous testons :

(i) H_0 : Homoscédasticité Conditionnelle, absence d'effets ARCH.

contre

(ii) H_a : Hétéroscédasticité Conditionnelle, présence d'effets ARCH.

Nous donnons les résultats des tests du lag 1 à 40 sur le tableau ci-dessous¹⁰.

p-value	p-value	p-value	p-value
lag=1	lag=10	lag=30	lag=40
6.007e-07	4.217e-13	6.5e-11	2.516e-07

Toutes les p-values sont inférieure à 0.05, nous pouvons donc affirmer que nous possédons des effets ARCH dans la série `rte`, il nous faut alors utiliser la procédure GARCH (1,1).

10. Se référer au code R donné pour vérifier les résultats

10 Queues épaisses conditionnelles

En ayant détecté la présence d'effet ARCH il nous faut donc appliquer la procédure GARCH (1,1) afin de les capturer sur les résidus de notre ARMA (0,11).

Les p-values du test GARCH sont inférieure à 0.05. Nous obtenons des coefficients significatifs sur le résultat du test GARCH.

a0	a1	b1
p-value	p-value	p-value
1.24e-14	< 2e-16	< 2e-16

Il nous faut dès lors extraire les résidus du modèle GARCH(1,1) afin de vérifier l'absence d'effet ARCH ces résidus-ci. Nous donnons les résultats dans le tableau ci-dessous.

p-value	p-value	p-value	p-value
lag=1	lag=10	lag=30	lag=40
0.3425	0.9507	0.9918	0.999

Le modèle GARCH a réussi à prendre en compte toute l'hétéroscédasticité présente dans le modèle précédent.

Désormais avec un ARMA(0,11) et un GARCH (1,1) nous pouvons affirmer que nous avons modélisé l'autocorrélation ainsi que l'hétéroscédasticité conditionnelle présentes dans le rendement logarithmique. Nous pouvons dire que nous avons terminé l'étude du PDG de la série rte.

Il nous reste désormais à savoir si nos queues de distribution des résidus de notre modèle ARMA-GARCH sont plus épaisses qu'une loi normale. Pour ce faire nous allons réaliser le test d'Anscombe. Nous affichons les résultats sous la forme du tableau ci-dessous.

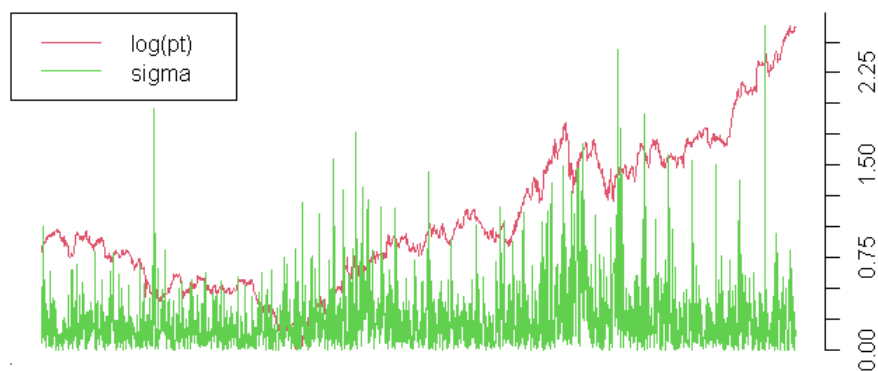
kurt	p-value
9.0543	< 2.2e-16

La kurtosis est bien supérieur à 6 et la p-value est significative. Nous pouvons donc affirmer que notre modèle ARMA-GARCH possède bien des résidus avec des queues de distribution plus épaisses que celle d'une loi normale.

11 Effet de levier

Nous allons désormais étudier des propriétés intéressante des actions. Dans les rendements des actions il est possible d'avoir la présence d'effet de levier. L'effet de levier se présente lorsque les baisses de cours tendent à engendrer une augmentation de la volatilité supérieure à celle induite par une hausse des cours de même ampleur.

Nous affichons le graphique en figure 25



Logarithme de Shiseido journalier et écart-type récursif journalier des rendements

FIGURE 25

Nous voyons donc ici que les baisses de cours entraînent une augmentation de la volatilité encore supérieur à celle engendré par une hausse de cours, nous avons donc ici une présence d'effet de levier.

12 La saisonnalité

12.1 Effet Week-End

Pour finir sur les propriétés intéressantes des séries, il serait adéquat de parler de saisonnalité. Nous avons tout d'abord l'effet week-end qui survient le lundi due à l'accumulation d'information pendant le week-end, cela provoquerait des variations de cours plus importantes que le reste des jours de la semaine.

Nous affichons les résultats en figure 26

	lundi	mardi	mercredi	jeudi	vendredi
moyenne en %	0.2271852	-0.05696796	0.1599501	0.06981862	-0.0640312
écart-type annuel en %	30.7954868	24.57812582	28.1141166	28.16473227	29.5307674
skewness	0.7003537	-1.30956517	0.5525219	0.90783781	-0.6209135
kurtosis	9.4836198	9.84963782	6.6340878	11.24587053	5.6431666

FIGURE 26 – Effet Week-End

Ici nous pouvons voir que l'écart type de lundi est légèrement plus grand que celui des autres jours, nous pourrions alors être tenté de croire qu'il y a présence d'effet week-end dans la série rte.

12.2 Effet Janvier

Nous allons terminer cette étude en parlant de l'effet janvier. Le mois de janvier est un mois favorable aux actions en générale.

En ce qui concerne l'effet janvier, il y aurait plusieurs explications à ce phénomène.

L'explication la plus plausible serait fiscale, les investisseurs souhaitent se débarrasser des actions qu'ils détiennent en fin d'année afin de bénéficier de crédits d'impôts.

Lorsque vient le mois de janvier ils décident alors de les racheter, ce qui mécaniquement ferait augmenter le prix de ces actions.

Nous donnons les résultats du test sur la figure 27.



FIGURE 27 – Effet Janvier

Nous ne sommes pas particulièrement enclin à conclure que les mois de Mars, Janvier et Décembre ont un rendement en moyenne plus élevé que les autres mois.

Nous pouvons affirmer que Décembre est dans les mois les plus volatile mais celui de Mars possède une volatilité légèrement plus élevée car il atteint des valeurs extrêmes par 2 fois.

Également, le mois de janvier ne possède pas la volatilité la plus élevée. Nous ne pouvons pas vraiment conclure qu'il y a un effet janvier dans notre série.

13 Conclusion

Donc en guise de conclusion nous pouvons dire que le modèle ARMA(0,11) est le meilleur modèle.

Le modèle GARCH (1,1) a réussi à prendre en compte toute l'hétéroscédasticité présente dans la série de base.

Le modèle ARMA-GARCH possède bien des résidus avec des queues de distribution plus épaisses que celle d'une loi normale.

La série rte possède un effet de levier, un effet week-end mais ne possède pas d'effet Janvier.

14 Annexes

4. les anciens MIREF, je veux qui vous réalisiez la procédure ci-dessous afin de vérifier si vos résultats sont robustes : vous allez estimer ADF avec lags=la valeur donnée par la formule de Schwert puis séquentiellement retirer les γ non significatifs de tel sorte que le dernier γ doit avoir une valeur de la statistique t , qui soit en valeur absolue > 1.6

FIGURE 28 – Consigne sur l'utilisation de pmax

```
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.096171 -0.008004  0.001120  0.009421  0.126414

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.055607   0.105700  -9.987 < 2e-16 ***
z.diff.lag1   0.049619   0.103266   0.480  0.63093
z.diff.lag2   0.053066   0.100721   0.527  0.59835
z.diff.lag3   0.049869   0.098095   0.508  0.61125
z.diff.lag4   0.038128   0.095242   0.400  0.68896
z.diff.lag5   0.050056   0.092148   0.543  0.58705
z.diff.lag6   0.043018   0.088997   0.483  0.62889
z.diff.lag7   0.059258   0.085754   0.691  0.48963
z.diff.lag8   0.067808   0.082389   0.823  0.41060
z.diff.lag9   0.098439   0.078737   1.250  0.21137
z.diff.lag10  0.081452   0.075112   1.084  0.27832
z.diff.lag11  0.018962   0.071734   0.264  0.79155
z.diff.lag12 -0.005734   0.068291  -0.084  0.93309
z.diff.lag13  0.015048   0.064431   0.234  0.81535
z.diff.lag14  0.013599   0.060326   0.225  0.82167
z.diff.lag15  0.007325   0.055792   0.131  0.89555
z.diff.lag16 -0.001678   0.050943  -0.033  0.97372
z.diff.lag17  0.022491   0.045481   0.495  0.62101
z.diff.lag18  0.049260   0.039397   1.250  0.21132
z.diff.lag19  0.054026   0.032182   1.679  0.09335 .
z.diff.lag20  0.064339   0.022704   2.834  0.00465 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01783 on 1924 degrees of freedom
Multiple R-squared:  0.5083,    Adjusted R-squared:  0.5029
F-statistic: 94.71 on 21 and 1924 DF,  p-value: < 2.2e-16
```

FIGURE 29 – Significativité du pmax-5


```

Coefficients:
(Intercept) -1.622e-03  9.780e-04 -1.659  0.09733 .
y.l1        -3.556e-02  7.009e-02 -0.507  0.61200
trend       3.036e-06  1.237e-06  2.455  0.01417 *
y.d11       2.438e-02  6.586e-02  0.370  0.71130
y.d12       1.677e-02  6.141e-02  0.273  0.78475
y.d13       3.133e-03  5.655e-02  0.055  0.95583
y.d14      -1.631e-02  5.142e-02 -0.317  0.75110
y.d15      -1.241e-02  4.571e-02 -0.272  0.78599
y.d16      -3.049e-02  3.943e-02 -0.773  0.43946
y.d17      -2.449e-02  3.209e-02 -0.763  0.44544
y.d18      -2.369e-02  2.259e-02 -1.049  0.29438
du         -4.867e-03  1.776e-03 -2.741  0.00618 **
dt          7.842e-06  4.470e-06  1.754  0.07952 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01779 on 1944 degrees of freedom
(9 observations effacées parce que manquantes)
Multiple R-squared:  0.007789, Adjusted R-squared:  0.001664
F-statistic: 1.272 on 12 and 1944 DF, p-value: 0.2285

Teststatistic: -14.7752
Critical values: 0.01= -5.57 0.05= -5.08 0.1= -4.82
Potential break point at position: 1373

```

FIGURE 30 – Zivot et Andrews modèle both avec lag= 8.

Pour le test de Zivot et Andrews il s'avère que le modèle both ne convient pas car le coefficient δ_2 n'est pas significatif, nous passons au modèle crash donc.

```

Residuals:
    Min       1Q   Median       3Q      Max
-0.095365 -0.008535  0.000471  0.008926  0.126297

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.010e-03  9.531e-04  -2.109  0.03505 *
y.l1         -2.153e-02  6.967e-02  -0.309  0.75734
trend        3.588e-06  1.197e-06   2.998  0.00275 **
y.d11        1.195e-02  6.551e-02   0.182  0.85529
y.d12        5.963e-03  6.113e-02   0.098  0.92231
y.d13       -5.982e-03  5.634e-02  -0.106  0.91546
y.d14       -2.369e-02  5.128e-02  -0.462  0.64409
y.d15       -1.816e-02  4.562e-02  -0.398  0.69035
y.d16       -3.483e-02  3.937e-02  -0.885  0.37650
y.d17       -2.736e-02  3.206e-02  -0.853  0.39368
y.d18       -2.509e-02  2.259e-02  -1.111  0.26690
du          -3.085e-03  1.457e-03  -2.117  0.03439 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0178 on 1945 degrees of freedom
(9 observations effacées parce que manquantes)
Multiple R-squared:  0.006218, Adjusted R-squared:  0.0005979
F-statistic: 1.106 on 11 and 1945 DF, p-value: 0.3517

Teststatistic: -14.663
Critical values: 0.01= -5.34 0.05= -4.8 0.1= -4.58

Potential break point at position: 1373

```

FIGURE 31 – Zivot et Andrews modèle crash avec lag= 8.

Nous avons ensuite réalisé une procédure Top-Down sans qu'aucun des coefficients se trouve avoir une statistique t supérieur à 1.6, nous affichons tout de même le dernier utilisé en figure 32.¹¹

```
Call:
lm(formula = testmat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.096304 -0.008597  0.000571  0.008943  0.126919

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.952e-02  7.276e-03   2.683  0.00735 **
y.ll        -1.335e-02  3.204e-02  -0.417  0.67706
trend        1.584e-06  7.134e-07   2.220  0.02651 *
y.dll        5.041e-03  2.257e-02   0.223  0.82330
du          -2.050e-02  7.321e-03  -2.799  0.00517 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0178 on 1959 degrees of freedom
(2 observations effacées parce que manquantes)
Multiple R-squared:  0.005929, Adjusted R-squared:  0.0039
F-statistic: 2.921 on 4 and 1959 DF, p-value: 0.0201
```

FIGURE 32 – Zivot et Andrews modèle crash avec lag= 1.

11. Voir le code R pour les autres lags

Nous avons dès lors une prise en compte avec le modèle MAIC. La statistique calculée est inférieure à 4.51, donc nous arrivons aux mêmes conclusions.

Seulement, il y a une erreur dans la prise en compte des lags à mettre en place, il est affiché une NA (valeurs manquantes). Voilà pourquoi il reste plus intéressant de pouvoir réaliser un test LS avec le nombre de lags optimale qui prennent en compte l'autocorrélation comme trouvé dans le test de Dickey-Fuller.

```
> myLS_test <- ur.ls(y=rte, model = myModel, breaks = myBreaks, lags = myLags, method = "GT03",pn = 0.1, print.results =
  "print" )
[1] -44.53101
[1] "First possible structural break at position: 1411"
[1] "The location of the first break - lambda_1: 0.7 , with the number of total observations: 1966"
Critical values - Break model:
      lambda      1%      5%     10%
[1,]    0.1 -5.11 -4.50 -4.21
[2,]    0.2 -5.07 -4.47 -4.20
[3,]    0.3 -5.15 -4.45 -4.18
[4,]    0.4 -5.05 -4.50 -4.18
[5,]    0.5 -5.11 -4.51 -4.17
[1] "Number of lags determined by general-to-specific lag selection: NA"
Runtime:
Time difference of 0.270479 mins
```

FIGURE 33 – Lee et Strazicich avec critère MAIC

Ici nous obtenons un résultat étrange qui ne montre pas d'autocorrélation des résidus pour la série `rte` normal (malgré le fait que les autres test ACF, PACF le montre), nous avons tout de même fait le choix de continuer l'étude avec les tests Arima.

```
> pvaluesrte
[1] 0.8566344 0.9833743 0.9933123 0.9883736 0.9900684 0.9890932 0.9909369 0.9959056 0.9684235 0.9684780 0.4501610
[12] 0.4143350 0.4474114 0.5249507 0.5964338 0.6515847 0.6409863 0.6169640 0.6804702 0.7338802 0.2807969 0.2926357
[23] 0.3133741 0.3455244 0.3760137 0.3693105 0.4172249 0.3663232 0.3285830 0.1590071 0.1899157 0.2216325 0.2547588
[34] 0.2752692 0.1952019 0.2239928 0.2477551 0.2505652 0.2833992 0.2700627
```

FIGURE 34 – Lundg-box sur les rendements normaux

Nous essayons de trouver un meilleur modèle que celui du ARMA(0,11) en comparant la valeur du BIC.

```
> reg2 <- Arima(rte, order=c(1,0,12),fixed=c(0,0,0,0,0,0,0,0,0,0,0,NA,0,0))
> coeftest(reg2)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ma11 -0.058084    0.022111 -2.6269 0.008616 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> BIC(reg2)
[1] -10245.17
```

FIGURE 35 – ARMA (1,12)

Des modèles nous ont amenés à avoir des valeurs manquantes, les résultats sont en figure 36.

```
> reg3 <- Arima(rte, order=c(2,0,2),fixed=c(NA,NA,NA,NA,NA))
> coeftest(reg3)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1    -0.00222852      NaN      NaN      NaN
ar2    -0.00014602      NaN      NaN      NaN
ma1    -0.00192185      NaN      NaN      NaN
ma2    -0.00057994      NaN      NaN      NaN
intercept 0.00064238 0.00040063  1.6034 0.1088

Warning message:
In sqrt(diag(se)) : NaNs produced
> BIC(reg3)
[1] -10210.56
```

FIGURE 36 – ARMA (2,2)

Des modèles commençaient avec des valeurs correctes puis nous avions des valeurs manquantes.

```
> reg4 <-Arima(rte, order=c(3,0,3),fixed=c(NA,NA,NA,NA,NA,NA))
> coeftest(reg4)

z test of coefficients:

              Estimate Std. Error z value Pr(>|z|)
ar1      0.97133739    0.60743192   1.5991   0.1098
ar2      0.35101911    1.06468770   0.3297   0.7416
ar3     -0.74170877    0.58590279  -1.2659   0.2055
ma1     -0.97395176    0.61766240  -1.5768   0.1148
ma2     -0.35615210    1.08392832  -0.3266   0.7425
ma3      0.75738688    0.60215088   1.2578   0.2085
intercept 0.00064278    0.00040950   1.5697   0.1165

> reg4 <-Arima(rte, order=c(3,0,3),fixed=c(NA,NA,NA,NA,0,NA))
> coeftest(reg4)

z test of coefficients:

              Estimate Std. Error z value Pr(>|z|)
ar1     -0.00198385    4.05465221  -0.0005   0.9996
ar2     -0.00038728         NaN         NaN         NaN
ar3     -0.00526124         NaN         NaN         NaN
ma1     -0.00213899    4.05508211  -0.0005   0.9996
ma2     -0.00026808         NaN         NaN         NaN
intercept 0.00064273    0.00039859   1.6125   0.1068

Warning message:
In sqrt(diag(se)) : NaNs produced
~|
```

FIGURE 37 – ARMA (3,3)

Plusieurs cas de figures avec des valeurs manquantes se présentaient à nous.

```
> reg4bis <-Arima(rte, order=c(4,0,4),fixed=c(NA,NA,NA,NA,0,NA,NA,NA))
> coeftest(reg4bis)

z test of coefficients:

              Estimate Std. Error z value Pr(>|z|)
ar1      0.00406455    0.01604353   0.2533   0.8000
ar2     -0.21153427         NaN         NaN         NaN
ar3     -0.32397201         NaN         NaN         NaN
ar4      0.70931353         NaN         NaN         NaN
ma2      0.20213369         NaN         NaN         NaN
ma3      0.31139941    0.02259133  13.7838 <2e-16 ***
ma4     -0.71115781         NaN         NaN         NaN
intercept 0.00064101    0.00039246   1.6333   0.1024
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Warning message:
In sqrt(diag(se)) : NaNs produced
|
```

FIGURE 38 – ARMA (4,4)

Il convient de se référer au code R correspondant afin d'avoir connaissance de l'ensemble des modèles testés.