

---

# Calcul de la Value-at-Risk à l'aide des modèles GARCH

---

BOUDCHICHI Oussama  
Actuariat-Finance

"large changes tend to be followed by large changes,  
of either sign, and small changes tend to be followed by small changes"  
**Benoit B. Mandelbrot (1963)**

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	Préliminaires . . . . .	6
<b>2</b>	<b>Value-at-Risk</b>	<b>8</b>
2.1	Définition et discussion . . . . .	8
<b>3</b>	<b>Propriétés et Analyse des séries temporelles financières</b>	<b>10</b>
3.1	Analyse descriptive des données . . . . .	11
3.2	Normalité des log-rendements ? . . . . .	12
3.3	Faits Stylisés (Stylized Facts) . . . . .	14
3.4	Stationnarité au sens large (au second ordre) : . . . . .	15
3.5	ACF/PACF et Test de "Portmanteau" . . . . .	16
3.5.1	ACF/PACF . . . . .	16
3.5.2	Test de Ljung-Box (Portmanteau) . . . . .	20
<b>4</b>	<b>Modélisation et prévisions des séries temporelles financières</b>	<b>21</b>
4.1	La prévision par lissage exponentiel . . . . .	21
4.1.1	Lissage exponentiel simple . . . . .	21
4.1.2	Lissage exponentiel double (ou de Holt) . . . . .	21
4.1.3	La méthode de Holt-Winters (non-saisonnnière) . . . . .	22
4.1.4	Application à la série des log-rendements du MADEX . . . . .	22
4.2	Modèles linéaires ARIMA . . . . .	23
4.2.1	Rappel . . . . .	23
4.2.2	Choix et estimation des paramètres du modèle . . . . .	24
4.2.3	Analyse des résidus . . . . .	26
4.3	Estimer la volatilité ? . . . . .	28
4.4	Modèles GARCH . . . . .	30
4.4.1	Définition . . . . .	31
4.4.2	Test des effets ARCH : (LM test) . . . . .	31
4.4.3	Estimation des paramètres du modèle . . . . .	32
4.5	Prévisions . . . . .	33
<b>5</b>	<b>Retour à la Value-at-Risk</b>	<b>36</b>
5.1	Méthodes de calcul de la VaR . . . . .	36
5.2	Quelques mots de la théorie des valeurs extrêmes . . . . .	37
5.3	Validation d'une prévision de la VaR (le Backtesting) . . . . .	38
5.3.1	Estimation et prévision par fenêtre glissante (Rolling window estimating and forecasting) . . . . .	38
5.3.2	Backtesting . . . . .	38
5.4	Application à la série des log-rendements du MADEX . . . . .	39
5.4.1	Modélisation . . . . .	41
5.4.2	Résultats du calcul de la VaR (5%) et du Backtest . . . . .	43
<b>6</b>	<b>Conclusion</b>	<b>46</b>
<b>7</b>	<b>Bibliographie</b>	<b>46</b>
<b>8</b>	<b>Code Python</b>	<b>46</b>

# Table des figures

1	Exemple d'une distribution de la perte et de la $\text{VaR}_{0.95}$ Source : Auteur . . . . .	9
2	Evolution du cours de fermeture du MADEX . . . . .	10
3	Log-rendements du MADEX . . . . .	11
4	Densité empirique des log-rendements du MADEX . . . . .	12
5	QQplot des log-rendements du MADEX . . . . .	13
6	ACF et PACF des log-rendements du MADEX . . . . .	17
7	ACF et PACF des log-rendements du MADEX . . . . .	17
8	ACF et PACF de la série des carrés des log-rendements du MADEX . . . . .	18
9	Lag-Plots de la série des valeurs absolues des log-rendements du MADEX . . . . .	19
10	Lag-Plot d'ordre 1 de la série des valeurs absolues des log-rendements du MADEX . . . . .	19
11	Lissage exponentiel simple et prévisions de la série des log-rendements du MADEX . . . . .	22
12	Lissage exponentiel double et prévisions de la série des log-rendements du MADEX . . . . .	23
13	ACF et PACF des résidus du modèle $ARIMA(1, 0, 1)$ . . . . .	24
14	ACF et PACF des résidus du modèle $ARIMA(1, 0, 2)$ . . . . .	24
15	ACF et PACF des résidus du modèle $ARIMA(2, 0, 1)$ . . . . .	25
16	ACF et PACF des résidus du modèle $ARIMA(2, 0, 2)$ . . . . .	25
17	Densité empirique des résidus du modèle $ARIMA(2, 0, 2)$ . . . . .	26
18	ACF et PACF de la série des carrés des résidus du modèle . . . . .	27
19	Graphique de la série des résidus du modèle . . . . .	27
20	Différents Estimateurs de la volatilité . . . . .	29
21	Intervalles de confiances de 0 (moyenne) au niveau 95%, en supposant la normalité conditionnelle . . . . .	30
22	Valeurs Ajustées et prévisions . . . . .	33
23	Volatilité Ajustée et prédite . . . . .	34
24	Valeurs prédites données par le modèle $ARIMA(2, 0, 2)$ contre celles fournies par $ARIMA(2, 0, 2) + GARCH(1, 1)$ . . . . .	35
25	Illustration de l'estimation et la prévision de la VaR par fenêtre glissante pour une fenêtre d'estimation de taille $W_E$ . . . . .	38
26	Série des log-rendements du MADEX de 2016-06-21 au 2021-06-21 . . . . .	40
27	ACF et PACF de la série des log-rendements du MADEX . . . . .	40
28	VaR au niveau de 5% calculée par le modèle $AR(1)$ - $GARCH(1,1)$ et série des dépassements . . . . .	43
29	VaR au niveau de 5% calculée par le modèle EWMA et série des dépassements . . . . .	44
30	VaR au niveau de 5% calculée par les deux modèles et séries des dépassements . . . . .	44

# Liste des tableaux

1	statistiques descriptives des log-rendements . . . . .	11
2	Coefficients d'asymétrie et d'applatissage des log-rendements du MADEX . . . . .	12
3	Test de Ljung-Box pour la série des log-rendements du MADEX . . . . .	20
4	ACF et PACF des résidus du modèles GARCH . . . . .	32
5	Test de Ljung-Box appliqué à la série des résidus du modèle $GARCH(1, 1)$ . . . . .	33
6	Un "Zoom" sur les prévisions . . . . .	34
7	Modèle de la moyenne conditionnelle des log-rendements : $AR(2)$ . . . . .	41
8	Test de Ljung-Box pour les résidus du modèle $AR(2)$ . . . . .	41
9	ARCH-LM test pour les résidus du modèle $AR(2)$ . . . . .	41
10	Test de Kolmogorov-Smirnov pour les résidus du modèle $AR(2)$ . . . . .	41
11	Modèle de la variance conditionnelle des log-rendements (résidus du modèle : $AR(2)$ ) : $GARCH(1, 1)$ . . . . .	42
12	Test de Ljung-Box pour les résidus du modèle $GARCH(1, 1)$ . . . . .	42
13	ARCH-LM test pour les résidus du modèle $GARCH(1, 1)$ . . . . .	42
14	Résultats du modèle $AR(1) - GARCH(1, 1)$ . . . . .	43
15	Résultats du test de Kupiec pour les deux modèles . . . . .	45

# 1 Introduction

À la fin des années 1970 et durant les années 1980, certaines institutions financières des Etats-Unis commençaient à développer des modèles pour mesurer et agréger les risques, pour des fins internes de gestion des risques. Toutefois, Les entreprises devenaient plus complexes et la tâche d'agréger leurs risques en tenant compte de la façon dont elles interagissent, devenait à son tour, de plus en plus difficile, mais aussi de plus en plus importante.

Le système RiskMetrics développé par JP Morgan est le plus connu de ces systèmes. Selon la légende de l'industrie, ce système serait né lorsque le président de JP Morgan, Dennis Weatherstone, a demandé à son personnel de lui remettre un rapport quotidien d'une page indiquant les pertes potentielles liées au portefeuille de la banque au cours des prochaines 24 heures. Ce rapport (le fameux "4 :15 report") devait lui être remis à 4h15 chaque jour, après la clôture des séances du Trading. Afin de combler cette demande, le personnel de JP Morgan a dû développer un système pour mesurer les risques de différentes positions prises par l'ensemble de l'établissement et agréger également ces risques en une seule mesure de risque : La Value-at-Risk (VaR), ou la perte maximale potentielle qui ne devrait être atteinte qu'avec une probabilité donnée sur un horizon temporel donné<sup>1</sup>. Elle a été estimée à partir d'un système basé sur la théorie du portefeuille standard, en utilisant des estimations de la volatilité et des corrélations entre les rendements des différents instruments financiers.

La Value-at-Risk est ensuite devenue, en moins d'une dizaine d'années, une mesure de référence du risque sur les marchés financiers, consacrée notamment par la réglementation prudentielle définie dans le cadre des accords de Bâle II.<sup>2</sup>

## 1.1 Préliminaires

Avant de donner une définition mathématique de la VaR, on commence par introduire un contexte probabiliste pour modéliser les risques financiers et qui va nous servir dans plusieurs définitions par la suite.

On modélise l'incertitude du monde, par un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ . Tel que :  $\Omega$  contient tous les états de natures futures (incertains). Toutes les variables introduites plus bas sont définies sur cet espace.<sup>3</sup>

Considérons un portefeuille noté  $P$ , constitué d'un ensemble d'instruments financiers (actions, obligations, ..) de valeur à un instant  $t$  notée  $V(t)$ .<sup>4</sup> Supposons de plus que  $V(t)$  est observable à l'instant  $t$ . Pour un horizon temporel  $\Delta$ , la perte associée au portefeuille  $P$  dans l'intervalle temporel  $[t, t + \Delta]$  est :

$$L_{t,t+\Delta} := -(V(t + \Delta) - V(t))$$

**Remarque :** Notons que la composition de notre portefeuille reste inchangée durant  $\Delta$ . Cette hypothèse peut être un peu réaliste pour une valeur de  $\Delta$  petite, mais elle devient de plus en plus irréaliste tant que  $\Delta$  augmente.

On mesure par la suite, le temps en terme de  $\Delta$  et on introduit une notation des séries temporelles  $(V_t)_{t \in \mathbb{N}}$ , à l'aide du processus stochastique générateur  $(V(t))_{t \in \mathbb{R}_+}$ , définie par :

$$(\forall t \in \mathbb{N}) \quad V_t := V(t\Delta)$$

En utilisant cette notation, la variable aléatoire perte (entre  $t$  et  $t+1$ ) devient :

$$L_{t+1} := L_{t\Delta, (t+1)\Delta} := -(V_{t+1} - V_t)$$

---

1. 1 jour à l'époque

2. un dispositif prudentiel préparé depuis 1988 par le Comité de Bâle et destiné à mieux appréhender les risques bancaires, pour garantir un niveau minimum de capitaux propres, afin d'assurer la solidarité financière.

3. Cette modélisation n'est pas exhaustive, cependant elle nous sera utile dans les définitions

4. La notation  $V(t)$  est dédiée à la valeur en temps continu ( $t \in \mathbb{R}_+$ ) tandis que la notation  $V_t$  est utilisée pour  $t \in \mathbb{N}$

La distribution de  $L_t$  à un instant  $t$  est appelée la distribution de perte à l'instant  $t$ . On distingue, en général, entre la distribution de perte conditionnelle et la distribution de perte inconditionnelle.

Dans la pratique standard de la gestion quantitative du risque,  $(V_t)$  est modélisé à chaque instant comme :

$$(\forall t \in \mathbb{N}) \quad V_t = f(t, Z_t) \quad (1.1)$$

Tel que :  $f : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}$  est une fonction borélienne et :

$$(\forall t \in \mathbb{N}) \quad Z_t = (Z_{i,t})_{1 \leq i \leq d}$$

pour chaque instant  $t$ ,  $Z_t$  est un vecteur aléatoire constitué des *facteurs de risque* et supposés observables à l'instant  $t$ .

Le choix du processus vectoriel  $(Z_t)_t$  et de  $f$  est un problème de modélisation et dépend du portefeuille en question. Les facteurs de risque les plus utilisés sont les prix logarithmiques des actifs financiers. La représentation de la valeur de  $P$  dans la relation 1.1 est appelé un : *mapping of risks*.

On définit la série des variations des facteurs de risque notée  $(X_t)_{t \in \mathbb{N}}$  comme :

$$(\forall t \in \mathbb{N}^*) \quad X_t := Z_t - Z_{t-1}$$

Ces séries sont les objets les plus intéressants dans l'étude statistique des séries temporelles financières.

En utilisant la relation 1.1, la variable perte se réécrit comme :

$$(\forall t \in \mathbb{N}) \quad L_{t+1} = -(f(t+1, Z_t + X_{t+1}) - f(t, Z_t))$$

**Exemple (Portefeuille d'actions) :** Supposons que le portefeuille  $P$  est constitué de  $d$  actions et  $\lambda_i$  est la quantité de l'action  $i$  présente dans le portefeuille à l'insant initial  $t$ . Pour  $1 \leq i \leq d$ , le process du prix associé à l'action  $i$  est noté  $(S_{i,t})_{t \in \mathbb{N}}$ . Dans la pratique on choisit comme facteurs de risque : les prix logarithmiques. Ainsi :

$$(\forall t \in \mathbb{N}) \quad (\forall 1 \leq i \leq d) \quad Z_{i,t} := \ln(S_{i,t})$$

De ce fait, les variations des facteurs de risque correspondent au log-rendements des actions :

$$(\forall t \in \mathbb{N}) \quad (\forall 1 \leq i \leq d) \quad X_{i,t+1} := \ln(S_{i,t+1}) - \ln(S_{i,t})$$

On a :

$$V_t = \sum_{i=1}^d \lambda_i \exp(Z_{i,t})$$

Donc :

$$L_{t+1} = -(V_{t+1} - V_t) = -\sum_{i=1}^d \lambda_i S_{i,t} (\exp(X_{i,t+1}) - 1) = -V_t \sum_{i=1}^d \omega_{i,t} (\exp(X_{i,t+1}) - 1)$$

Tel que :  $\omega_{i,t} = \frac{\lambda_i S_{i,t}}{V_t}$  est le poids de l'action  $i$  dans le portefeuille à l'insant  $t$ .

Comme déjà mentionné, on distingue entre deux types de distributions de  $L_t$  : conditionnelle et inconditionnelle. la différence entre les deux est fortement liée aux propriétés des séries temporelles de  $(X_t)_{t \in \mathbb{N}}$ .

**Stationnarité au sens strict :** On dit qu'un processus  $(X_t)_{t \in \mathbb{N}}$  est stationnaire au sens stricte si : pour tout  $n, h \in \mathbb{N}^*$  et des instants  $t_1, \dots, t_n$ , on a :

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{=} (X_{t_1+h}, \dots, X_{t_n+h})$$

Supposons que le processus  $(X_t)_{t \in \mathbb{N}}$  est stationnaire au sens strict. Soit  $F_{\mathbf{X}}$  sa fonction de répartition défini sur  $\mathbb{R}^d$ . On muni notre espace probabilisé par la filtration canonique du processus  $(X_t)_{t \in \mathbb{N}}$ ,  $(\mathcal{F}_t^{\mathbf{X}})_t$  définit par :

$$(\forall t \in \mathbb{N}) \quad \mathcal{F}_t^{\mathbf{X}} = \sigma(X_s, s \leq t)$$

La distribution conditionnelle de la perte à un instant donné  $t \in \mathbb{N}^*$ , notée  $F_{L_t | \mathcal{F}_{t-1}^{\mathbf{X}}}$  est définit comme la distribution de  $L_t$  sachant l'information contenue dans  $\mathcal{F}_{t-1}^{\mathbf{X}}$  par :

$$(\forall l \in \mathbb{R}) \quad F_{L_t | \mathcal{F}_{t-1}^{\mathbf{X}}}(l) := \mathbb{P}(L_t \leq l | \mathcal{F}_{t-1}^{\mathbf{X}})$$

D'autre part, la distribution inconditionnelle de  $L_t$  est sa distribution sous la fonction de répartition  $F_{\mathbf{X}}$ .

Ces deux fonctions ne sont pas égales en général<sup>5</sup>. Cependant si  $(X_t)_{t \in \mathbb{N}}$  sont i.i.d, on a l'égalité.

## 2 Value-at-Risk

Considérons un portefeuille constitué d'un nombre d'actifs risqués jusqu'à un horizon temporel  $\Delta$ . Notons par  $F_L$  la fonction de répartition de la perte correspondante. Il est à noter qu'on ne fait pas là une distinction entre la distribution conditionnelle et inconditionnelle de la perte, mais on suppose que le choix de  $F_L$  est justifié par une analyse et elle représente la distribution d'intérêt.

### 2.1 Définition et discussion

Etant donné un niveau de confiance  $\alpha \in ]0, 1[$ , la  $\text{VaR}_\alpha$  est le quantile d'ordre  $\alpha$  de la loi de  $L$ . Ainsi :

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R} : F_L \geq \alpha\}$$

Ou encore :

$$\mathbb{P}(L \leq \text{VaR}_\alpha) = \alpha$$

Si  $F_L$  est inversible, ça sera tout simplement

$$\text{VaR}_\alpha = F_L^{-1}(\alpha)$$

---

5. Un exemple est donné par le modèle GARCH



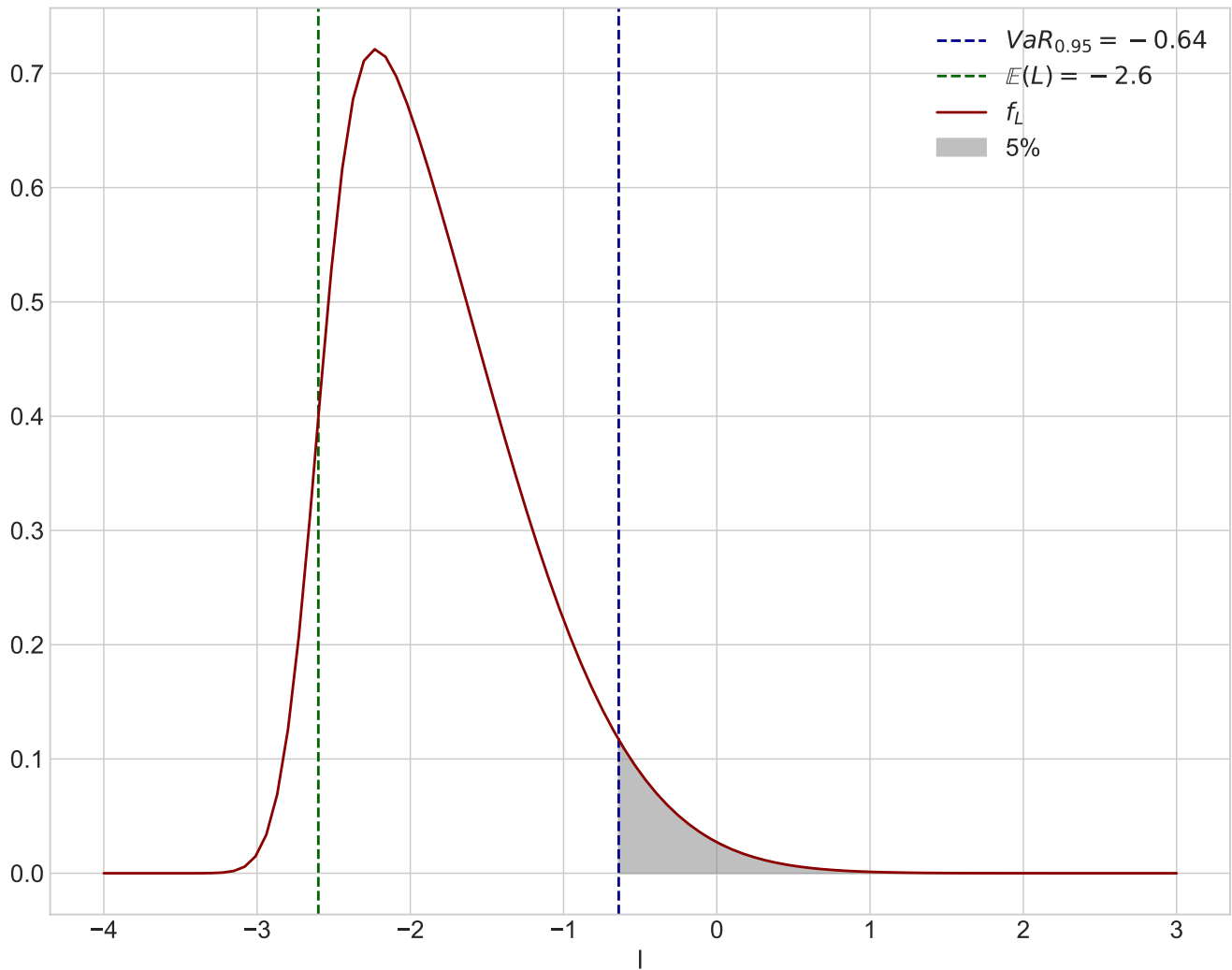


FIGURE 1 – Exemple d’une distribution de la perte et de la  $VaR_{0.95}$   
Source : Auteur

Dans la figure ci-dessus, on constate que la probabilité que  $L$  dépasse au moins  $-0.64$  ( $VaR_{0.95}$ ) est inférieure à 5%.

#### Remarques :

- Notons que par sa définition, la  $VaR_\alpha$  ne donne aucune information sur la sévérité des pertes générées, avec une probabilité inférieure au égale à  $1 - \alpha$ . C’est pour cela qu’il existe d’autres mesures de risque qui prennent ceci en considération, comme : l’Expected Shortfall.
- La distinction entre la distribution conditionnelle et inconditionnelle de la perte, affecte bien l’estimation et la prévision de la VaR puisqu’elle se définit à partir de la fonction de répartition de la perte. C’est pour cela qu’on distingue là aussi, entre la VaR inconditionnelle et la VaR conditionnelle.

**Problématique :** Le calcul de la VaR dépend des propriétés statistiques et surtout distributionnelles du processus des pertes. De ce fait, comment peut-on modéliser la dynamique d’un tel processus ? Comment peut-on calculer la VaR ?

L’étude des pertes générées par un portefeuille d’actifs financiers est équivalent à l’étude des rendements de ces actifs. Ce qui est l’objet de la section suivante.

### 3 Propriétés et Analyse des séries temporelles financières

Afin d'illustrer les propriétés statistiques des données financières, on a choisi l'indice compact MADEX (Moroccan Most Active shares Index) de la Bourse des Valeurs de Casablanca, comme un portefeuille constitué des valeurs les plus actives de la côte, en termes de liquidité. On s'intéressera plus précisément au cours de fermeture de cet indice pendant la période allant de 2020-01-02 jusqu'au 2021-05-31.

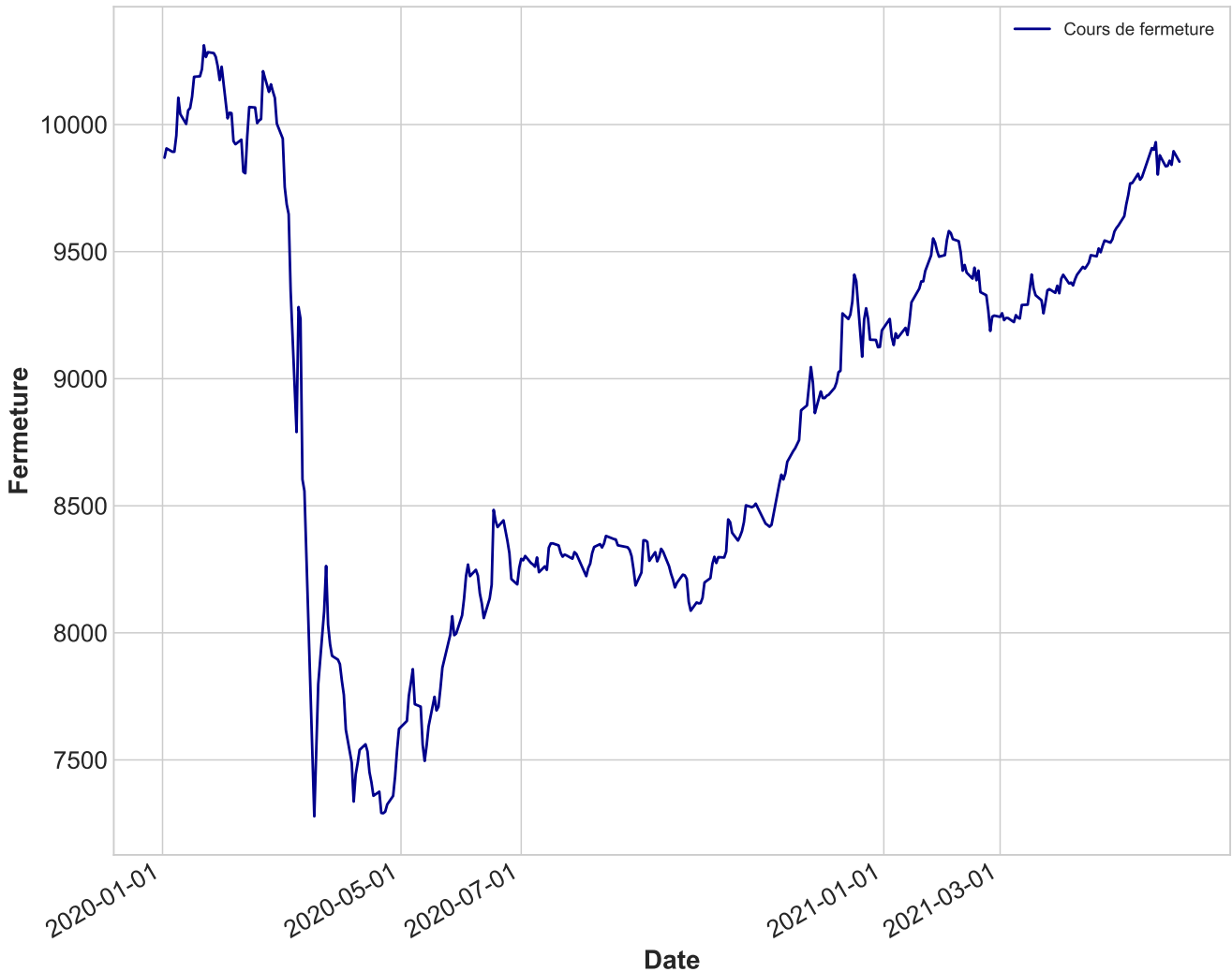


FIGURE 2 – Evolution du cours de fermeture du MADEX

On note par  $P_t$ , le cours de fermeture de l'indice MADEX. Le rendement simple entre  $t$  et  $t+1$  est défini par<sup>6</sup> :

$$R_t := \frac{P_{t+1} - P_t}{P_t}$$

Et le log-rendement entre  $t$  et  $t+1$  sera :

$$r_t := \ln(1 + R_t) := \ln\left(\frac{P_t}{P_{t-1}}\right)$$

On préfère travailler dans la pratique, avec les log-rendements, car les prix sont supposés, généralement, de loi log-normal dans la littérature de la finance<sup>7</sup>. De ce fait,  $r_t$  suit une loi normale<sup>8</sup>. De plus, ce choix est justifié par l'exemple donné dans la section 1.1, dans lequel on a mentionné que

---

6. En absence de Dividendes

7. Le Modèle de Black-Scholes

8. On va tester la validité de cette hypothèse par la suite

le processus *variation des facteurs de risques*, qui est le processus le plus important dans l'analyse des pertes, ne correspond qu'au celui des log-rendements. On pourra ajouter que, généralement, les log-rendements forme une série stationnaire au sens large.

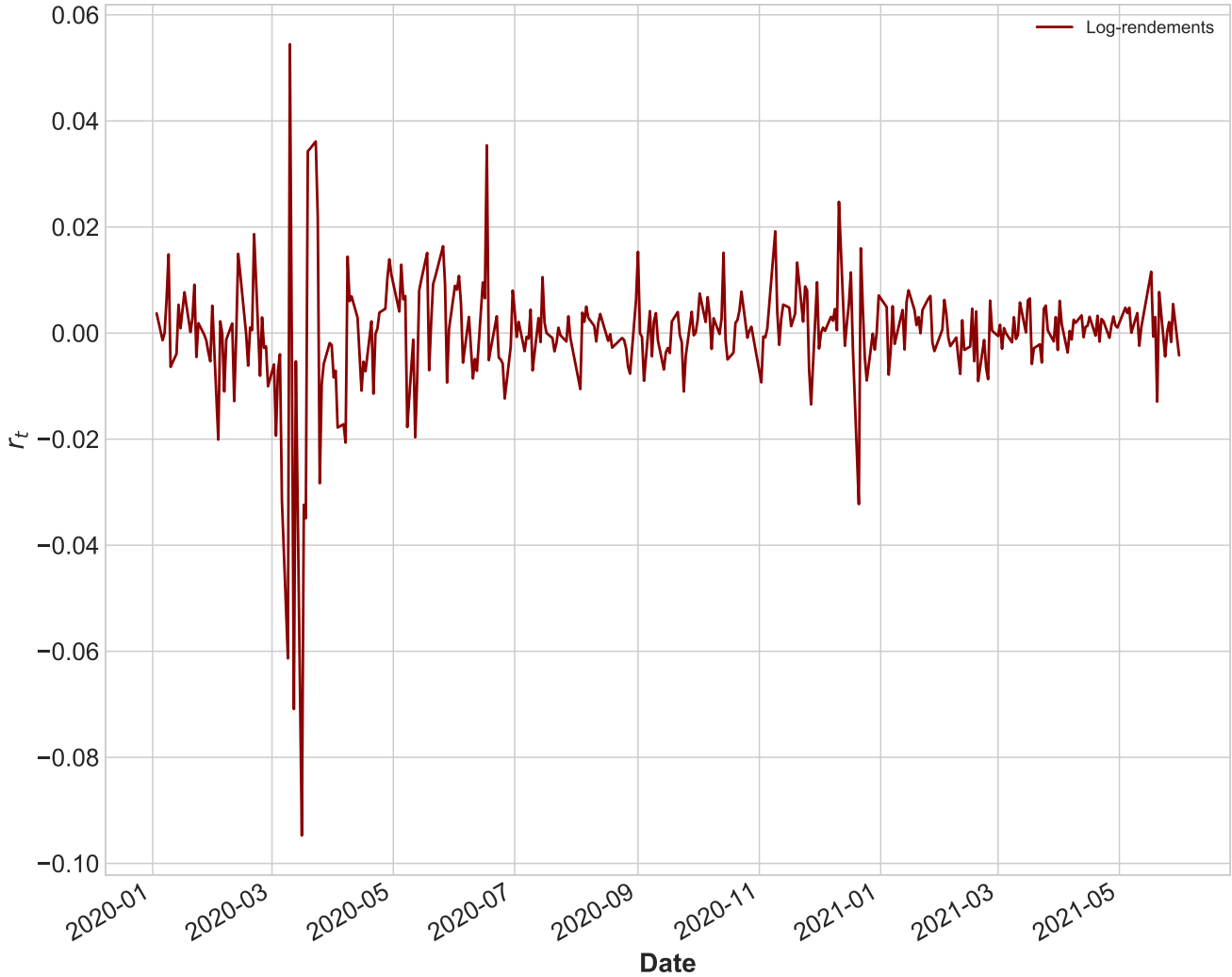


FIGURE 3 – Log-rendements du MADEX

On note qu'on va diviser notre base de données en deux parties, une utilisée dans la modélisation et l'autre sera utilisée pour tester le pouvoir prédictif des modèles

### 3.1 Analyse descriptive des données

N	Moyenne	Ecart-type	Min	$Q_1$	Médiane	$Q_3$	Max
249	-0.000287	0.013279	-0.094733	-0.003998	0.000015	0.004420	0.054460

TABLE 1 – statistiques descriptives des log-rendements

On constate que l'estimation de la moyenne des log-rendements est très proche de zéro, ainsi que 25% des valeurs des log-rendements sont négatives.

**Rappel - Coefficients d'asymétrie et d'applatissment :** Les coefficients d'asymétrie (Skewness) et d'applatissment (Kurtosis) d'une variable aléatoire  $X$ <sup>9</sup> sont définis par :

$$Sk(X) = \frac{\mathbb{E}[(X - \mathbb{E}(X))^3]}{\mathbb{E}[(X - \mathbb{E}(X))^2]^{\frac{3}{2}}} \quad ; \quad K(X) = \frac{\mathbb{E}[(X - \mathbb{E}(X))^4]}{\mathbb{E}[(X - \mathbb{E}(X))^2]^2}$$

9. Sous réserve d'existence des moments d'ordre 3 et 4

Pour  $X \sim \mathcal{N}(0, 1)$  on a :  $Sk(X) = 0$  et  $K(X) = 3$ .

Si on se donne d'un échantillon  $(X_1, \dots, X_n) \stackrel{i.i.d}{\sim} X$ , les estimateurs de ces deux coefficients sont (resp) :

$$\widehat{Sk}(X) = \frac{\frac{1}{n-1} \sum_{j=1}^n (X_i - \bar{X})^3}{\left( \frac{1}{n-1} \sum_{j=1}^n (X_i - \bar{X})^2 \right)^{\frac{3}{2}}} ; \quad \widehat{K}(X) = \frac{\frac{1}{n-1} \sum_{j=1}^n (X_i - \bar{X})^4}{\left( \frac{1}{n-1} \sum_{j=1}^n (X_i - \bar{X})^2 \right)^2}$$

Pour notre série, les estimations sont :

Skewness	Kurtosis
-2.0795	15.4293

TABLE 2 – Coefficients d'asymétrie et d'aplatissement des log-rendements du MADEX

On constate une valeur négative de -2.0705 du coefficient d'asymétrie empirique, ce qui veut dire que la distribution des log-rendements est plus étalée à gauche (vers les valeurs négatives). Pour le kurtosis empirique, on remarque une valeur de 15.4293, soit un excess kurtosis<sup>10</sup> de 12.4293 ! Ce qui veut dire que les queues de la distribution des log-rendements sont beaucoup plus épaisses que celles de la loi normale et que la distribution est leptokurtique. Ce dernier constat, nous pousse à revoir l'hypothèse de la normalité des log-rendements, en commençant par une analyse graphique et en effectuant des tests statistiques.

### 3.2 Normalité des log-rendements ?

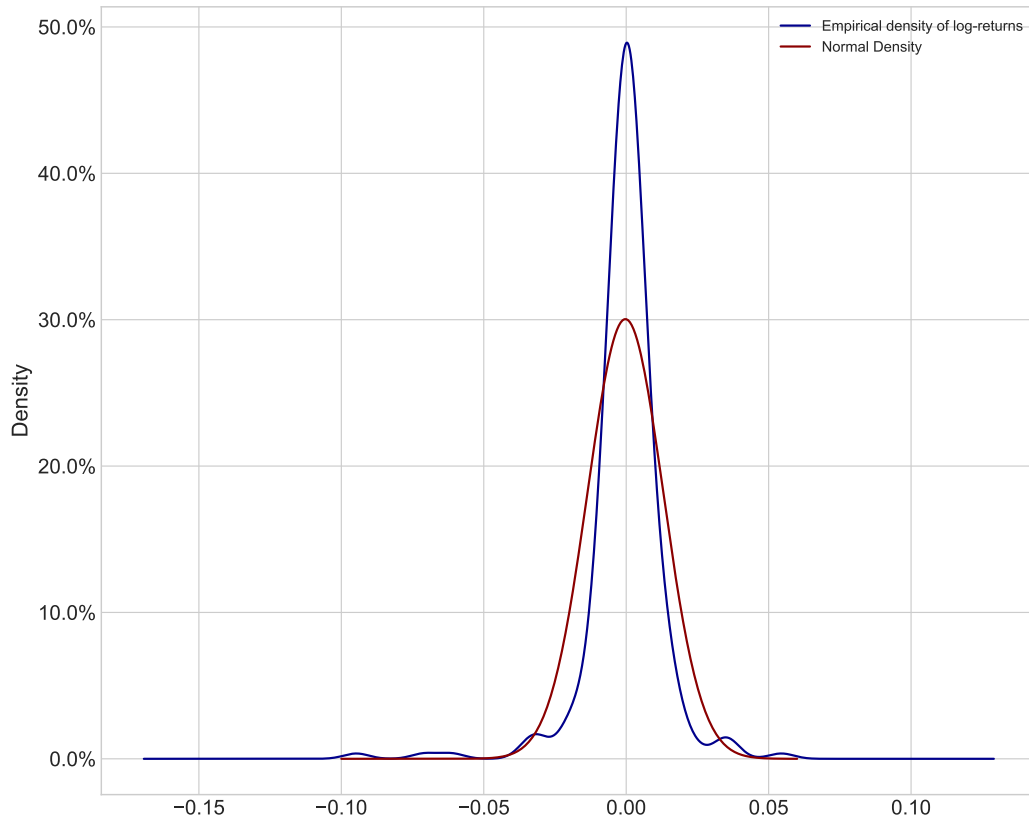


FIGURE 4 – Densité empirique des log-rendements du MADEX

10.  $\widehat{K}(X) - 3$

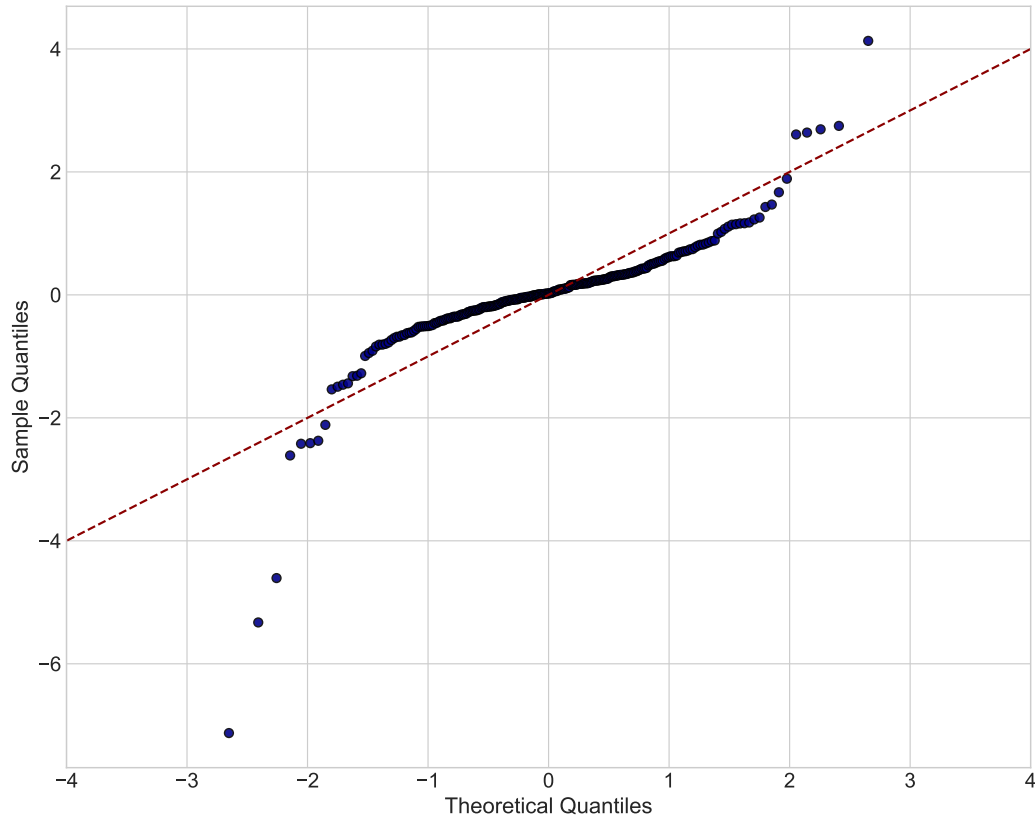


FIGURE 5 – QQplot des log-rendements du MADEX

D'après ces graphiques on constate d'une part, que la densité empirique des log-rendements est clairement leptokurtique et loin d'être similaire de celle de la loi normale. D'autre part, on remarque d'après le QQplot que dès qu'on commence à s'éloigner d'un voisinage de 0, les quantiles empiriques de l'échantillon sont différents des quantiles théoriques de la loi normale, notamment les quantiles négatifs qui correspondent au queue gauche (Perte) de la distribution.

**Test de Shapiro-Wilk :** Etant donné un échantillon :  $(X_1, \dots, X_n) \stackrel{i.i.d}{\sim} X$ . Le test de Shapiro-Wilk teste l'hypothèse suivante :

$$\mathcal{H}_0 : X \sim \mathcal{N}(\mu, \sigma^2)$$

La statistique du test est :

$$W = \frac{\left( \sum_{j=1}^n a_j X_{(j)} \right)^2}{\sum_{j=1}^n (X_i - \bar{X})^2}$$

Tel que :  $(X_{(1)}, \dots, X_{(n)})$  est la statistique d'ordre associée à l'échantillon.  $(a_1, \dots, a_n)$  sont des valeurs tabulées. La loi de W est inconnue, c'est pour cela que les quantiles de cette loi sont estimés à l'aide des simulations Monte Carlo.<sup>11</sup>

D'après la sortie du logiciel, on constate que l'hypothèse de la normalité est rejetée au seuil de 5%.

```
ShapiroResult(statistic = 0.774, pvalue = 3.046e-18)
```

11. Méthode d'estimation des intégrales à l'aide des simulations aléatoires. Elle s'appuie notamment sur la loi forte des grands nombre et est très utilisée dans la Finance mathématique (Génération des scénarios, estimations ...)

**Test de Kolmogorov-Smirnov :** Etant donné un échantillon :  $(X_1, \dots, X_n) \stackrel{i.i.d}{\sim} X$  dont  $F_X$  est continue et  $F_0$  est la fonction de répartition d'une loi continue. On désire tester l'hypothèse suivante contre son alternative :

$$\mathcal{H}_0 : F = F_0$$

On note par  $F_n$  la fonction de répartition empirique de l'échantillon. La statistique du test est définie par :

$$\zeta_n = \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$$

On peut démontrer que sous  $\mathcal{H}_0$  :

$$\zeta_n \xrightarrow{\mathcal{L}} W$$

Tel que  $W$  est une v.a de fonction de répartition définie par :

$$(\forall w > 0) \quad F_W(w) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2 w^2}$$

D'après la sortie du logiciel, on constate que la p-value asymptotique est très inférieure à 0.05, donc là aussi, l'hypothèse de la normalité des log-rendements est significativement rejetée au seuil de 5%

```
KstestResult(statistic = 0.482, pvalue = 1.196e-53)
```

Tous ces résultats nous indiquent qu'il faut rejeter l'hypothèse de la normalité.

### 3.3 Faits Stylisés (Stylized Facts)

Les faits stylisés des séries temporelles financières désignent une collection d'observations empiriques et des inférences déduites de ces données. En général, la majorité des séries journalières<sup>12</sup> des variations des facteurs de risque<sup>13</sup> (log-rendements des actions, des indices, des taux de change ...) exhibent ces propriétés statistiques.

On peut citer :

- Les séries des rendements (log-rendement) **ne sont pas i.i.d** et exhibent une très **faible auto-corrélation**.
- Les séries de la valeur absolue ou le carré des rendements présentent de fortes auto-corrélations.
- Conditionnellement au passé, les rendements espérés sont très proches de zéro.
- Les Queues des distributions des rendements sont épaisses, de ce fait leur décroissance et généralement plus faibles que dans le cas gaussien (en  $\exp(\frac{-x^2}{2})$ ). On parle de distributions **leptokurtiques**.
- **Volatility Clustering** : empiriquement, à court terme, de fortes variations tendent à être suivies par d'autres grandes variations (*Regroupement des extrêmes*).
- **Effet de levier** : cette propriété, notée par Brock en 1976 repose sur l'observation du fait qu'il existe une asymétrie entre l'impact des valeurs passées positives et des valeurs passées négatives. Les baisses de cours tendent à provoquer un accroissement de la volatilité supérieure à celui induit par une hausse du cours de même amplitude.
- **Time patterns** : La volatilité des rendements tend à augmenter sur des périodes bien précises de l'année (*January effect, Week-end effect ...*), on parle d'accumulation d'information.

On va par la suite, observer et tester la présence de quelques faits stylisés dans notre série des log-rendements.

12. Ils s'appliquent dans certains cas aux séries hebdomadaires, mensuelles, ..

13. notés  $(X_t)_{t \in \mathbb{N}}$  dans la section 1.1

### 3.4 Stationnarité au sens large (au second ordre) :

Un processus  $(X_t)_{t \in \mathbb{N}}$  est stationnaire au sens large si :

- $(\forall t \in \mathbb{N}) \quad \mathbb{E}(X_t^2) < +\infty$
- $(\forall t \in \mathbb{N}) \quad \mathbb{E}(X_t) = \mu$ , constante indépendante du temps
- $(\forall t, h \in \mathbb{N}) \quad \text{cov}(X_t, X_{t+h}) = \mathbb{E}[(X_t - \mu)(X_{t+h} - \mu)] = \gamma(h)$ , indépendante de  $t$ .

#### Remarques :

- Une conséquence est que la variance  $V(X_t) = \gamma(0)$  est constante et indépendante du temps.
- la fonction  $\gamma$  est appelé *fonction d'autocovariance*.
- la fonction définie par :  $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$  est appelée *fonction d'autocorrélation*

On confondra par la suite la stationnarité et la stationnarité au sens large.

**Le Test de Dickey & Fuller augmenté :** Soit  $(r_t)_{t \in \mathbb{N}}$  une série temporelle, ce test permet de tester l'hypothèse :

$$\mathcal{H}_0 : (r_t)_{t \in \mathbb{N}} \text{ est intégré d'ordre au moins 1} \quad \mathcal{H}_1 : (r_t)_{t \in \mathbb{N}} \text{ suit un modèle stationnaire}^{14}$$

Le test se fait, brièvement, suivant la stratégie suivante (en commençant par le troisième modèle). Si on dispose de T observations de cette série : .

- (1) On suppose que  $\Delta r_t = \phi r_{t-1} + \sum_{j=1}^p \alpha_j \Delta r_{t-j} + \varepsilon_t$  et on teste  $\mathcal{H}_0 : \phi = 0$
- (2) On suppose que  $\Delta r_t = \alpha + \phi r_{t-1} + \sum_{j=1}^p \alpha_j \Delta r_{t-j} + \varepsilon_t$  et on teste  $\mathcal{H}_0 : \phi = 0$  et  $\alpha = 0$
- (3) On suppose que  $\Delta r_t = \alpha + \beta t + \phi r_{t-1} + \sum_{j=1}^p \alpha_j \Delta r_{t-j} + \varepsilon_t$  et on teste  $\mathcal{H}_0 : \phi = 0, \alpha = 0$  et  $\beta = 0$

D'après les sorties du logiciel, on constate que le processus des log-rendements du MADEX est stationnaire au sens large, au seuil de 5%<sup>15</sup>

#### Augmented Dickey-Fuller Results

=====

Test Statistic	-5.316
P-value	0.000
Lags	7

-----

Trend: **Constant and Linear Time Trend**

Critical Values: -4.00 (1%), -3.43 (5%), -3.14 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

#### Augmented Dickey-Fuller Results

=====

Test Statistic	-5.043
P-value	0.000
Lags	7

-----

---

14. AR(p)

15. Le choix du nombre de retards est basé sur l'AIC

Trend: **Constant**

Critical Values: -3.46 (1%), -2.87 (5%), -2.57 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

#### Augmented Dickey-Fuller Results

=====

Test Statistic            -5.042

P-value                    0.000

Lags                        7

-----

Trend: **No Trend**

Critical Values: -2.58 (1%), -1.94 (5%), -1.62 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

**Bruit blanc faible :** On appelle bruit blanc faible de variance  $\sigma^2$  et on note  $WWN(0, \sigma^2)$ , tout processus  $(\varepsilon_t)_{t \in \mathbb{N}}$  stationnaire dont les auto-covariances sont tous nulles :  $\forall h \neq 0, \gamma(h) = 0$

### 3.5 ACF/PACF et Test de "Portmanteau"

#### 3.5.1 ACF/PACF

Considérons  $T$  observations d'une série temporelle,  $(r_t)_{t \in \llbracket 1, T \rrbracket}$ . On définit la **fonction d'autocorrélation empirique** par :

$$(\forall h \in \llbracket 1, T-1 \rrbracket) \quad \hat{\rho}(h) = \frac{\frac{1}{T-h} \sum_{t=1}^{T-h} (r_t - \bar{r}_T) (r_{t-h} - \bar{r}_T)}{\frac{1}{T} \sum_{t=1}^T (r_t - \bar{r}_T)^2}$$

Et pour une série stationnaire,  $(r_t)_{t \in \mathbb{N}}$ , on définit la fonction d'autocorrélation partielle par :

$$(\forall h \in \mathbb{N}) \quad \psi(h) = \text{corr}(\widehat{X}_t, \widehat{X}_{t-h})$$

Avec :<sup>16</sup>

$$\widehat{X}_t = X_t - \mathbb{E}\mathbb{L}(X_t | X_{t-1}, \dots, X_{t-h+1}) \quad ; \quad \widehat{X}_{t-h} = X_{t-h} - \mathbb{E}\mathbb{L}(X_{t-h} | X_{t-1}, \dots, X_{t-h+1})$$

L'estimateur de cette dernière noté  $\hat{\psi}$ , s'appelle **fonction d'autocorrélation partielle empirique**

Les graphiques de  $\hat{\rho}$  et de  $\hat{\psi}$  s'appellent respectivement *L'autocorrélogramme* (*L'autocorrélogramme partielle*)

---

16.  $\mathbb{E}\mathbb{L}(X_t | X_{t-1}, \dots, X_{t-h+1})$  est l'espérance linéaire, c'est la projection orthogonale dans  $\mathbb{L}^2$  de  $X_t$  sur  $\text{vect}\{\mathbb{I}, X_{t-1}, \dots, X_{t-h+1}\}$



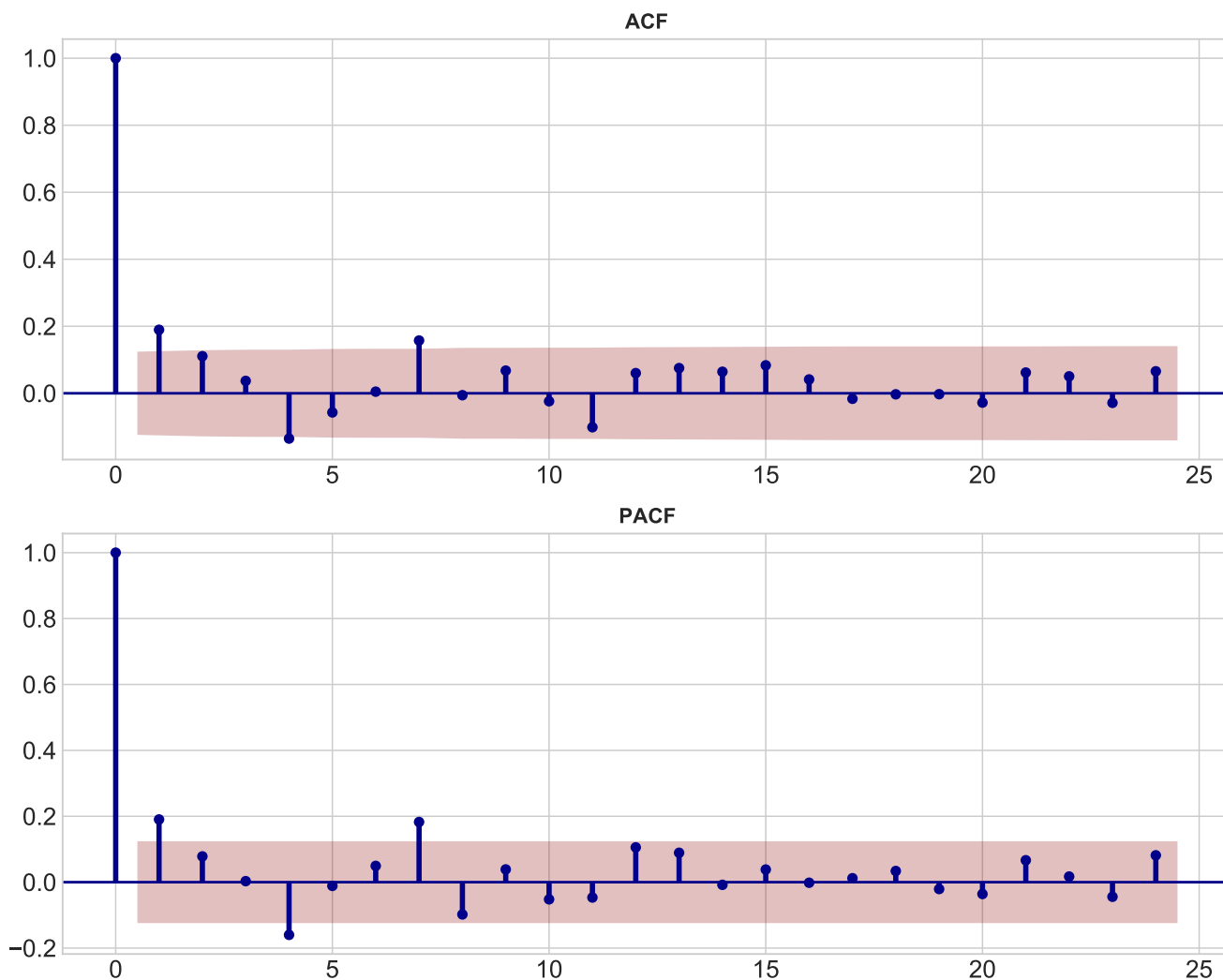


FIGURE 6 – ACF et PACF des log-rendements du MADEX

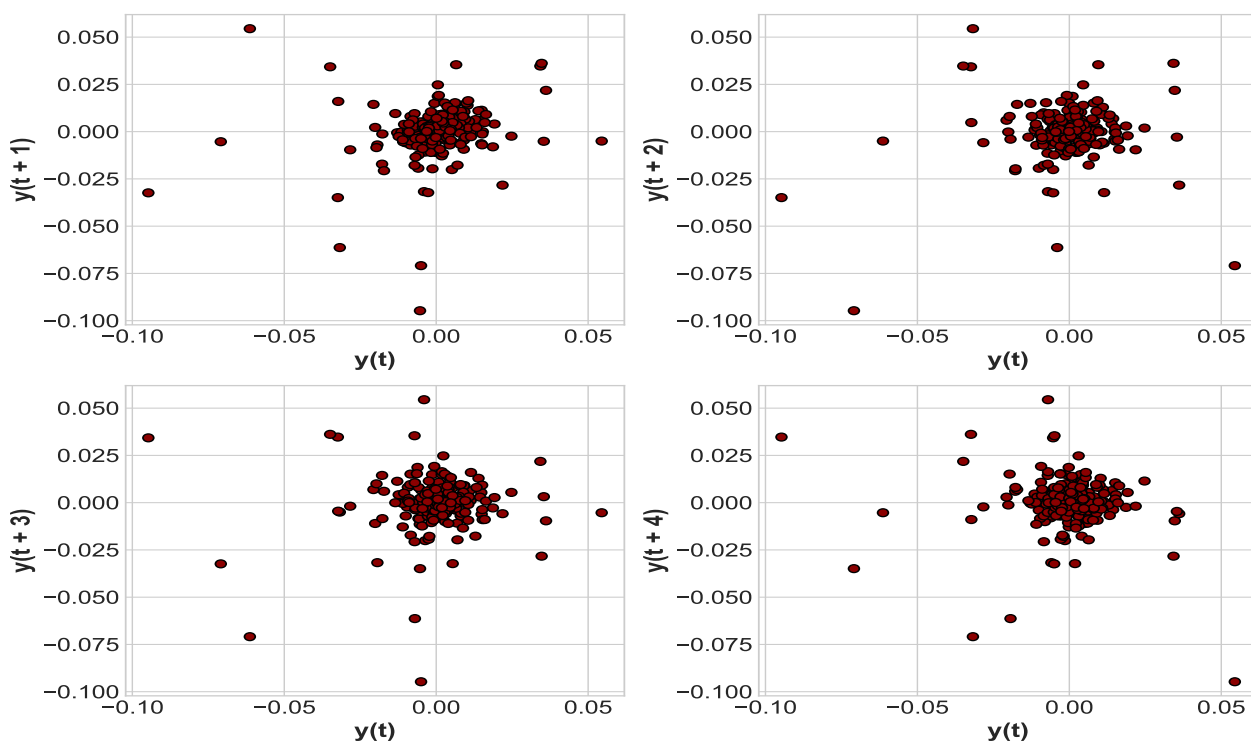


FIGURE 7 – ACF et PACF des log-rendements du MADEX

On remarque dans l'ACF que presque la majorité des autocorrélations, à l'exception de la première, la deuxième et la septième, sont nulles au risque d'erreur de 5%. Le même constat s'applique au PACF. De plus, ceci est également confirmé par les *Lag-plots* jusqu'à l'ordre 4 de la série des log-rendements. Il est à noter que ceci est en accord avec le premier fait stylisé cité auparavant.

En fait, ce résultat est très connu dans les marchés financiers liquides et est souvent cité comme un argument "empirique" de l'hypothèse d'efficience des marchés financiers. Si de telles autocorrélations sont significativement non nulles (pour certains ordres de retard), les agents peuvent les exploiter pour concevoir des stratégies d'investissements à rendement positif sous forme d'opportunités d'arbitrage, c'est ce qu'on appelle : *l'arbitrage statistique* qui regroupe plusieurs méthodes en ce sens, dont la plus célèbre est le : *Pairs Trading*.

Considérant maintenant, l'ACF et le PACF de la série des carrés des log-rendements, ainsi que les Lag-Plots (jusqu'à l'ordre 4) de la série de leurs valeur absolues. On remarque que les carrés des log-rendements exhibent des autocorrélations significativement non nulles, jusqu'à l'ordre 5 ! le même constat s'applique au PACF. Ceci contredit l'hypothèse des log-rendements i.i.d !

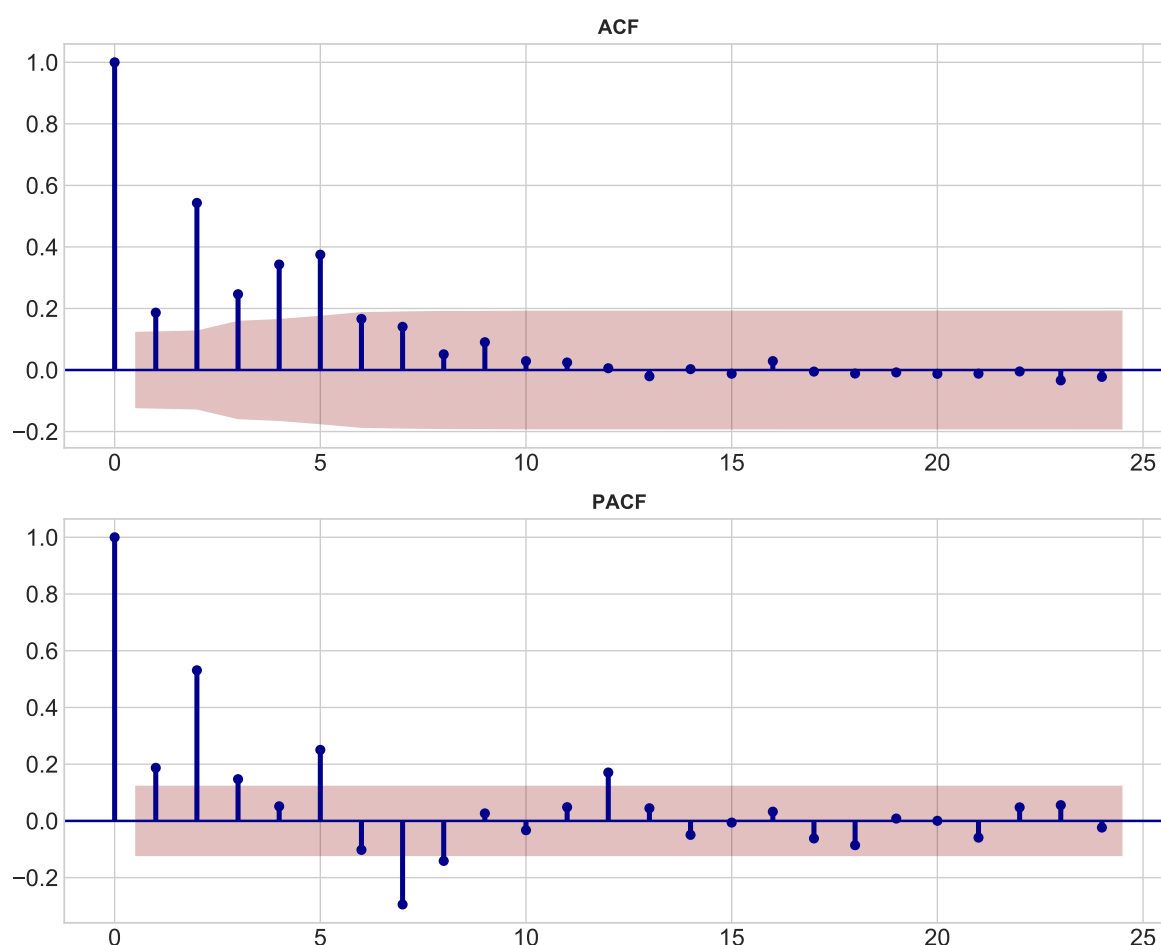


FIGURE 8 – ACF et PACF de la série des carrés des log-rendements du MADEX

Ceci est également observé chez la série des valeurs absolues des log-rendements dans les Lag-plots. Il s'agit d'une signature du phénomène : *Volatility clustering*, observé pour la première fois (1963) par le mathématicien Benoît Mandelbrot. La volatilité se définit comme l'écart type conditionnel de la série des rendements. Cette mesure est censé quantifier l'incertitude associée à l'évolution des rendements. Comme déjà cité, ce phénomène se manifeste par le fait que les grandes variations tendent à être suivies par de grandes variations, peu importe le signe.<sup>17</sup>

17. A court terme

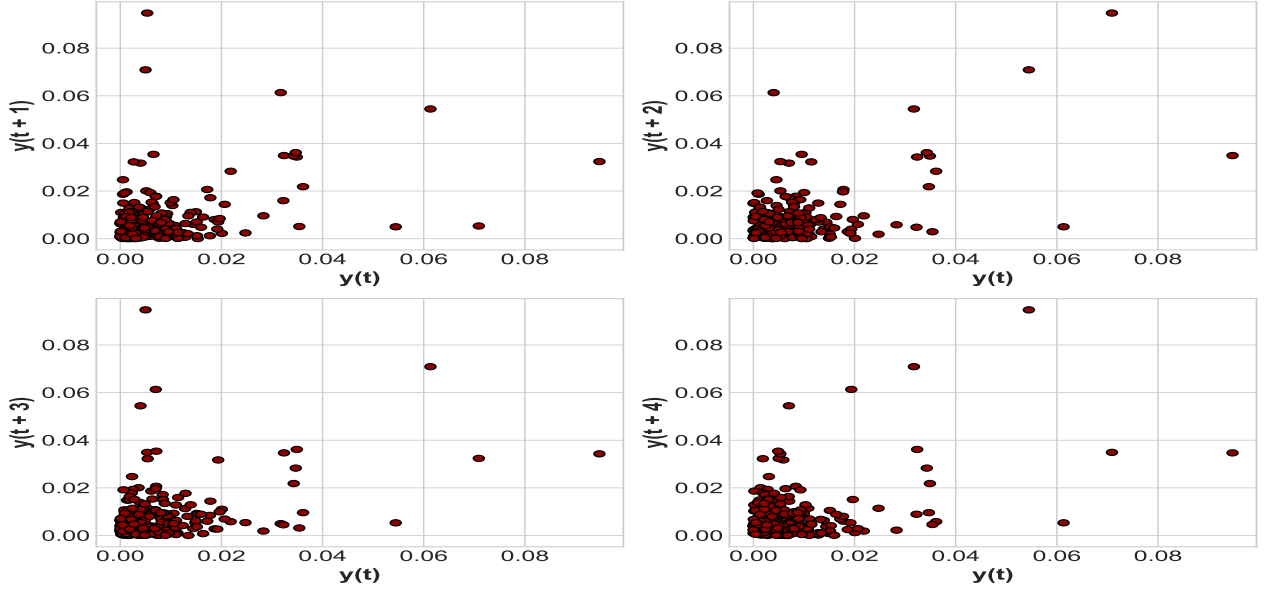


FIGURE 9 – Lag-Plots de la série des valeurs absolues des log-rendements du MADEX

Pour illustrer ce phénomène, revenons au Lag-plot d'ordre 1 de la série des log rendements. Les points sont numérotés selon leur succession temporelle.

46 :  $(-0.0317, -0.0613)$     47 :  $(-0.0613, 0.0545)$     48 :  $(0.0545, -0.0050)$   
 49 :  $(-0.0050, -0.0709)$     50 :  $(-0.0709, -0.0053)$     51 :  $(-0.0053, -0.0947)$   
 52 :  $(-0.0947, -0.0324)$     53 :  $(-0.0324, -0.0349)$     54 :  $(-0.0349, 0.0343)$

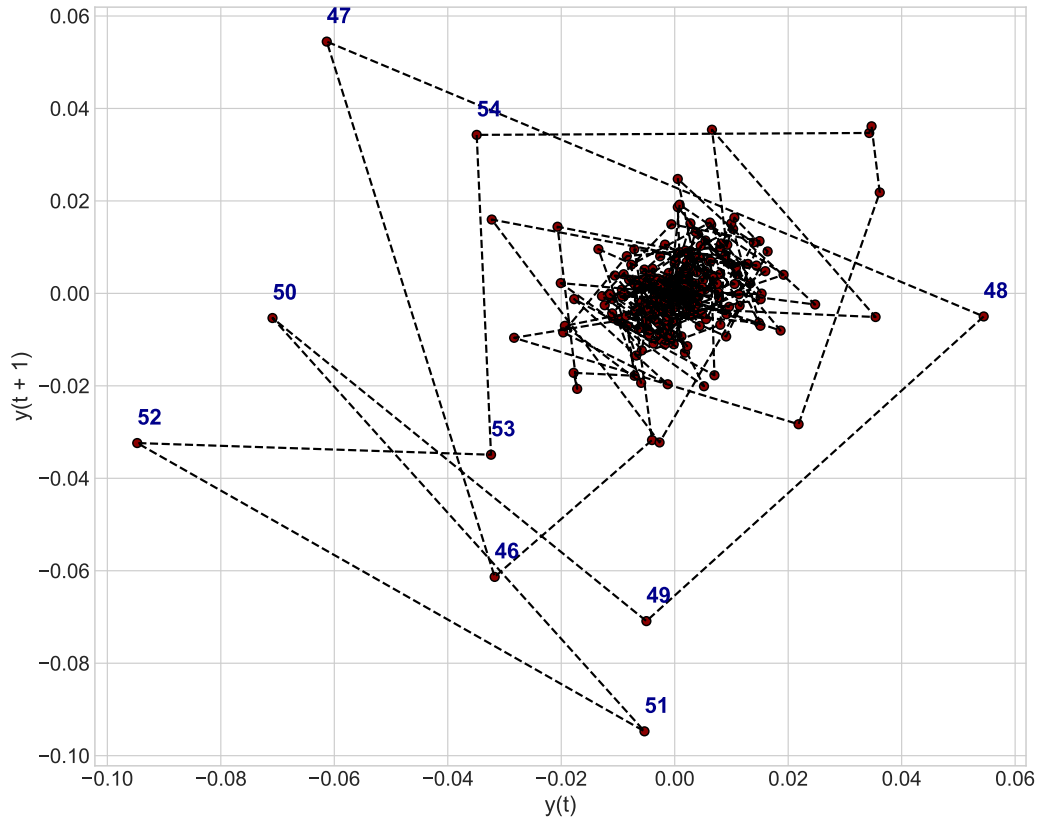


FIGURE 10 – Lag-Plot d'ordre 1 de la série des valeurs absolues des log-rendements du MADEX

On remarque ainsi, que la série des log-rendement a passé de la valeur -0.06 à 0.05, puis de 0.05 à -0.005 ...<sup>18</sup> ce qui illustre ce qu'on a déjà cité. Il est à noter que ceci motive le fait que la volatilité varie dans le temps et est, en quelque sorte, prévisible.

### 3.5.2 Test de Ljung-Box (Portmanteau)

Etant donné  $T$  observations d'une série temporelle,  $(\varepsilon)_{t \in \llbracket 1, T \rrbracket}$ , ce test permet de tester si elles sont des réalisations d'un bruit blanc faible. Plus précisément, il s'agit de tester pour un nombre  $h$  de retards :

$$\mathcal{H}_0 : \rho(1) = \dots = \rho(h) = 0 \quad \mathcal{H}_1 : \exists j \in \llbracket 1, h \rrbracket, \rho(j) \neq 0$$

La statistique du test est la statistique de Ljung-Box, définit par :

$$Q_h = T(T+2) \sum_{j=1}^h \frac{\hat{\rho}(j)}{T-j}$$

Asymptotiquement, sous  $\mathcal{H}_0$  :

$$Q_h \sim \chi_h^2$$

Pour la série des log-rendements du MADEX, on remarque que le coefficient d'autocorrélation d'ordre 1 et 2 sont significativement différents de 0 au seuil de 5%. En plus, il ne s'agit pas d'un bruit blanc au risque d'erreur de 5% . La raison pour laquelle on a effectuée ce test, est le fait que les log-rendements (ou les rendements) ont un ACF similaire à celui d'un bruit blanc. On finit cette partie par une observation en ce sens de B. Mandelbrot : *"arbitrage tends to whiten the spectrum of price changes"*.

Lag	Lb-stat	Lb-pvalue
1	9.055162	0.002620
2	12.156962	0.002292
3	12.504289	0.005841
4	17.167393	0.001793
5	18.003246	0.002942
6	18.009263	0.006209
7	24.417000	0.000962

TABLE 3 – Test de Ljung-Box pour la série des log-rendements du MADEX

---

18. N'oublions pas qu'on est sur une échelle logarithmique

## 4 Modélisation et prévisions des séries temporelles financières

### 4.1 La prévision par lissage exponentiel

Introduites par Holt en 1958, Winters en 1960 et popularisées par Brown en 1963, les méthodes de lissage constituent l'ensemble des techniques empiriques de prévision qui accordent plus ou moins d'importance aux valeurs du passé d'une série.

On ne fait ici, qu'un simple rappel de ces méthodes. Pour plus de détails, le lecteur peut voir : *Time Series and Dynamic Models [1995]. by C. Gouriéroux and A. Monfort*

On dispose de  $T$  observations d'une série temporelle  $(r_t)_{t \in \llbracket 1, T \rrbracket}$ . On souhaite prévoir, à la date  $t \in \llbracket 1, T \rrbracket$ , la valeur de la série, à un horizon quelconque  $h$ . On note  $\hat{r}_t(h)$  cette valeur.

#### 4.1.1 Lissage exponentiel simple

La prévision fournie par la méthode de **lissage exponentiel simple** est :

$$\hat{r}_t(h) = (1 - \beta) \sum_{j=0}^{t-1} \beta^j r_{t-j}$$

Tel que,  $\beta \in ]0, 1[$  est une constante appelée : *constante de lissage*.

On donne un poids d'autant moins important que les observations sont loins, avec une décroissance exponentielle.<sup>19</sup>

#### Remarques

- Si  $\beta$  ne dépend pas de  $h$ ,  $\hat{r}_t(h)$  ne dépend pas de  $h$  et on la note  $\hat{r}_t$  (série lissée à la date  $t$ )
- Méthode adaptative de mise à jour (ordre 1) :  $\hat{r}_t = (1 - \beta)r_t + \beta\hat{r}_{t-1}$
- Le lissage exponentiel simple est adapté à des séries pouvant être ajustées par une constante au voisinage de  $t$

#### 4.1.2 Lissage exponentiel double (ou de Holt)

Le principe de **lissage exponentiel double**, permet d'approcher la série  $(r_t)_{t \in \llbracket 1, T \rrbracket}$ , à un instant donné  $t$ , par  $(X_t)_{t \in \llbracket 1, T \rrbracket}$  où :<sup>20</sup>

$$X_j = \alpha_t + (j - t)\beta_t \quad (4.1)$$

La prévision à un horizon  $h$  s'écrit :

$$\hat{r}_t(h) = \hat{\alpha}_t + h\hat{\beta}_t$$

Tel que, pour tout instant  $t$  :

$$(\hat{\alpha}_t, \hat{\beta}_t) = \underset{(\alpha_t, \beta_t)}{\operatorname{argmin}} \left\{ \sum_{j=1}^{t-1} \gamma^j [r_{t-j} - (\alpha_t + \beta_t j)]^2 \right\}$$

Avec,  $\gamma \in ]0, 1[$  est une constante de lissage.

**Remarque :** Vu 4.1.2, Cette méthode est mieux adaptée à des séries montrant une tendance localement linéaire.

---

19. Si  $\beta$  est proche de 1, on prend compte de tout le passé. Si elle est proche de 0, on prend en compte les valeurs récentes plus que celles du passé

20. au voisinage de  $t$

### 4.1.3 La méthode de Holt-Winters (non-saisonnière)

De la même façon que pour le lissage exponentiel double, l'ajustement se fait de façon linéaire au voisinage de  $t$ , la différence se trouve au niveau des formules de mise à jour d'estimation des coefficients :

$$\begin{cases} \hat{\alpha}_{t+1} = (1 - \lambda)r_t + \lambda(\hat{\alpha}_t + \hat{\beta}_t) & \text{où } \lambda \in ]0, 1[ \\ \hat{\beta}_{t+1} = (1 - \mu)(\hat{\alpha}_{t+1} - \hat{\alpha}_t) + \mu\hat{\beta}_t & \text{où } \mu \in ]0, 1[ \end{cases}$$

### 4.1.4 Application à la série des log-rendements du MADEX



FIGURE 11 – Lissage exponentiel simple et prévisions de la série des log-rendements du MADEX

On remarque que le modèle n'arrive pas à capter les fluctuations des log-rendements, en plus les prévisions sont constantes quelque soit l'horizon temporel considéré. L'erreur de prévision empirique de ce modèle est :

$$\text{RMSE} = 0.004185$$

La même remarque s'applique au lissage exponentiel double. Chose qui était prévisible, puisqu'il suppose qu'au voisinage de chaque instant, la série exhibe un caractère localement linéaire. L'erreur de prévision empirique de ce modèle est légèrement supérieur à celui du modèle précédent. La deuxième remarque porte sur le fait que les intervalles de prévision au niveau de 95%<sup>21</sup> du deuxième modèle sont plus larges que ceux du premier modèle.

$$\text{RMSE} = 0.004959$$

21. Il faut bien sur, vérifier l'hypothèse de la normalité, de l'absence de l'autocorrélation et de l'homoscédasticité des résidus des deux modèles.



FIGURE 12 – Lissage exponentiel double et prévisions de la série des log-rendements du MADEX

## 4.2 Modèles linéaires ARIMA

### 4.2.1 Rappel

Un processus  $(r_t)$  est un processus  $ARIMA(p, d, q)$  s'il vérifie une équation du type :

$$(\mathbb{I} - L)^d r_t = \sum_{j=1}^p \phi_j r_{t-j} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$

où :  $\phi_p \neq 0$ ,  $\theta_q \neq 0$ , les racines des polynômes auto-régressif et moyenne-mobile sont tous de module supérieur strictement à 1<sup>22</sup> et les conditions initiales :<sup>23</sup>

$$CI = \{r_{-1}, \dots, r_{-p-d}, \varepsilon_{-1}, \dots, \varepsilon_{-q}\}$$

sont non-corrélées avec  $(\varepsilon_t)_{t \in \mathbb{N}}$ .  $(\varepsilon_t)_{t \geq -q}$  est un SWN(0,  $\sigma^2$ ).

**Remarque :** Un processus  $(r_t)$  est  $ARIMA(p, d, q)$  si :  $((\mathbb{I} - L)^d r_t)_t$  **converge** vers un processus  $ARMA(p, q)$ .

On suppose dans un premier temps que la série des log-rendements du MADEX peut être modélisée par un processus  $ARIMA(p, d, q)$ . Rappelons qu'on a déjà trouvé qu'il s'agit d'un processus stationnaire et ne contient pas une racine unité, ainsi  $d = 0$ . Une deuxième remarque se porte sur l'ACF et le PACF de la série dans lesquelles on a observé que la série montre des autocorrélations (simples et partielles) jusqu'à l'ordre 2.

22. Pour l'unicité et la stationnarité de la représentation  $ARMA(p, q)$

23. Il est à noter que cette condition est nécessaire pour les processus  $ARIMA(p, d, q)$  à la différence des processus  $ARMA(p, q)$  qui peuvent être définie sur  $\mathbb{Z}$ . Voir : *Time Series and Dynamic Models [1995]. by C. Gouriéroux and A. Monfort*

### 4.2.2 Choix et estimation des paramètres du modèle

On commence par l'estimation des paramètres des modèles candidats, calculer leurs AIC et présenter l'ACF et le PACF des résidus de chaque modèle afin de choisir l'ordre de retards qu'on va utiliser dans le test de Ljung-Box.

```
model : ARIMA ( 1 ,0, 1 ) ; AIC : -1448.6115  
model : ARIMA ( 1 ,0, 2 ) ; AIC : -1447.6899  
model : ARIMA ( 2 ,0, 1 ) ; AIC : -1447.0775  
model : ARIMA ( 2 ,0, 2 ) ; AIC : -1455.6938
```

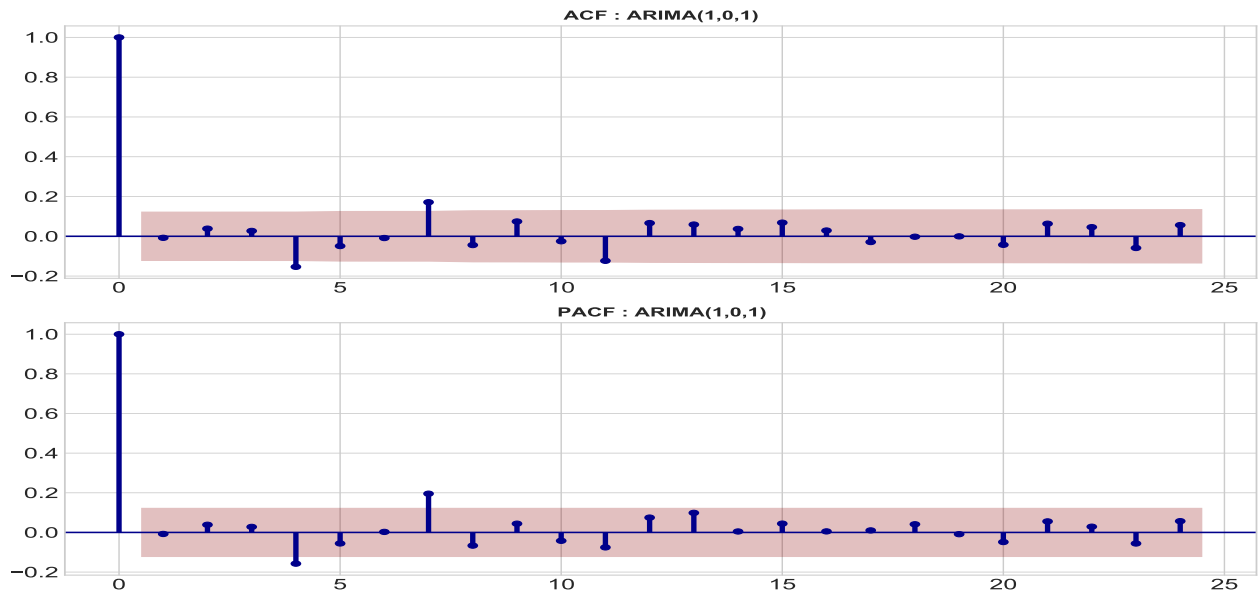


FIGURE 13 – ACF et PACF des résidus du modèle  $ARIMA(1,0,1)$

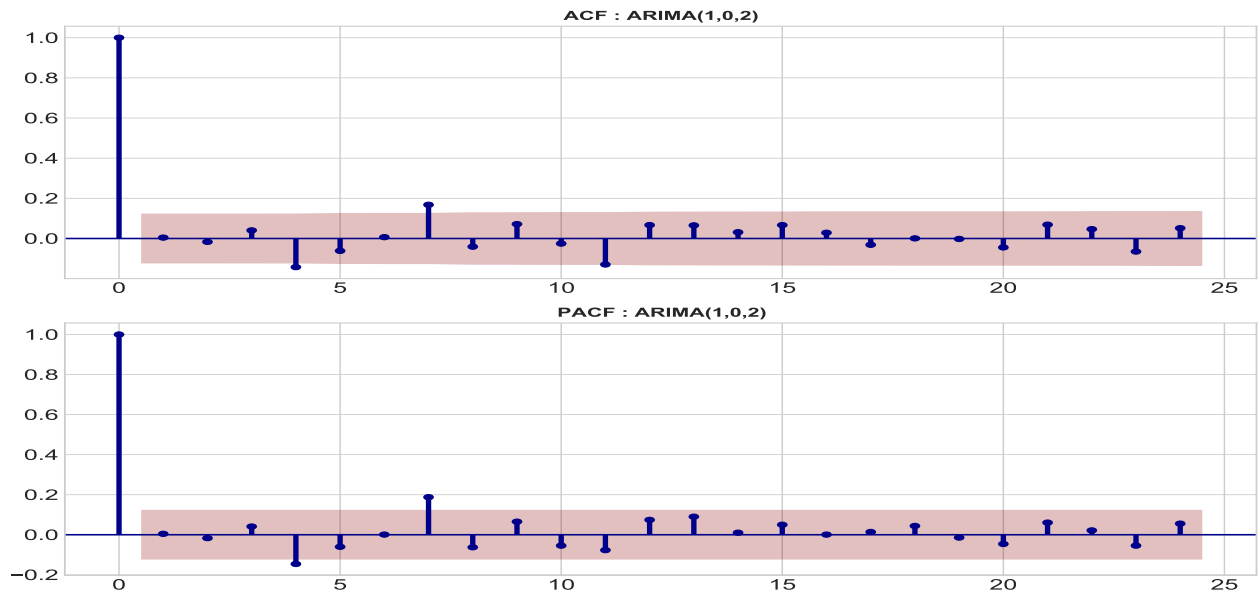


FIGURE 14 – ACF et PACF des résidus du modèle  $ARIMA(1,0,2)$



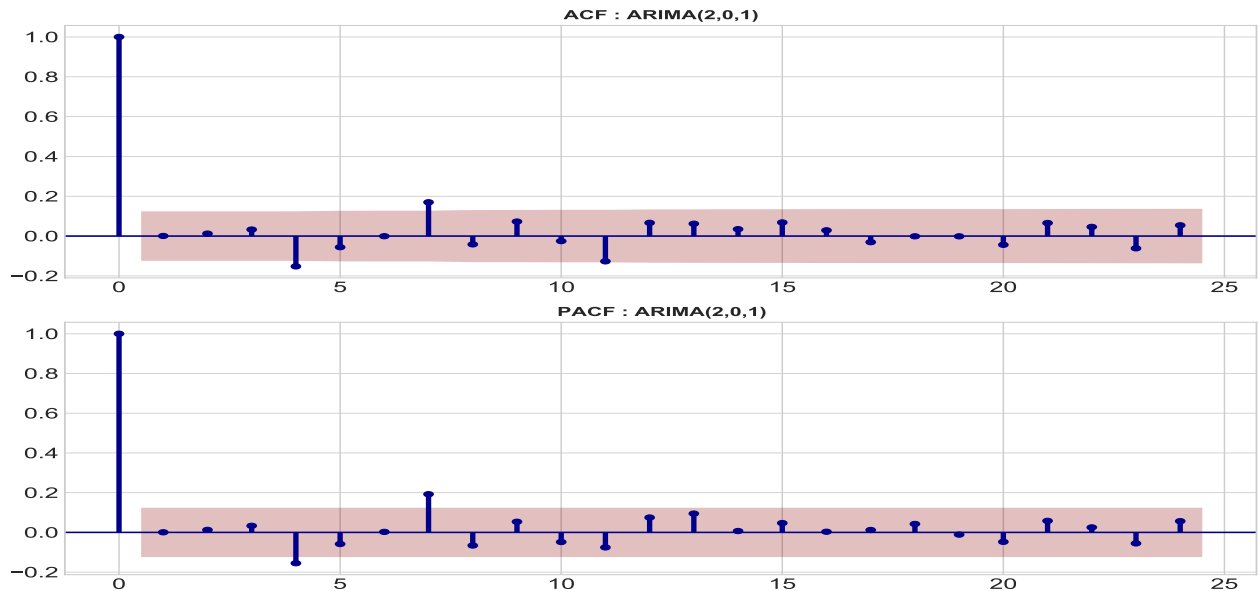


FIGURE 15 – ACF et PACF des résidus du modèle  $ARIMA(2,0,1)$

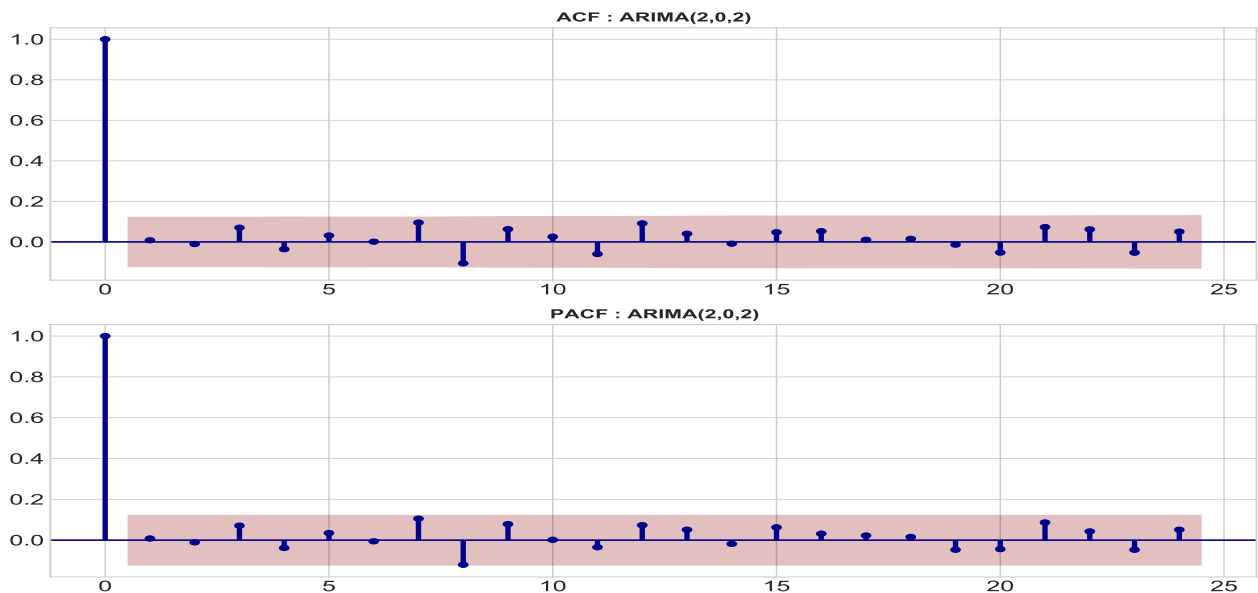


FIGURE 16 – ACF et PACF des résidus du modèle  $ARIMA(2,0,2)$

On constate que les résidus des 3 premiers modèles présentent des pics au niveau de l'ACF et le PACF pour un ordre de retards maximal égal à 7. l'ACF et le PACF des résidus du dernier modèle correspondent exactement au ceux d'un bruit blanc. Les résultats des tests de Ljung-Box correspondants sont :

```
model : ARIMA ( 1 ,0, 1 ) ; AIC : -1448.6115 ; LB p-value : 0.0112
model : ARIMA ( 1 ,0, 2 ) ; AIC : -1447.6899 ; LB p-value : 0.0075
model : ARIMA ( 2 ,0, 1 ) ; AIC : -1447.0775 ; LB p-value : 0.0059
model : ARIMA ( 2 ,0, 2 ) ; AIC : -1455.6938 ; LB p-value : 0.2377
```

On voit bien que le modèle  $ARIMA(2,0,2)$  est celui vérifie l'hypothèse du bruit blanc pour les résidus et possède l'AIC minimal. D'ailleurs c'est le modèle proposé par `auto_arma` du package `pmdarima` de Python :

Dep. Variable:	y	No. Observations:	249
Model:	<b>SARIMAX(2, 0, 2)</b>	Log Likelihood	733.806
Date:	Sat, 19 Jun 2021	AIC	-1457.611
Time:	21:02:51	BIC	-1440.024
Sample:	0	HQIC	-1450.532
	- 249		

	coef	std err	z	P>  z	[0.025	0.975]
ar.L1	0.9991	0.115	8.698	0.000	0.774	1.224
ar.L2	-0.7935	0.082	-9.718	0.000	-0.954	-0.633
ma.L1	-0.8501	0.104	-8.212	0.000	-1.053	-0.647
ma.L2	0.7958	0.083	9.640	0.000	0.634	0.958
sigma2	0.0002	5.79e-06	27.829	0.000	0.000	0.000

**Remarque :** On a estimé d'abord un modèle avec dérive, mais la constante était significativement nulle au risque de 5%.

On remarque que les coefficients du modèle sont significativement non nuls au risque d'erreur de 5%. Il est à noter que ces tests statistiques sont fondés sur les hypothèses de la normalité et de l'homoscédasticité du bruit blanc. Des hypothèses qu'on va vérifier par la suite. Il convient également de vérifier que les polynômes auto-régressif et moyenne-mobile n'ont pas de racines communes pour s'assurer que la représentation soit minimale, dans le sens qu'il n'existe pas d'autres représentations qui donnent les mêmes résultats.<sup>24</sup>

AR roots : [0.62951836-0.92946248j , 0.62951836+0.92946248j]

MA roots : [0.53409531-0.98557502j , 0.53409531+0.98557502j]

### 4.2.3 Analyse des résidus

**Absence d'autocorrélations :** Cette hypothèse a été déjà vérifiée auparavant.

**Normalité :**

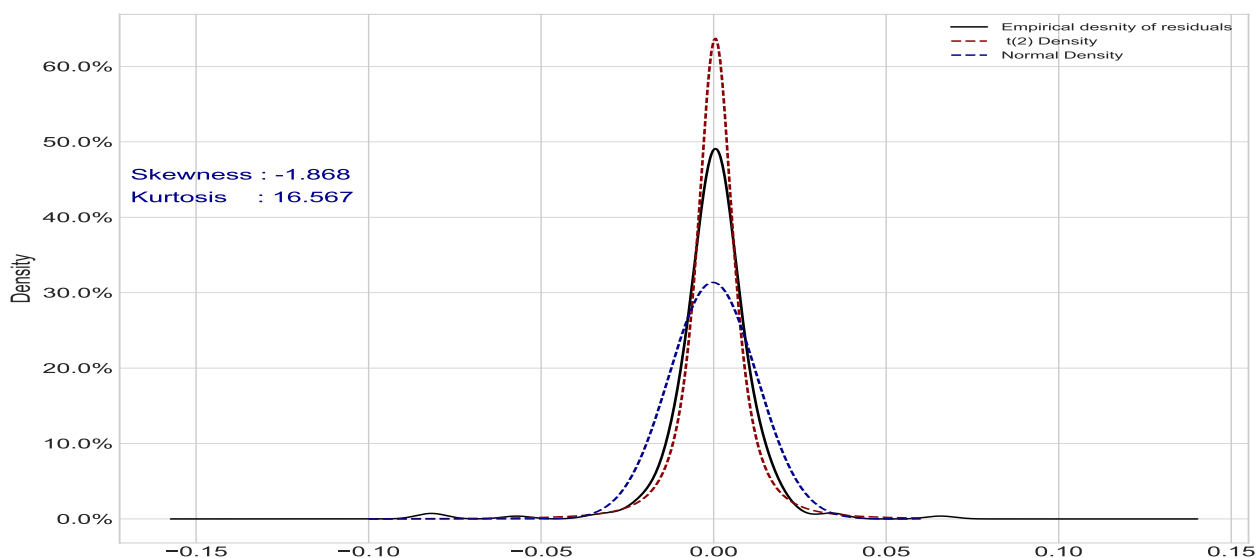


FIGURE 17 – Densité empirique des résidus du modèle  $ARIMA(2, 0, 2)$

24. Si c'est le cas, ceci peut conduire à des erreurs lors de la prévision. Voir *Time Series and Dynamic Models*. pages 159-160

```
ShapiroResult(statistic = 0.769576907157898, pvalue = 2.0762396743251974e-18)
```

```
KstestResult(statistic = 0.48229592475762917, pvalue = 8.062751837234154e-54)
```

On remarque d'après la figure, que les résidus du modèle possèdent une densité leptokurtique avec un kurtosis égal à 16,567. En plus, les tests de Shapiro-Wilk et de Kolmogorov-Smirnov confirment ce constat en rejetant l'hypothèse de la normalité des résidus au risque de 5%.

Un autre constat est le fait que la loi de Student  $t(2)$  s'ajuste bien à la distribution des résidus, notamment au niveau des queues. Ce constat est confirmé par le test de Kolmogorov-Smirnov.

```
KstestResult(statistic = 0.027439347029434513, pvalue = 0.9894807176773781)
```

**Homoscédasticité :** On commence par présenter l'ACF et le PACF de la série des carrés des résidus, ainsi que le graphique de la série des résidus.

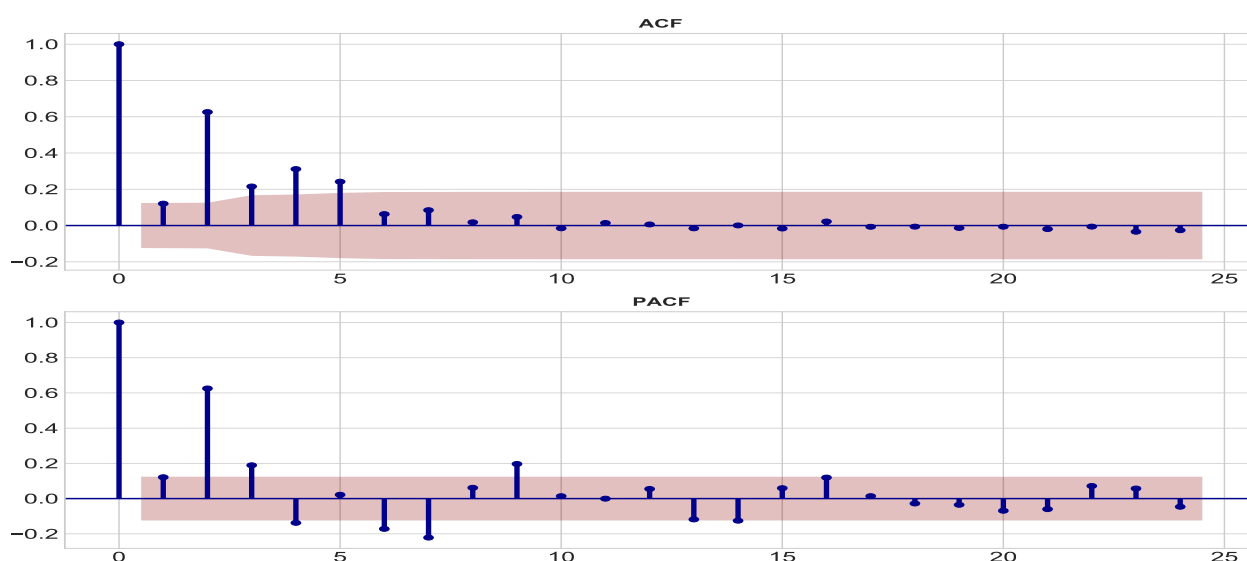


FIGURE 18 – ACF et PACF de la série des carrés des résidus du modèle

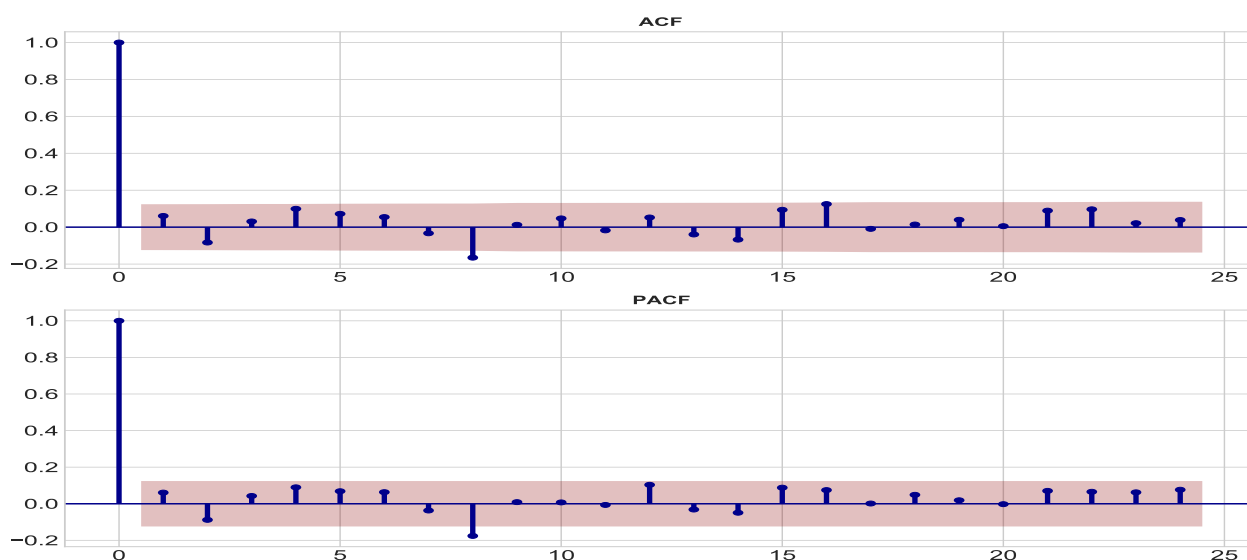


FIGURE 19 – Graphique de la série des résidus du modèle

Si les résidus du modèle montraient une absence d'autocorrélation à tout ordre (au risque d'erreur de 5%) ; qui correspond à une absence de relation linéaire entre un résidu et ses valeurs retardées.

La série des carrés des résidus montre des pics au niveau de l'ACF et du PACF. De ce fait, les résidus montrent une *dépendance sérielle* (Serial dependence). En plus, on observe sur le graphique de la série des résidus que la variance conditionnelle varie dans le temps. Ainsi, l'hypothèse : " les résidus sont i.i.d" semble être non-valide pour cette série !

### 4.3 Estimer la volatilité ?

On a vu que la volatilité de la série des log-rendements n'est pas constante. En plus, elle n'est pas directement observable. Si on note par  $\mathcal{F}_t$  la tribu canonique associée au processus  $(r_t)_{t \in \mathbb{N}}$ , la volatilité des log-rendements peut être définie, à un instant  $t$ , par : <sup>25</sup>

$$\sigma_t = \sqrt{\text{Var}(r_t \mid \mathcal{F}_{t-1})}$$

Plusieurs estimateurs ont été proposés pour estimer et prédire la volatilité, basés sur l'historique des log-rendements. On parle de volatilité historique.

#### Estimateur de type Moyenne-Mobile (Volatilité glissante)

$$\hat{\sigma}_t = \left( \frac{1}{h} \sum_{j=1}^h r_{t-j}^2 \right)^{\frac{1}{2}}$$

Tel que  $h$  est une fenêtre à choisir. <sup>26</sup>

**Estimateur de type EWMA** Le système RiskMetrics de JP Morgan, discuté dans l'introduction, a proposé un estimateur de type moyenne mobile à poids exponentiels **EWMA** - Eponential Weighted Moving Average, défini par :

$$\hat{\sigma}_t^2 = \beta \hat{\sigma}_{t-1}^2 + (1 - \beta) r_{t-1}^2$$

Ou bien :

$$\hat{\sigma}_t = \left( \frac{1 - \beta}{\beta(1 - \beta^h)} \sum_{j=1}^h \beta^j r_{t-j}^2 \right)^{\frac{1}{2}}$$

RiskMetrics recommande une valeur  $\beta = 0.94$  pour les rendements journaliers.

D'autres estimateurs ont été proposés, qui n'utilisent pas forcément que les prix de clôture. On note :

- $M_t$  : le prix maximal observé le jour  $t$ .
- $L_t$  : le prix minimal observé le jour  $t$ .
- $O_t$  : le prix d'ouverture observé le jour  $t$ .

On peut citer : <sup>27</sup>

#### Rogers & Satchell (1991) :

$$\hat{\sigma}_t = \left( \frac{1}{h} \sum_{j=t-h}^t \log \left( \frac{M_j}{O_j} \right) \cdot \log \left( \frac{M_j}{P_j} \right) + \log \left( \frac{L_j}{O_j} \right) \cdot \log \left( \frac{L_j}{P_j} \right) \right)^{\frac{1}{2}}$$

25. On voit bien que dans le cadre des modèles ARIMA, cette quantité est constante dans le temps.

26. Figlewski a noté dans *Forecasting Volatility. (1997)* que le fait de considérer les fluctuations autour de 0, au lieu de la moyenne empirique augmente la précision des prévisions.

27. Pour plus de détails, voir : <https://www.cmegroup.com/trading/fx/files/volEst.pdf>

Parkinson (1980) :

$$\hat{\sigma}_t = \left( \frac{1}{4h \log(2)} \sum_{j=t-h}^t \left( \log \left( \frac{M_j}{L_j} \right) \right)^2 \right)^{\frac{1}{2}}$$

Garman & Klass (1980) :

$$\hat{\sigma}_t = \left( \frac{0.51}{h-1} \sum_{j=t-h}^t \left( \log \left( \frac{M_j}{L_j} \right) \right)^2 - \frac{0.39}{h-1} \sum_{j=t-h}^t r_j^2 \right)^{\frac{1}{2}}$$

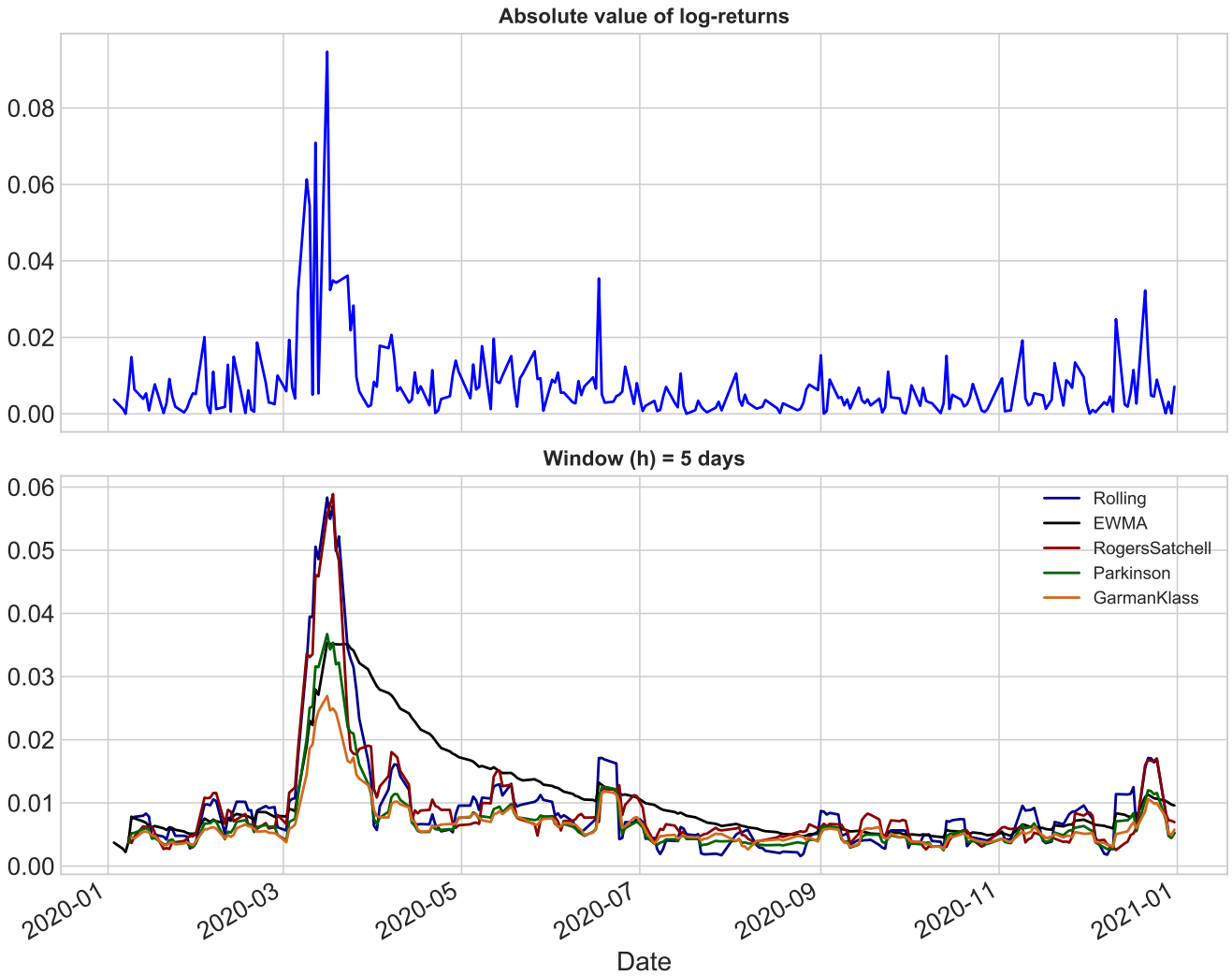


FIGURE 20 – Différents Estimateurs de la volatilité

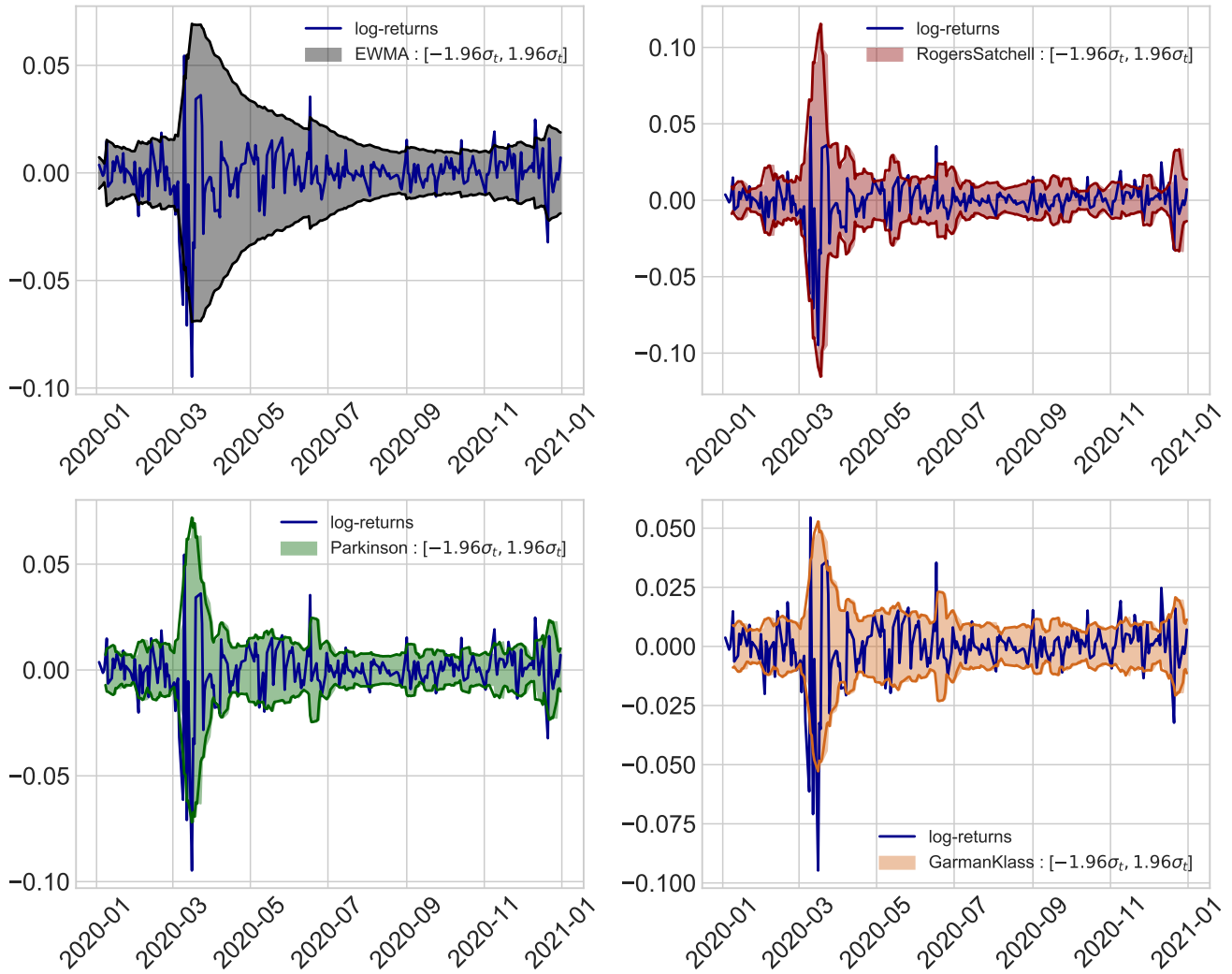


FIGURE 21 – Intervalles de confiances de 0 (moyenne) au niveau 95%, en supposant la normalité conditionnelle

#### 4.4 Modèles GARCH

Les modèles ARCH (AutoRegressive Conditionally Heteroscedastic) ont été introduits par Engle en 1982, puis généralisés en modèles GARCH (Generalized ARCH) en 1986 par Bollerslev. Dans ces modèles, le concept clé est le fait qu'ils prennent en considération la variation de la variance conditionnelle dans le temps. D'où le mot *Conditionnellement hétéroscédastique*

*Econometrica*, Vol. 50, No. 4 (July, 1982)

### AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY WITH ESTIMATES OF THE VARIANCE OF UNITED KINGDOM INFLATION<sup>1</sup>

BY ROBERT F. ENGLE

Traditional econometric models assume a constant one-period forecast variance. To generalize this implausible assumption, a new class of stochastic processes called autoregressive conditional heteroscedastic (ARCH) processes are introduced in this paper. These are mean zero, serially uncorrelated processes with nonconstant variances conditional on the past, but constant unconditional variances. For such processes, the recent past gives information about the one-period forecast variance.

#### 4.4.1 Définition

Un processus  $(\varepsilon_t)_{t \in \mathbb{N}}$  est un processus  $GARCH(p, q)$  fort et stationnaire si :

$$\begin{cases} \varepsilon_t = \eta_t \sigma_t \\ \sigma_t^2 = \omega + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \end{cases}$$

où :

- $\omega > 0$ ,  $\forall j \in \llbracket 1, p \rrbracket \alpha_j \geq 0$  et  $\forall j \in \llbracket 1, q \rrbracket \beta_j \geq 0$
- $(\eta_t)_{t \in \mathbb{N}}$  est un SWN(0, 1), tel que pour tout instant  $t$  :  $\eta_t$  est indépendant de  $\sigma(\sigma_s : s \leq t)$
- $\sigma_t^2 = \text{Var}(\varepsilon_t \mid \mathcal{F}_{t-1})$ . Avec,  $(\mathcal{F}_t)_{t \in \mathbb{N}}$  est la filtration canonique associée au processus  $(\varepsilon_t)_{t \in \mathbb{N}}$ .
- En absence de la partie autorégressive dans l'équation de la variance conditionnelle, le processus est dit :  $ARCH(p)$ .
- $\sum_{j=1}^p \alpha_j + \sum_{j=1}^q \beta_j < 1$  (Condition de stationnarité au second ordre). Si cette dernière somme vaut 1, le processus est dit :  $IGARCH(p, q)$  (Integrated GARCH).<sup>28</sup>

#### Remarques :

- Si la dernière condition de stationnarité au sens large est vérifiée, la *variance inconditionnelle* s'écrit comme :

$$\text{Var}(\varepsilon_t) = \frac{\omega}{1 - \sum_{j=1}^p \alpha_j - \sum_{j=1}^q \beta_j}$$

- $\mathbb{E}(\varepsilon_t \mid \mathcal{F}_{t-1}) = 0$ . On en déduit :  $\mathbb{E}(\varepsilon_t) = \mathbb{E}(\mathbb{E}(\varepsilon_t \mid \mathcal{F}_{t-1})) = 0$
- Les distributions classiques des innovations  $(\eta_t)_{t \in \mathbb{N}}$  sont : La loi normale centrée réduite, Student réduite, Skew-Student, Generalized error distribution (GED), Skew-GED, Skew-Normal ..

#### 4.4.2 Test des effets ARCH : (LM test)

On dispose de  $T$  observation d'une série temporelle de processus générique  $(\varepsilon_t)_{t \in \mathbb{N}}$  supposé  $ARCH(p)$ . On désire tester :

$$\mathcal{H}_0 : \alpha_1 = \dots = \alpha_p = 0$$

En posant  $\theta = (\omega, \alpha_1, \dots, \alpha_p)$ , la statistique du test s'écrit :

$$LM_n = \frac{1}{n} {}^t(\nabla_{\theta} QL(\theta)) \hat{I}^{-1}(\nabla_{\theta} QL(\theta))$$

où :  $\nabla_{\theta}$  désigne le gradient sous  $\theta$ ,  $QL$  est la quasi-vraisemblance de l'échantillon.  $I$  est la matrice d'information de Fisher.

Asymptotiquement, sous  $\mathcal{H}_0$ , on a :

$$LM_n \sim \chi_p^2$$

Testons maintenant l'homoscédasticité conditionnelle des résidus du modèle  $ARIMA(2, 0, 2)$ . D'après la sortie du logiciel, cette hypothèse est rejetée au risque d'erreur de 5%. Ainsi, il y a présence des effets ARCH dans les résidus.

LM : 1.185288e+02 ; LM p-value : 1.053025e-17

28. Ce processus est stationnaire au sens strict (fort), cependant il n'admet pas un moment d'ordre 2 et n'est pas stationnaire au sens large (faible)

#### 4.4.3 Estimation des paramètres du modèle

Rappelons qu'on a déjà trouvé que la distribution de Student à deux degré de liberté s'ajuste bien à la distribution des résidus. Pour des raisons concernant la convergence de l'algorithme d'optimisation, **on a multiplié les données par 100** (rescale). Le modèle choisi par le logiciel est  $GARCH(1, 1)$ .<sup>29</sup>

Dep. Variable:	None	R-squared:	0.000
Mean Model:	Zero Mean	Adj. R-squared:	0.004
Vol Model:	<b>GARCH</b>	Log-Likelihood:	-314.044
Distribution:	<b>Standardized Student's t</b>	AIC:	636.088
Method:	Maximum Likelihood	BIC:	650.158
		No. Observations:	249
Date:	Mon, Jun 21 2021	Df Residuals:	249

	coef	std err	t	P>  t	95.0% Conf. Int.
omega	0.0917	5.458e-02	1.680	<b>9.289e-02</b>	[-1.526e-02, 0.199]
alpha[1]	0.2657	0.134	1.978	<b>4.797e-02</b>	[2.380e-03, 0.529]
beta[1]	0.7253	8.514e-02	8.519	<b>1.614e-17</b>	[ 0.558, 0.892]
	coef	std err	t	P>  t	95.0% Conf. Int.
nu	3.1681	0.671	4.720	2.359e-06	[ 1.853, 4.484]

Unconditional variance : 0.0014454061370732165

On remarque que les coefficients du modèles sont significativement non nuls au risque d'erreur de 5% et d'après l'ACF et le PACF des innovations du modèle, il parraît qu'il s'agit d'un WWN.

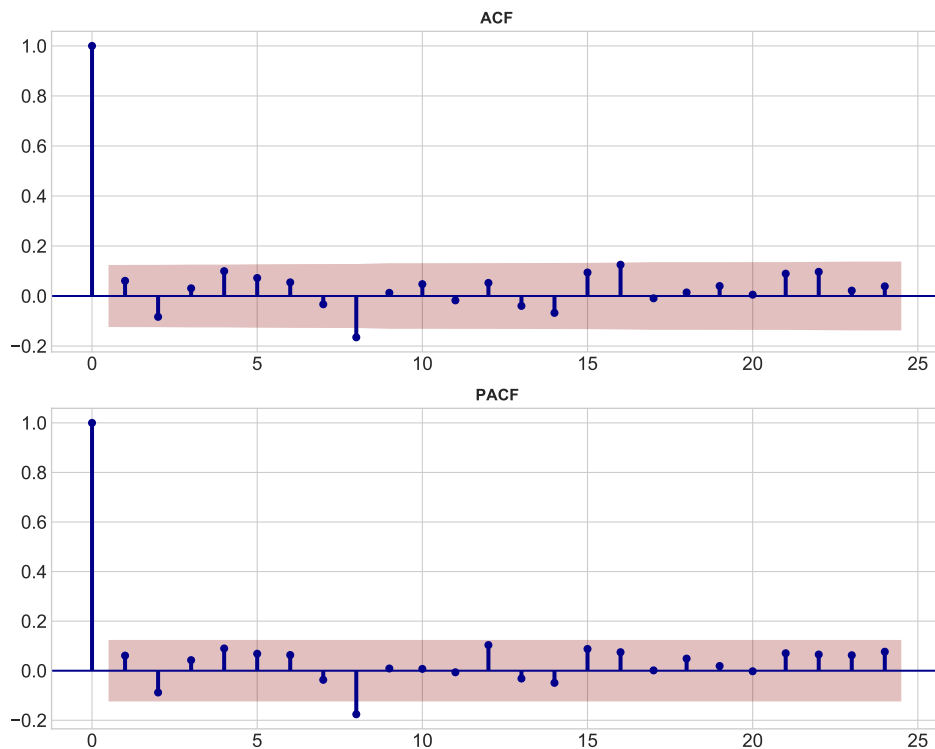


TABLE 4 – ACF et PACF des résidus du modèles GARCH

29. Normalement, pour identifier les paramètres  $p$  et  $q$ , il faut voir l'ACF des carrés des résidus, **en corrigeant les bornes des intervalles de confiance** ou bien utiliser la méthode du coin.



Ce dernier constat est confirmé par les résultats du test de Ljung-Box :

lag	lb_stat	lb_pvalue
1	0.928382	0.335284
2	2.682071	0.261575
3	2.923457	0.403578
4	5.444749	0.244629
5	6.776300	0.237816
6	7.543489	0.273491
7	7.827084	0.348088

TABLE 5 – Test de Ljung-Box appliqué à la série des résidus du modèle  $GARCH(1, 1)$

On teste par la suite la présence des effets ARCH pour ces résidus. Les résultats du LM-test sont :

LM : 7.444435 ; LM p-value : 0.963714

Donc, on accepte l'hypothèse de l'homoscédasticité conditionnelle au risque d'erreur de 5%.

## 4.5 Prévisions

On rappelle que le modèle choisi pour modéliser la série des log-rendements du MADEX est un  $ARIMA(2, 0, 2)$  à erreurs  $GARCH(1, 1)$  dont les erreurs sont de loi Student  $t(3)$  réduite.

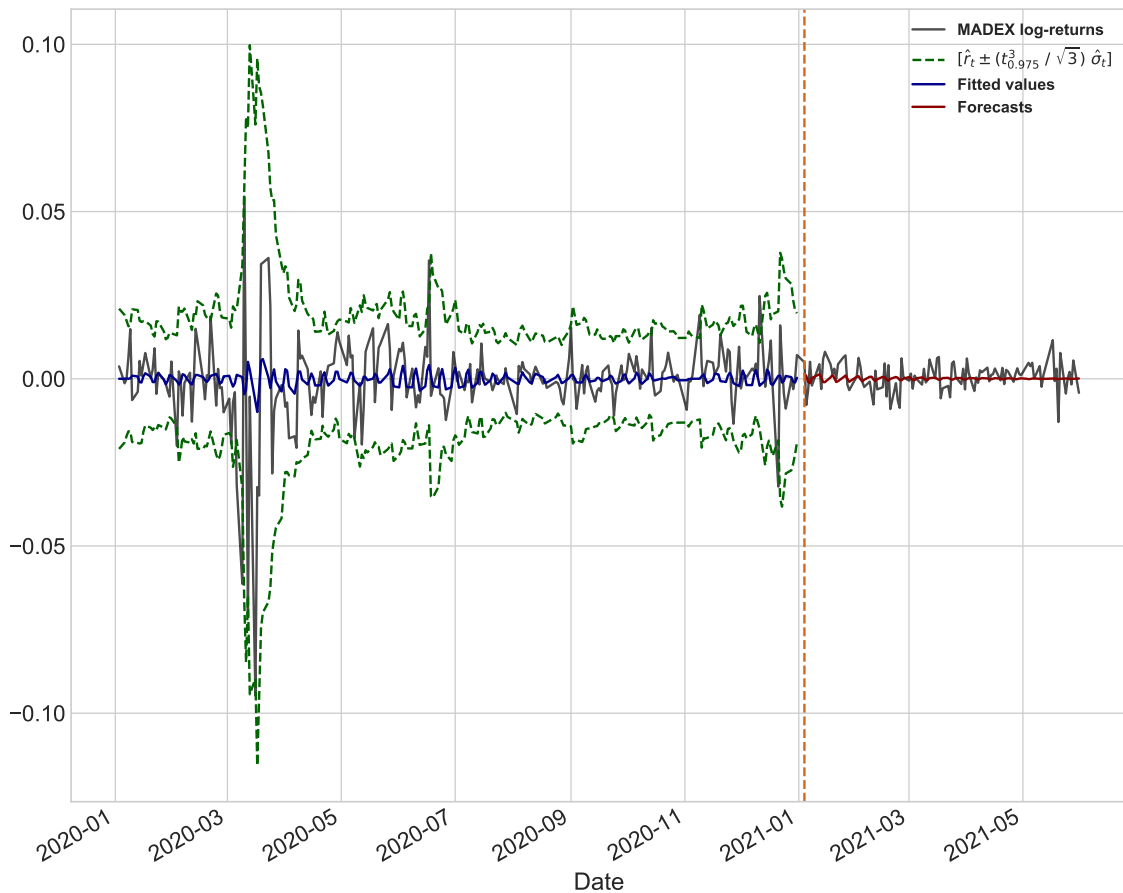


FIGURE 22 – Valeurs Ajustées et prévisions

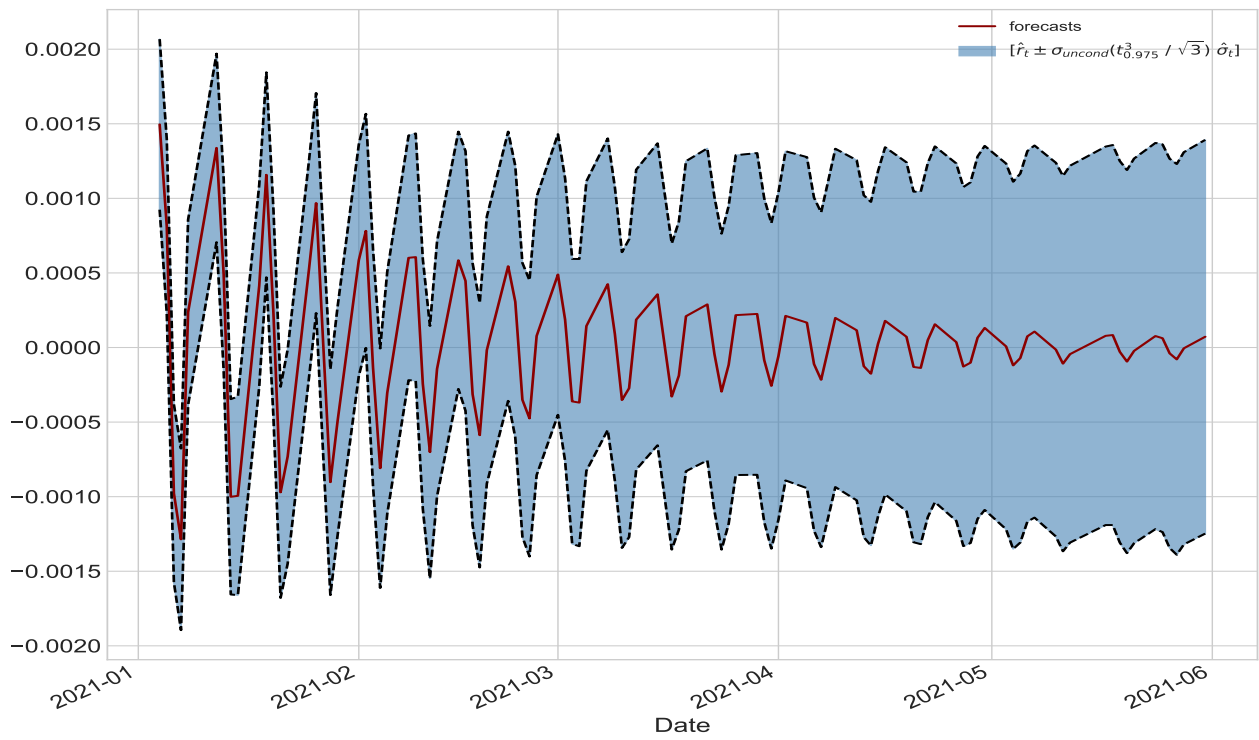


TABLE 6 – Un "Zoom" sur les prévisions

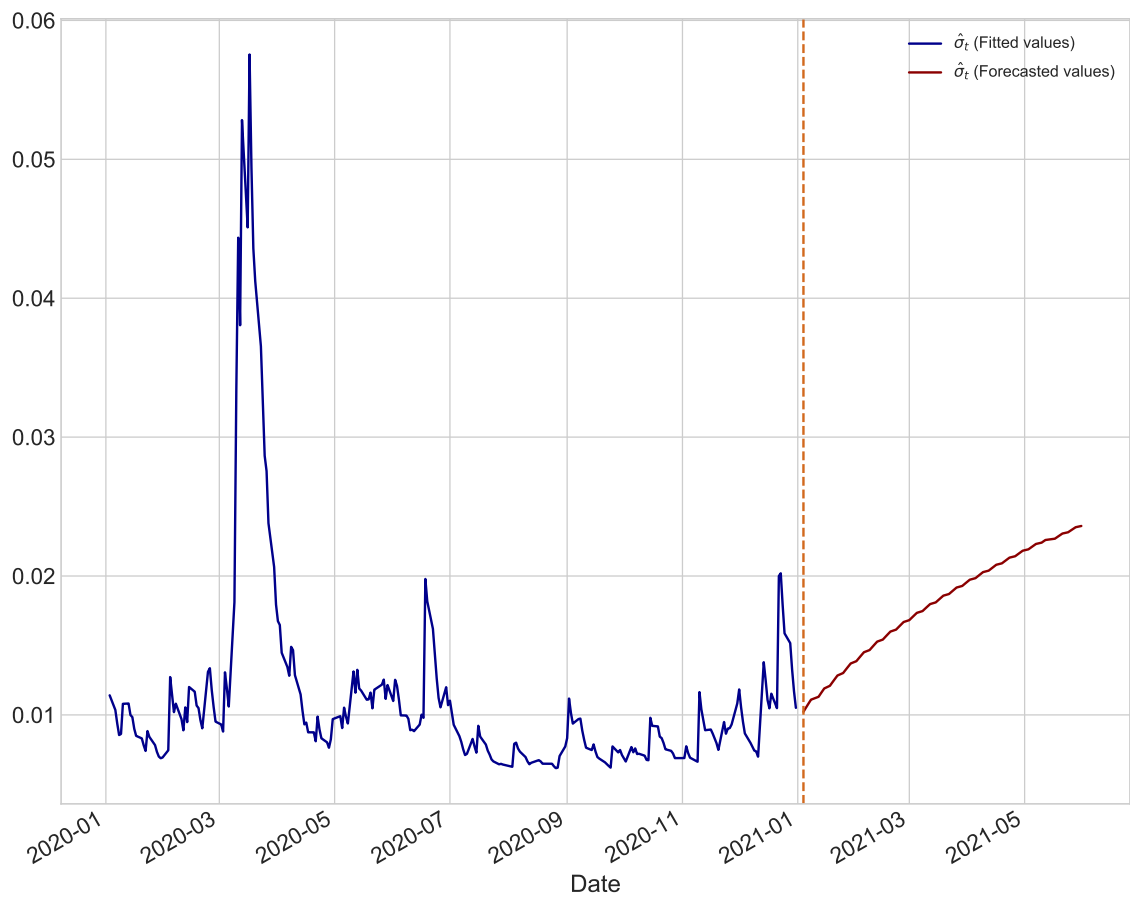


FIGURE 23 – Volatilité Ajustée et prédite

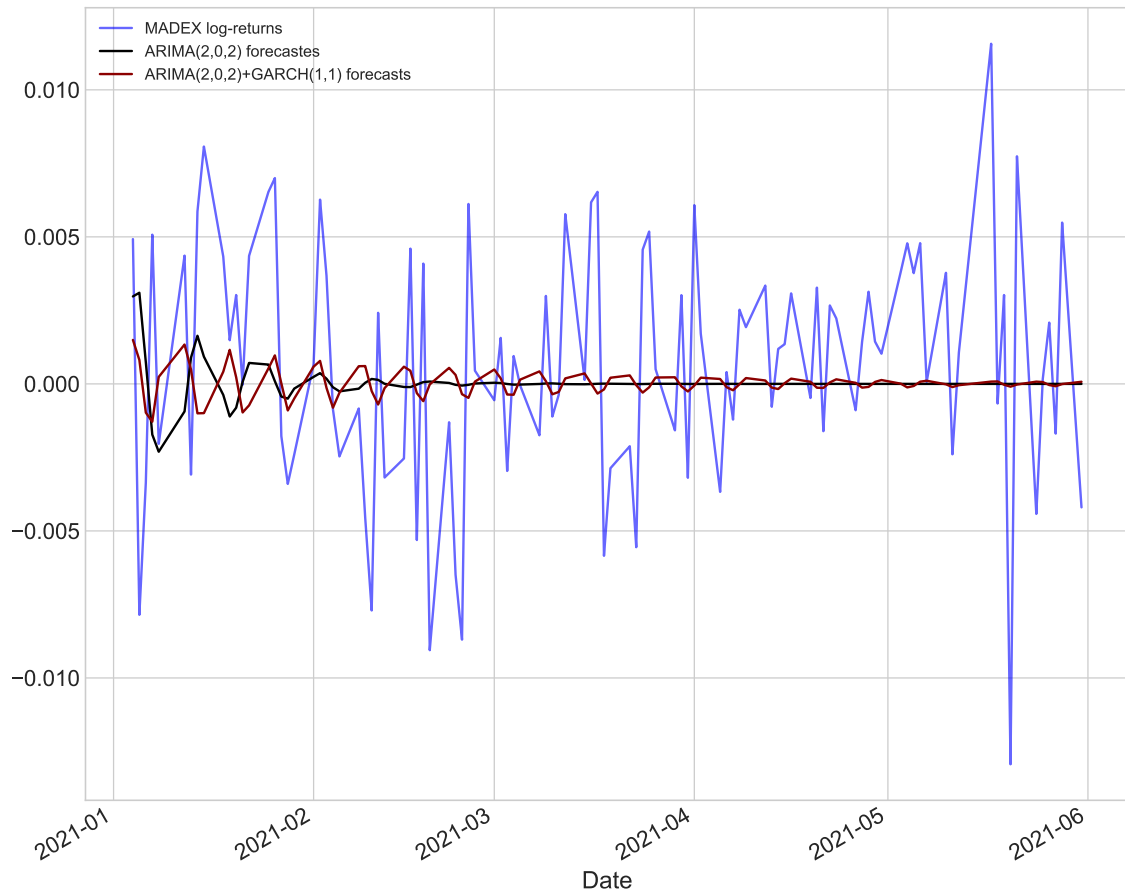


FIGURE 24 – Valeurs prédites données par le modèle  $ARIMA(2,0,2)$  contre celles fournies par  $ARIMA(2,0,2) + GARCH(1,1)$

Les erreurs de prévisions des deux modèles sont :

$ARIMA(2,0,2)$  RMSE : 0.00426881

$ARIMA(2,0,2)+GARCH(1,1)$  RMSE : 0.00424652

On remarque que le deuxième modèle a une erreur de prévision inférieure à celle du premier. De plus, d'après les graphiques, on constate que le deuxième a capté la variation de la variance conditionnelle au cours du temps. Ainsi, on retient le modèle  $ARIMA(2,0,2)$  à erreurs  $GARCH(1,1)$  pour modéliser les log-rendements du MADEX. Il est à noter que les intervalles (en vert) de la figure 22, ne sont pas des intervalles de confiances des valeurs ajustées. On remarque également d'après les graphiques que le modèle a un faible pouvoir prédictif. Ceci est, peut-être, dû au fait qu'il ne distingue pas entre les chocs qui résultent en une perte et ceux qui résultent en un gain. Ceci est un inconvénient car on sait d'après l'effet de levier, qu'il existe une asymétrie entre l'impact des chocs positifs et négatifs. C'est pour cela qu'il existe une famille des variants du modèle GARCH classique qui sont mieux adaptés aux faits stylisés des séries financières. On peut citer : *EGARCH* (*Exponential GARCH*), *GJR-GARCH* (*Glosten-Jagannathan-Runkle GARCH*) ...

On finira cette partie par un aphorisme du statisticien George Box : "*All models are wrong, but some are useful*". Dans le cadre de notre modélisation, on peut le modifier en : All models are useful, but they come with limitations.

## 5 Retour à la Value-at-Risk

Rappelons qu'on a défini (section 2.1) la  $\text{VaR}_\alpha$  comme le quantile d'ordre  $\alpha$  de la loi de la perte et on distingue entre une VaR conditionnelle et inconditionnelle. On note également que la VaR dépend de 3 paramètres, à savoir : l'horizon  $\Delta$ , le niveau de confiance  $\alpha$  et la distribution de la perte.

On se place dans le cadre introduit dans la section 1.1.

### 5.1 Méthodes de calcul de la VaR

Selon J. McNeil et R. Frey (2005), il existe généralement trois méthodes standards pour calculer la VaR.

**La méthode Variance-Covariance (paramétrique) :** Dans cette méthode, on suppose que la variation des facteurs de risque, à l'instant  $t + 1$ ,  $X_{t+1}$  est de loi normale multivariée :

$$X_{t+1} \sim \mathcal{N}_d(\mu, \Sigma)$$

De plus, on suppose un profil de risque linéaire pour le processus des pertes :

$$L_{t+1} = -(c_t + {}^t b_t X_{t+1})$$

Tel que :  $c_t$  est une constante et  $b_t$  est un vecteur de taille  $d$ , supposés connus à l'instante  $t$   
Alors :

$$L_{t+1} \sim \mathcal{N}(-c_t - {}^t b_t \mu, {}^t b_t \Sigma b_t)$$

De ce fait la  $\text{VaR}_\alpha$  **inconditionnelle** s'écrit :<sup>30</sup>

$$\text{VaR}_\alpha = -c_t - {}^t b_t \mu + \sqrt{{}^t b_t \Sigma b_t} \phi^{-1}(\alpha)$$

$\mu$  et  $\Sigma$  sont estimés à partir d'un échantillon des variations des facteurs de risque.

Pour la version conditionnelle de cette méthode, on suppose que :

$$X_{t+1} \mid \mathcal{F}_t^{\mathbf{X}} \sim \mathcal{N}_d(\mu_{t+1}, \Sigma_{t+1})$$

Dans ce cas la VaR calculée et notée  $\text{VaR}^{t+1|t}$  est dite : **conditionnelle**.  $\mu_{t+1}$ , et  $\Sigma_{t+1}$  sont obtenues par des prévisions après la construction d'un modèle pour le processus des variations des facteurs de risque : *GARCH*, *ARIMA - GARCH*, *EWMA* ..

Parmi les inconvénients de cette méthode, on cite l'hypothèse de la normalité, qui est peu réaliste. Une extension de cette méthode peut-être faite en supposant que  $X_{t+1}$  est de loi de Student multivariée :

$$X_{t+1} \sim t_d(\nu, \mu, \Sigma)$$

Tel que :

$$f_{t_d(\nu, \mu, \Sigma)}(\mathbf{x}) = \frac{\Gamma(\frac{1}{2}(\nu + d))}{\Gamma(\frac{1}{2}\nu)(\pi\nu)^{\frac{1}{2}}|\Sigma|^{\frac{1}{2}}} \left(1 + \frac{{}^t(\mathbf{x} - \mu)\Sigma^{-1}(\mathbf{x} - \mu)}{\nu}\right)^{\frac{-(\nu+d)}{2}}$$

où :  $\Gamma$  est la fonction gamma d'Euler.

---

30.  $\phi$  est la fonction de répartition de la loi normale standard

**La méthode de la Simulation Historique (non-paramétrique) :** Dans cette méthode on utilise les données historiques. Le calcul de la VaR est basé sur une hypothèse assez forte, à savoir : Les variations des facteurs de risque sont i.i.d. Par conséquent, la VaR est estimée par le quantile empirique de la distribution des variations des facteurs de risque. Il s'agit d'une méthode inconditionnelle.

**La méthode de la Simulation Monte-Carlo (paramétrique) :** Cette méthode garde le même principe de la simulation historique sauf qu'au lieu d'estimer la distribution des rendements en se basant sur l'historique, la simulation de Monte-Carlo utilise des scénarios qui sont générés selon un modèle qui est extrait de l'historique des données. On suppose que les variations des facteurs de risque des titres sont régis par une distribution paramétrique avec des paramètres connus. Après l'obtention de scénarios possibles pour cette série. Le problème de l'obtention de la VaR est résolu de manière similaire à l'approche de la Simulation Historique. Cette simulation est répétée plusieurs fois. Les étapes de l'estimation sont connues par : *Crude Monte-Carlo Simulation*. Cette méthode peut-être conditionnelle ou inconditionnelle.

**Remarque :** On a défini la VaR en utilisant le processus des pertes. Il peut être défini également par le processus des rendements simples ou celui des log-rendements. On note  $\text{VaR}(X)$  avec  $X \in \{L, R, r\}$ , la VaR défini avec chacun de ces processus. On a :

$$\begin{aligned}\mathbb{P}(L_{t+1} \leq \text{VaR}_\alpha^{t+1|t}(L) \mid \mathcal{F}_t) = \alpha &\Leftrightarrow \mathbb{P}(R_{t+1} \geq -\frac{\text{VaR}_\alpha^{t+1|t}(L)}{V_t} \mid \mathcal{F}_t) = \alpha \\ &\Leftrightarrow \mathbb{P}(r_{t+1} \geq \ln(1 - \frac{\text{VaR}_\alpha^{t+1|t}(L)}{V_t}) \mid \mathcal{F}_t) = \alpha\end{aligned}$$

Ainsi :

$$\text{VaR}_\alpha^{t+1|t}(L) = -V_t \text{VaR}_{1-\alpha}^{t+1|t}(R) = V_t \left[ 1 - \exp\left(\text{VaR}_{1-\alpha}^{t+1|t}(r)\right) \right]$$

Ainsi, le calcul de la  $\text{VaR}(L)$ , revient au calcul de la  $\text{VaR}(r)$ .

## 5.2 Quelques mots de la théorie des valeurs extrêmes

La théorie des valeurs extrêmes peut se résumer aux mathématiques des événements extrêmes. elle s'intéresse uniquement à la modélisation des queues de distribution à partir des lois spécifiques propres. Cela marque une différence importante comparativement aux théories qui font une modélisation totale de la distribution. Elle est basée principalement sur deux approches : l'approche du Block Maxima et l'approche Peak Over Threshold (POT). L'approche POT basée sur la distribution de Pareto généralisée ou en anglais Generalized Pareto Distribution (GPD), fournit une méthode pour modéliser les excès au-delà d'un seuil relativement élevé. La fonction de répartition de la loi GPD est défini par :

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}} & \text{si : } \xi \neq 0 \\ 1 - \exp\left(-\frac{x}{\beta}\right) & \text{si : } \xi = 0 \end{cases}$$

Tel que :  $x \geq 0$  si  $\xi \geq 0$  et  $x \in [0, -\frac{\beta}{\xi}]$  si  $\xi < 0$

Dans le cadre de la gestion quantitative du risque, les techniques de la théorie des valeurs extrêmes (EVT) <sup>31</sup> sont utilisées pour étendre la méthode non-paramétrique de la simulation historique, en une méthode **semi-paramétrique**, en donnant des estimateurs plus précis des quantiles de la distribution que ceux donnés par la première méthode. Pour plus de détails, le lecteur peut voir : *Modelling Extremal Events for Insurance and Finance*. Embrechts, P., Klüppelberg, C. and Mikosch, T.

---

31. Extreme Value Theory

### 5.3 Validation d'une prévision de la VaR (le Backtesting)

Il est nécessaire de tester la précision du calcul de la VaR. Étant donné que la véritable valeur de la VaR n'est pas observable, les techniques usuelles de validation des prévisions telles que l'erreur quadratique moyenne et l'erreur absolue moyenne en pourcentage ne peuvent pas être utilisées. Dès lors, des procédures de *backtesting* ont été développées.

Le Backtesting est un ensemble de procédures statistiques dont le but est de vérifier que les pertes réelles observées ex-post sont en adéquation avec les pertes prévues. Cela implique de comparer systématiquement l'historique des prévisions de la Value-at-Risk aux rendements observés du portefeuille. Avant de donner une définition plus formelle de ces tests, on introduit une approche d'estimation et de prévision qui nous sera utile dans les applications.

#### 5.3.1 Estimation et prévision par fenêtre glissante (Rolling window estimating and forecasting)

Afin de capter les caractéristiques dynamiques de la série d'intérêt qui varient en fonction du temps, l'approche d'estimation par fenêtre glissante peut être envisagée. Considérons une série de  $T$  observations des log-rendements  $(r_t)_{t \in \llbracket 1, T \rrbracket}$  (par exemple). Nous fixons une taille de fenêtre notée  $\omega$ . Pour tout  $t \in \llbracket \omega, T \rrbracket$ , on estime les paramètres du modèle en utilisant l'échantillon  $(r_t)_{t \in \llbracket t-\omega+1, t \rrbracket}$  et on calcule la prévision à horizon 1.

Dans la pratique, on divise notre échantillon en deux parties, une sera utilisée pour estimer les paramètres du modèle et l'autre pour le test. Cette méthode sera utilisée dans le Backtesting.

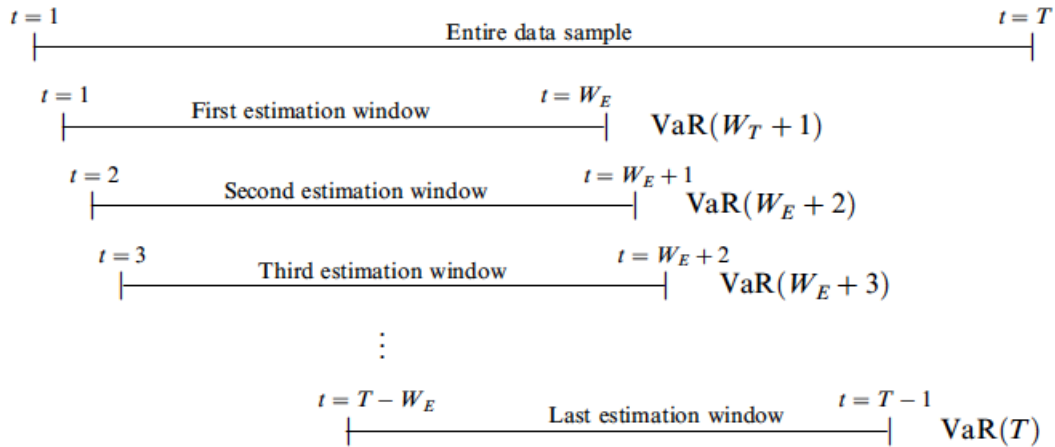


FIGURE 25 – Illustration de l'estimation et la prévision de la VaR par fenêtre glissante pour une fenêtre d'estimation de taille  $W_E$

#### 5.3.2 Backtesting

Le but du Backtesting est de vérifier que la fréquence de dépassements de la VaR estimée par les pertes réelles observées est conforme avec le niveau de confiance de l'estimation. On adopte la définition de la VaR conditionnelle basée sur le processus des log-rendements. Soit :

$$\mathbb{P}(r_{t+1} \leq \text{VaR}_\alpha^{t+1|t} \mid \mathcal{F}_t) = \alpha$$

On appelle violation (ou un hit), une situation dans laquelle à la date  $t+1$  le log-rendement observé est inférieur à la valeur prévue  $\text{VaR}_\alpha^{t+1|t}$  (ou d'une manière équivalente, la perte est supérieur à la VaR prévue associée). Pour une fenêtre d'estimation de taille  $\omega$  et un taux de couverture  $\alpha$ , on définit la suite  $(I_t(\alpha))_{t \geq \omega}$  des dépassements de la VaR comme suit :

$$(\forall t \geq \omega) \quad I_t(\alpha) = \mathbb{1}_{\{r_t \leq \text{VaR}_\alpha^{t|t-1}\}}$$

Selon Christoffersen<sup>32</sup>, le problème de validation d'un modèle dans l'estimation de la VaR est réduit à l'étude de la suite de dépassements. Plus précisément, elle doit satisfaire les deux conditions suivantes :

- **Condition de couverture non-conditionnelle** : Chaque jour la probabilité d'avoir un dépassement doit être exactement égale au niveau de couverture  $\alpha$  :

$$\mathbb{P}(I_t(\alpha) = 1) = \mathbb{E}(I_t(\alpha)) = \alpha$$

- **Condition d'indépendance** : La suite des dépassements issues d'un même niveau de risque doit être une suite de variables aléatoires indépendantes.

Plusieurs Backtests de la VaR ont été proposés. Ces tests examinent soit l'une des deux conditions (test du ratio de dépassement, test de couverture non-conditionnelle (Kupiec test)) soit les deux ensemble (le test de couverture conditionnelle). Pour une description de ces tests, le lecteur peut voir : *Une Evaluation des Procédures de Backtesting*. Hurlin C. Tokpavi, S. (2008)

## 5.4 Application à la série des log-rendements du MADEX

La procédure de Backtesting nécessite un nombre très grand d'observations. C'est pour cela qu'on va utiliser tout l'historique disponible sur internet de l'indice MADEX : de 2016-06-20, jusqu'au 2021-06-21. Lorsqu'on travaille avec les log-rendements, on perd une observation, soit donc un nombre totale d'observations égal à 1241. On utilise les 400 dernières observations pour la validation et une fenêtre glissante d'un jour de taille 841 pour l'estimation. Ceci ne représente qu'une simple illustration puisqu'on ne possède pas des données pour des années antérieures à 2016. Dans la réalité, le modèle doit capter tous les dynamiques de la série des log-rendements. Pour ce faire, on doit utiliser un historique suffisamment large (au sens temporel), pour qu'il puisse apprendre tous les scénarios possibles pour la série des log-rendements. Il est à noter, qu'un Backtest est très coûteux d'un point de vue computationnel<sup>33</sup>. On peut généralement utiliser des techniques de *Multiprocessing* pour réduire le temps de calcul. On s'intéresse ici, au calcul de la Value-at-Risk conditionnelle **associée au processus des log-rendements** en utilisant la méthode Variance-Covariance et pour un horizon d'un jour. On rappelle que la VaR est donnée par :

$$\text{VaR}_\alpha^{t+1|t} = \mu_{t+1|t} + \sigma_{t+1|t} q_\alpha$$

Tel que :  $\mu_{t+1|t} = \mathbb{E}(r_{t+1} | \mathcal{F}_t)$ ,  $\sigma_{t+1|t}^2 = \text{Var}(r_{t+1} | \mathcal{F}_t)$  et  $q_\alpha$  est le quantile d'ordre  $\alpha$  de la loi conditionnelle de  $r_{t+1}$ .

On désire comparer les performances de deux modèles : *ARIMA – GARCH* et *EWMA* de RiskMetrics. Le premier étant déjà défini. Le deuxième se présente comme suit :

$$\begin{cases} r_t = \eta_t \sigma_t \\ \sigma_t^2 = \beta \sigma_{t-1}^2 + (1 - \beta) r_{t-1}^2 \end{cases} \quad (\eta)_t \stackrel{\text{i.i.d}}{\sim} \mathcal{N}(0, 1)$$

$\beta$  sera fixée en 0,94. Notons que dans ce modèle, on suppose implicitement que :  $r_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t^2)$

On présentera par la suite, les différents résultats.

32. 'Evaluating interval forecasts', *International economic review* (1998)

33. Pour notre Backtest, on doit ré-estimer le modèle 400 fois

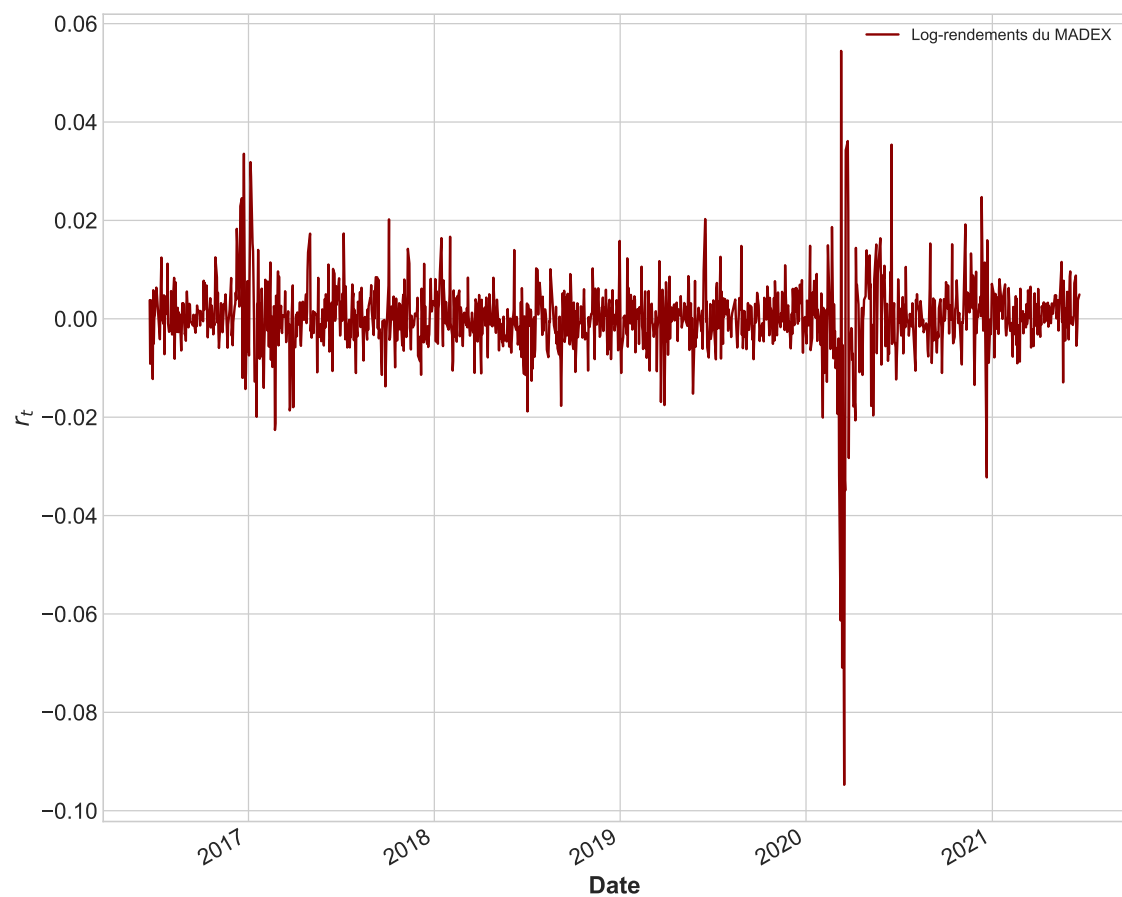


FIGURE 26 – Série des log-rendements du MADEX de 2016-06-21 au 2021-06-21

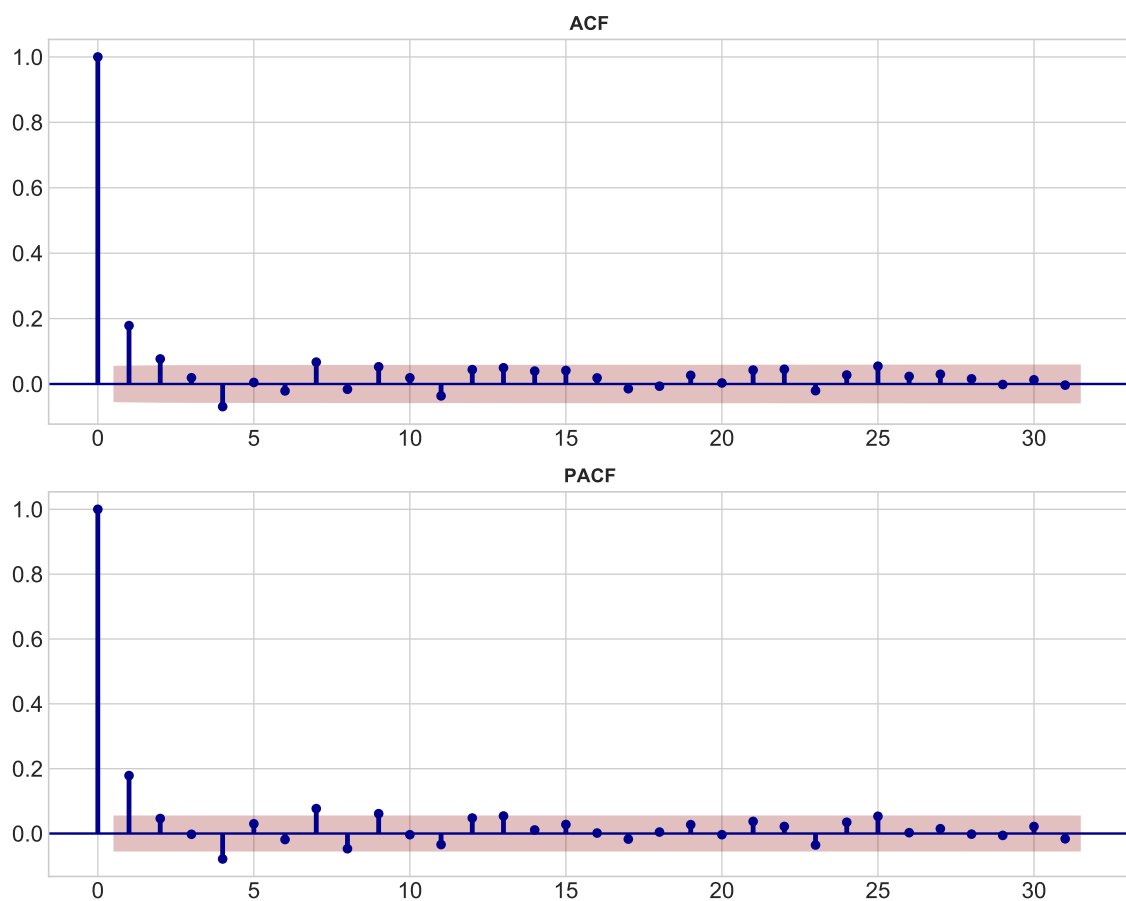


FIGURE 27 – ACF et PACF de la série des log-rendements du MADEX



### 5.4.1 Modélisation

Dep. Variable:	y	No. Observations:	1241
Model:	<b>SARIMAX(2, 0, 0)</b>	Log Likelihood	4266.114
Date:		AIC	-8526.228
Time:		BIC	-8510.857
Sample:	0	HQIC	-8520.448
	- 1241		

	coef	std err	z	P>  z	[0.025	0.975]
ar.L1	0.1713	0.011	15.333	0.000	0.149	0.193
ar.L2	0.0467	0.008	6.163	0.000	0.032	0.062
sigma2	6.047e-05	6.45e-07	93.709	0.000	5.92e-05	6.17e-05

TABLE 7 – Modèle de la moyenne conditionnelle des log-rendements :  $AR(2)$

lag	lb_stat	lb_pvalue
1	0.000895	0.976140
2	0.016389	0.991839
3	0.146474	0.985729
4	8.258474	0.082556
5	8.579726	0.127048
6	9.742240	0.135935
7	16.784816	0.018838

TABLE 8 – Test de Ljung-Box pour les résidus du modèle  $AR(2)$

LM	5.242148e+02
LM p-value	<b>3.194436e-96</b>

TABLE 9 – ARCH-LM test pour les résidus du modèle  $AR(2)$

Distribution	KS statistic	KS p-value
Student (df=3, loc=0.0002, scale=0.0043)	0.0141,	<b>0.96201</b>
Normal (loc=0.0001, scale= 0.0077)	0.0999	2.9580e-11

TABLE 10 – Test de Kolmogorov-Smirnov pour les résidus du modèle  $AR(2)$

Dep. Variable:	None	R-squared:	0.000
Mean Model:	Zero Mean	Adj. R-squared:	0.001
Vol Model:	<b>GARCH</b>	Log-Likelihood:	-1090.08
Distribution:	<b>Standardized Student's t</b>	AIC:	2188.17
Method:	Maximum Likelihood	BIC:	2208.66
		No. Observations:	1241
Date:	Wed, Jun 23 2021	Df Residuals:	1241

	coef	std err	t	P>  t	95.0% Conf. Int.
omega	0.0473	2.037e-02	2.321	2.026e-02	[7.363e-03, 8.720e-02]
alpha[1]	0.1938	6.177e-02	3.137	1.709e-03	[7.269e-02, 0.315]
beta[1]	0.7045	9.440e-02	7.462	8.493e-14	[ 0.519, 0.889]
	coef	std err	t	P>  t	95.0% Conf. Int.
nu	5.0632	0.641	7.894	2.936e-15	[ 3.806, 6.320]

TABLE 11 – Modèle de la variance conditionnelle des log-rendements (résidus du modèle:  $AR(2)$ ) :  $GARCH(1, 1)$

lag	lb_stat	lb_pvalue
1	0.174671	0.675994
2	0.315092	0.854237
3	0.494548	0.920088
4	0.568601	0.966488
5	6.590343	0.252932
6	6.593103	0.360119

TABLE 12 – Test de Ljung-Box pour les résidus du modèle  $GARCH(1, 1)$

LM	22.306460
LM p-value	<b>0.501821</b>

TABLE 13 – ARCH-LM test pour les résidus du modèle  $GARCH(1, 1)$

**Remarque :** Après la ré-estimation du modèle  $AR(2) - GARCH(1, 1)$ , on a trouvé que le coefficient associé à la deuxième composante auto-régressive ( $AR(2)$ ) est significativement nul au risque d'erreur de 5%. **On retient un modèle  $AR(1) - GARCH(1, 1)$**  pour la série des log-rendements du MADEX.

ARFIMA-GARCH Model Fit				
Conditional Variance Dynamics				
GARCH Model : sGARCH(1,1)				
Mean Model : ARFIMA(1,0,0)				
Distribution : std				
	Estimate	Std. Error	t value	Pr(> t )
ar1	0.086297	0.030252	2.8526	0.004336
omega	0.000004	0.000001	3.7696	0.000163
alpha1	0.168253	0.025238	6.6666	0.000000
beta1	0.744594	0.024694	30.1529	0.000000
shape	4.694708	0.523305	8.9713	0.000000

TABLE 14 – Résultats du modèle  $AR(1) - GARCH(1,1)$

#### 5.4.2 Résultats du calcul de la VaR (5%) et du Backtest

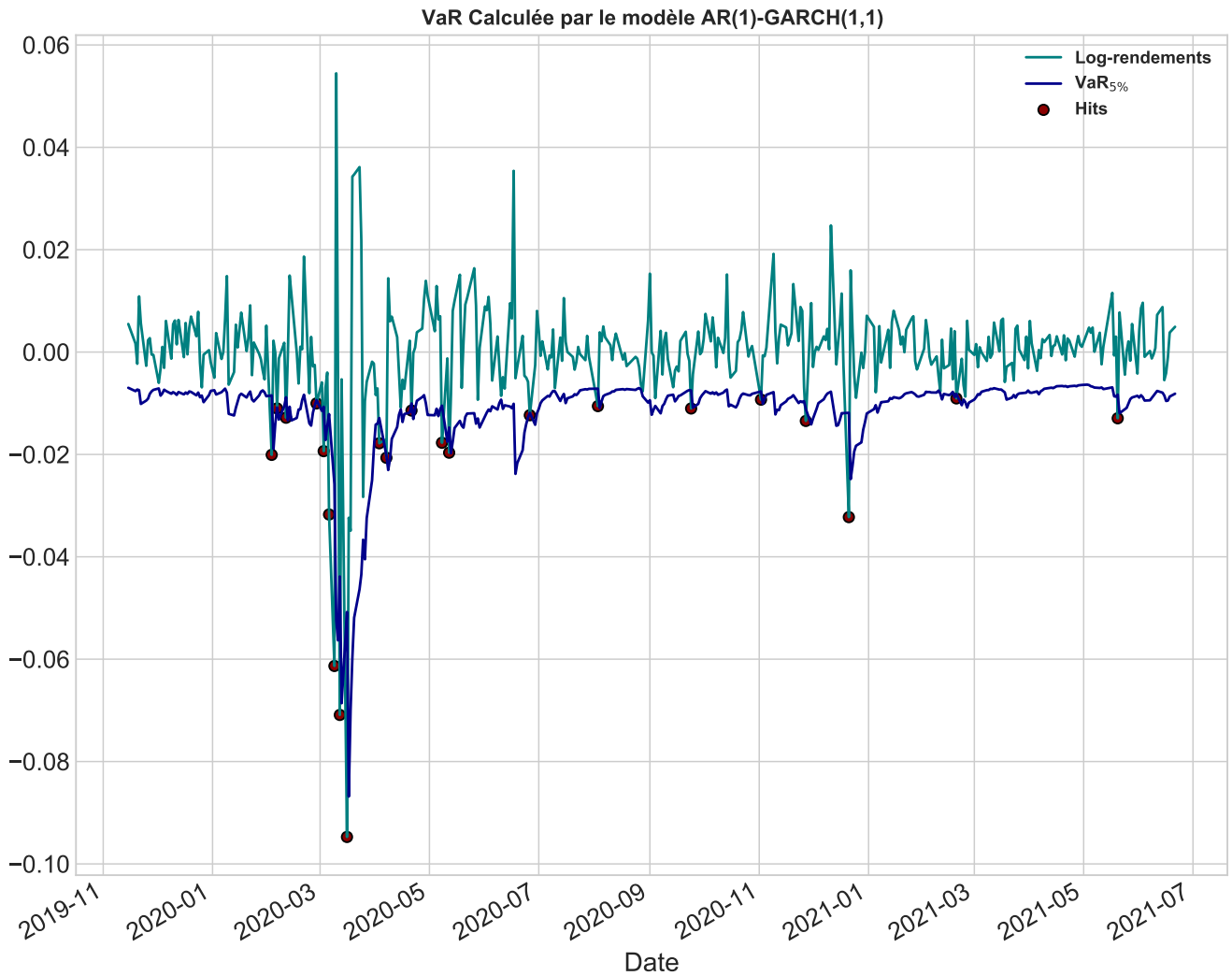


FIGURE 28 – VaR au niveau de 5% calculée par le modèle  $AR(1)$ - $GARCH(1,1)$  et série des dépassements

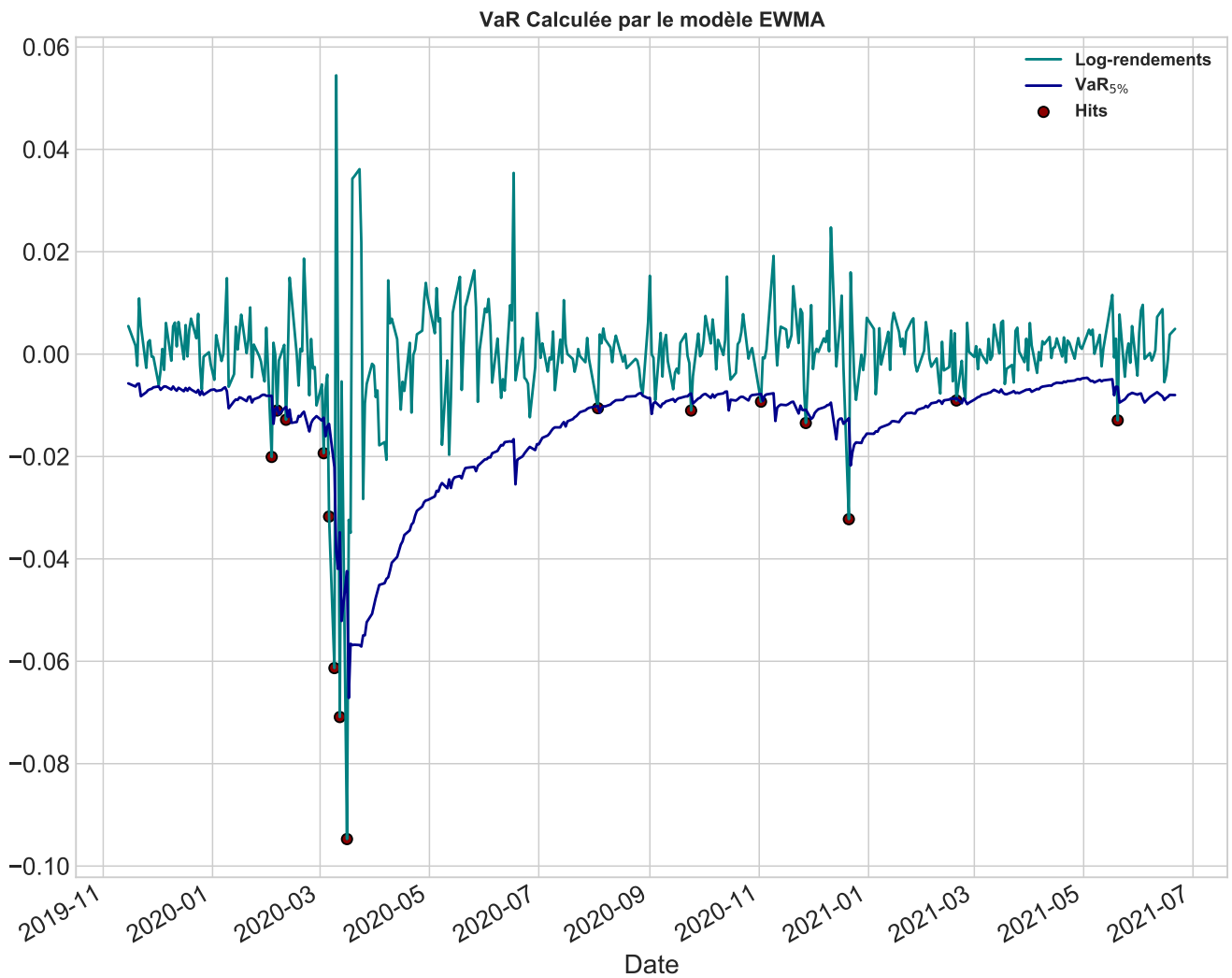


FIGURE 29 – VaR au niveau de 5% calculée par le modèle EWMA et série des dépassements

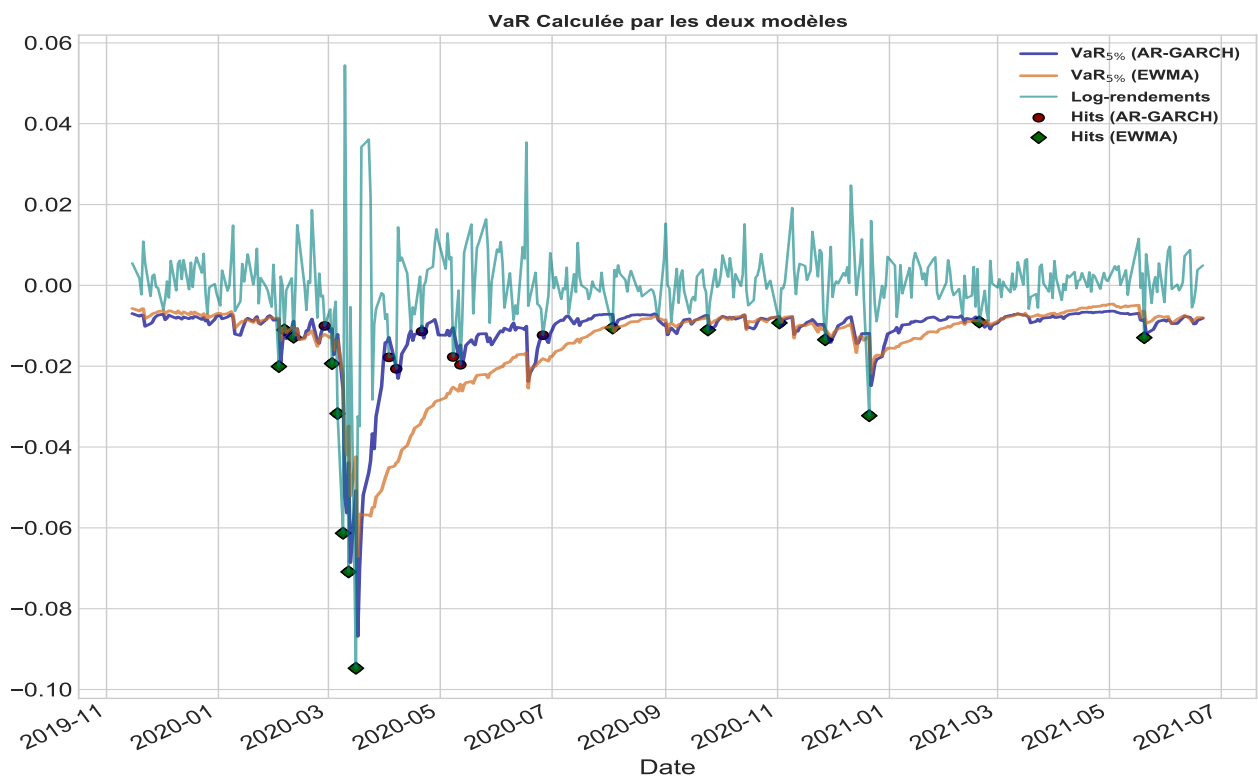


FIGURE 30 – VaR au niveau de 5% calculée par les deux modèles et séries des dépassements

### Ratio de dépassement :

Le ratio de dépassement est défini par :

$$p = \frac{1}{W_T} \sum_{t=W_E+1}^T I_t(\alpha)$$

Tel que :  $W_E$  est la taille de la fenêtre d'estimation,  $W_T$  est la taille de la fenêtre de prévisions et  $T$  désigne le nombre total de données.

D'après la sortie du logiciel, On constate que l'estimation du ratio de dépassement fournie par le modèle GARCH est supérieur à 5%, donc il semble que ce modèle a une tendance à sur-estimer le risque. Par contre, l'estimation fournie par le modèle EWMA est inférieur à 5%, soit donc une sous-estimation du risque. On valide ces constats par l'effectuation des tests d'hypothèses.

AR-GARCH RATIO : 0.055 ; EWMA RATIO : 0.0375

**Test de couverture inconditionnelle (Kupiec test) :** Sous les mêmes notations. On dispose d'un échantillon  $(I_t(\alpha))_{W_E+1 \leq t \leq T}$  i.i.d et on désire tester :

$$\mathcal{H}_0 : \mathbb{E}(I_t(\alpha)) = \alpha$$

Il s'agit d'un test du rapport de vraisemblance et la statistique du test s'écrit :

$$LR_{W_T} = -2W_T \ln \left( \left( \frac{1-\alpha}{1-\hat{p}} \right)^{1-\hat{p}} \left( \frac{\alpha}{\hat{p}} \right)^{\hat{p}} \right)$$

Asymptotiquement, sous  $\mathcal{H}_0$  :

$$LR_{W_T} \sim \chi_1^2$$

Dans notre cas,  $\alpha = 0.05$ . Les résultats de ce test pour les deux modèles sont :

Model	Kupiec-statistic	Kupiec p-value
AR(1)-GARCH(1,1)	0.63350	0.42607
EWMA	4.4522	0.03485

TABLE 15 – Résultats du test de Kupiec pour les deux modèles

On remarque que le modèle  $AR(1)-ARCH(1,1)$  vérifie la condition de couverture non-conditionnelle au risque d'erreur de 5%. Tandis que le modèle EWMA a violé cette hypothèse au seuil de 5%. Les résultats fournis par ce test, ne sont valides que si l'hypothèse d'indépendance des dépassements est vérifiée. Il existe un test d'indépendance en ce sens, introduit par Christoffersen (1998), basé sur les chaînes de MARKOV.

Enfinement, **sous l'hypothèse de l'indépendance des dépassements, Le modèle AR – GARCH a bien performé par rapport au modèle EWMA.**

## 6 Conclusion

Dans la première partie de ce travail, nous avons donné une histoire brève de la naissance de la VaR ainsi que sa définition. Ensuite, nous avons illustré les faits stylisés des séries financières à travers la série des log-rendements du MADEX et nous avons défini le concept de la volatilité ainsi que ces estimateurs.

Nous avons construit, par la suite, un modèle décrivant la dynamique de la série des log-rendements et nous avons conclut quant à la faiblesse de son pouvoir prédictif. Ceci est peut-être dû à la période utilisée dans l'apprentissage, qui a été marquée par la crise du Covid-19.

Dans la dernière partie, nous avons présenté les méthodes standards utilisée dans le Calcul de la VaR, ainsi que la procédure du Backtesting utilisée dans la validation du modèle de calcul. Finalement, nous avons illustré ces méthodes en utilisant la série des log-rendements du MADEX en comparant les performances de deux modèles.

## 7 Bibliographie

- 1 Quantitative Risk Management Concepts, Techniques and Tools Alexander J. McNeil, rüdiger Frey, Paul Embrechts. *Princeton University Press Princeton and Oxford (2005)*
- 2 GARCH Models Structure, Statistical Inference and Financial Applications. Christian Francq. Jean-Michel Zakoïan (2010)
- 3 Time Series and Dynamic Models . by C. Gouriéroux and A. Monfort. (1995)
- 4 Forecasting Volatility. Figlewski (1997)
- 5 Evaluating interval forecasts. Peter F. Christoffersen .*International economic review* (1998)
- 6 Modelling Extremal Events for Insurance and Finance. Embrechts, P., Klüppelberg, C. and Mikosch, T.

## 8 Code Python

Le lecteur pourra trouver le code (VaR\_GARCH.ipynb) et les bases de données utilisées dans le lien suivant :

<https://github.com/OussamaBOUDCHICHI/VaR-estimation-forecast-and-backtest>

Pour une bonne clarté du texte, **il est conseillé de télécharger le fichier 'VaR\_GARCH.html'**.