
Les modèles ARCH et GARCH : Application au CAC40



Réalisé par :

Rim Amiri

Mariem Bouhadda

Encadrement :

Mme. Selma Jelassi

Année universitaire

2018-2019

REMERCIEMENT

Nous tenons à remercier notre encadrante, Mme Selma Jelassi, pour sa présence, son encadrement, ses conseils ainsi que son attention fournies de façon efficace tout au long de la réalisation de ce projet et qui, par son expérience et son savoir-faire nous a permis de surmonter les difficultés et progresser dans notre projet.

Table des matières

1	Introduction	2
2	Étude théorique	3
2.1	Topologies des modèles ARMA	3
2.1.1	Processus autorégressif d'ordre p AR (p)	3
2.1.2	Processus à moyenne mobile MA(q)	3
2.1.3	Modèle ARMA(p,q)	4
2.2	Méthodologie de Box-Jenkins	4
2.2.1	Transformation	4
2.2.2	L'identification du modèle adéquat	5
2.2.3	Estimation des paramètres d'un modèle ARMA	6
2.2.4	Validation	8
2.2.5	Choix d'un modèle parmi plusieurs :	10
2.2.6	Prévision dans les ARMA :	11
2.3	Lissage exponentiel	17
2.3.1	Le lissage exponentiel simple (LES)	17
2.3.2	Le lissage exponentiel double (LED)	19
2.4	Modèles ARCH et GARCH	20
2.4.1	Modèle ARCH	20
2.4.2	Modèle GARCH	22
2.4.3	Estimation et prévision	24
3	Étude Pratique	28
3.1	L'indice boursier CAC 40	28
3.2	Description de la base de données	28

3.3	Etude de la stationnarité	29
3.4	Modélisation	34
3.5	Prévision	45
3.5.1	Prévision dynamique	45
3.5.2	Prévision statique	46
3.5.3	Prévision avec lissage exponentielle	47
4	Conclusion	50

1 Introduction

L'analyse de certaines séries chronologiques émanant du monde économique et financier (taux de change, taux de rendement d'action, indices,...) montre des caractéristiques spécifiques qui ne sont pas théoriquement prises en compte dans les modélisations ARIMA notamment celle de Box Jenkins. La première caractéristique concernée est la dépendance de la variance conditionnelle du temps, autrement dit le comportement hétéroscédastique de la variance (en plus, les périodes de grande variabilité sont groupées, tandis que les petites variations sont plutôt suivies par de petits changements). La deuxième caractéristique concerne la distribution probabiliste de la série (L'apparition de chocs n'est pas compatible avec la loi normale). Alors pour une modélisation plus réaliste de ces séries, les modèles ARCH/GARCH sont les plus utiles. Au cours de cette étude, nous allons d'abord étudier les modèles ARCH et GARCH d'un point de vue théorique où nous allons aborder la majorité des propriétés et des conditions à remplir par les modèles, puis nous allons les appliquer sur notre indice boursier (CAC40) pour pouvoir mener une bonne prévision.

2 Étude théorique

2.1 Topologies des modèles ARMA

2.1.1 Processus autorégressif d'ordre p AR (p)

La suite $X_t : t \geq 0$ est un processus autorégressif d'ordre p ($p > 0$) s'il peut s'écrire sous la forme suivante :

$$X_t = \sum_{k=1}^p \phi_k X_{t-k} + \epsilon_t, \epsilon_t \sim WN(0, (\sigma_\epsilon)^2)$$

Les $\phi_k (k = 1, 2, \dots, p)$ constituent les paramètres du modèle.

Dans ce cas, $X_t \sim AR(p)$. De la même façon, on peut réécrire un processus AR(p) avec un polynôme $\phi(B)$ qui multipliera X_t cette fois-ci :

$$\phi(B)X_t = \epsilon_t \text{ avec } \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

2.1.2 Processus à moyenne mobile MA(q)

La suite $X_t : t \geq 0$ est un processus à moyenne mobile d'ordre q ($q > 0$) si celui-ci peut s'écrire sous la forme suivante :

$$X_t = \sum_{k=1}^q \theta_k \epsilon_{t-k} + \epsilon_t \text{ avec } \epsilon_t \sim WN(0, (\sigma_\epsilon)^2)$$

Les $\theta_k (k = 1, 2, \dots, q)$ constituent les paramètres du modèle.

2.1.3 Modèle ARMA(p,q)

Un processus est dit ARMA(p,q) s'il existe des suites réelles ϕ_k et θ_k telles que :

$$X_t - \sum_{k=1}^p \phi_k X_{t-k} = \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \text{ avec } \epsilon_t \sim WN(0, (\sigma_\epsilon)^2)$$

2.2 Méthodologie de Box-Jenkins

La méthode de Box et Jenkins est un outil systématique qui permet de :

- Déterminer le meilleur modèle de type ARMA décrivant le processus stochastique d'une série observée ou d'une transformation stationnaire de celle-ci ;
- Estimer ce modèle ;
- L'utiliser pour extrapoler les valeurs de la série. La méthodologie de Box et Jenkins comporte essentiellement cinq étapes :

Étape 1 : Transformation des données afin de stabiliser la variance (log, sqrt,...) et différenciation des données pour les stationariser.

Étape 2 : Visualiser les ACF et les PACF empiriques pour identifier les paramètres p et q appropriés.

Étape 3 : Estimation des paramètres du(des) modèle(s) sélectionné(s).

Étape 4 : Diagnostique et tests adéquation du modèle.

Étape 5 : Prévision : La dernière étape consiste la prévision des valeurs futures à travers le modèle retenu.

2.2.1 Transformation

Les modèles ARMA sont assez restrictifs dans la mesure où ils supposent que les observations sont stationnaires, que les paramètres sont constants et que les erreurs sont de variance constante. Pour se ramener à un tel cas, il faut souvent transformer les données. La transformation log permet de stabiliser la variance d'une série si la variance de la série originale croît avec les valeurs de celle-ci. Cette transformation permet aussi parfois de se rapprocher de la normalité. Une série non stationnaire est une série qui en

général croît avec le temps. On peut donc être tenté de retirer par régression un trend temporel de cette série. L'autre solution consiste à utiliser un filtre du type $(1 - L)$. On va alors différencier la série. La série transformée est donc :

$$(1 - L)\log y_t = \log y_t - \log y_{t-1}$$

2.2.2 L'identification du modèle adéquat

La méthode d'identification est essentiellement fondée sur l'analyse conjointe des auto-corrélations et des auto-corrélations partielles. Nous supposons le processus stationnaire ARMA(p,q). La méthode s'appuie sur les résultats suivants :

-Choix de p pour un processus AR(p) :

Les auto-corrélations partielles sont identiquement nulles au delà de l'ordre p.

-Test sur le paramètre p :

Soit $X = (X_t, t \in \mathbb{Z})$ un processus AR(p). Alors pour tout $r > p$, l'autocorrélation partielle empirique $r_X^{(N)}(r)$ est approximativement une gaussienne centrée de variance $1/N$. On en déduit comme précédemment un test de $H_0 : X$ est un AR(p),

pour (presque) tout $r > p$ et $r < N/4$, on a $|r_X^{(N)}(r)| < 1.96\sqrt{\frac{1}{N}}$ alors on accepte H_0 .

-Choix de q pour un processus MA(q) :

Les auto-corrélations sont identiquement nulles au delà de l'ordre q.

-Test sur le paramètre q :

Si $X = (X_t, t \in \mathbb{Z})$ un processus MA(p). Alors pour tout $r > q$, l'autocorrélation empirique $\rho_X^{(N)}(r)$ est approximativement une gaussienne centrée de variance $\frac{1}{N}(1 + 2 \sum_{u=1}^q \rho_X(u)^2) = \frac{1}{N}(1 + 2 \sum_{u=1}^r \rho_X(u)^2)$.

En pratique, on fait donc l'approximation :

$$\forall r > q, \sqrt{N} \frac{\rho_X^{(N)}(r)}{\sqrt{1 + 2 \sum_{u=1}^r \rho_X(u)^2}} \rightsquigarrow N(0, 1).$$

On en déduit un test (pratique mais peu régoureux) de $H_0 : X$ est un MA(q),

si pour (presque) tout $r > p$ et $r < N/4$, on a $|\rho_X^{(N)}(r)| < 1.96\sqrt{\frac{1 + 2 \sum_{u=1}^q \rho_X^{(N)}(u)^2}{N}}$, alors on accepte H_0 .

Si l'on se limite à $r < N/4$, c'est que l'estimation de $\rho_X^{(N)}(r)$ devient de moins en moins

précise quand r croît vers N .

-Choix de p et q pour un processus ARMA(p, q) :

Les auto-corrélations partielles sont identiquement nulles au delà de l'ordre p et les auto-corrélations sont identiquement nulles au delà de l'ordre q .

-Test sur les paramètres (p, q) :

Si X est un processus ARMA($p; q$) vérifiant :

$$\phi(L)(X_t) = \theta(L)(\epsilon_t)$$

avec $\theta(z)$ et $\phi(z)$ sans racine commune, et ϵ_t bruit blanc d'innovation de variance σ^2 alors il peut se mettre sous forme

$$MA(\infty) : X_t = \epsilon_t + \sum_{u \geq 1} \psi_u \epsilon_{t-u}$$

ou $AR(\infty) : X_t + \sum_{v \geq 1} \eta_v X_{t-v} = \epsilon_t$

Comme on a les inégalités :

$$E[(X_t - \epsilon_t - \sum_{u=1}^{q'} \psi_u \epsilon_{t-u})^2] \leq \sigma^2 \left(\sum_{u \geq q'+1} |\psi_u| \right)^2$$

$$E[(X_t - \epsilon_t - \sum_{v=1}^{p'} \eta_v X_{t-v})^2] \leq \gamma_X(0) \left(\sum_{v \geq p'+1} |\eta_v| \right)^2$$

et que $\sum_{u \geq 0} |\psi_u|$ et $\sum_{v \geq 0} |\eta_v|$ sont finis ; on en déduit qu'il est possible d'approcher d'aussi près qu'on le souhaite X par un processus de type $MA(q')$ ou $AR(p')$ avec $p' \leq p$ et $q' \leq q$. En pratique, on choisit pour valeurs de p' et q' celles que l'on détermine par les méthodes vues précédemment, en supposant le processus AR puis MA. Cela fournit une première modélisation de X comme processus ARMA(p', q'), avec en générale plusieurs couples ($p'; q'$) possibles.

2.2.3 Estimation des paramètres d'un modèle ARMA

Une fois que l'on a décidé de chercher pour la série temporelle X_t un modèle ARMA(p, q), reste à estimer les coefficients qui le caractérisent.

Estimation de la partie autorégressive

Si X est un processus ARMA(p,q), vérifiant $\phi(L)X = \theta(L)\epsilon$, alors sa fonction d'autocovariance vérifie la relation de récurrence suivante :

$$\forall t \geq q+1; \gamma_X(t) + \phi_1\gamma_X(t-1) + \dots + \phi_p\gamma_X(t-p) = 0$$

Les estimateurs de Yule Walker de $\phi_1 \dots \phi_p$ sont donc naturellement définis par :

$$\begin{pmatrix} \hat{\rho}(q) & \dots & \hat{\rho}(q-p+1) \\ \vdots & \ddots & \vdots \\ \hat{\rho}(q+p-1) & \dots & \hat{\rho}(q) \end{pmatrix} \begin{pmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_p \end{pmatrix} = - \begin{pmatrix} \hat{\rho}(q+1) \\ \vdots \\ \hat{\rho}(q+p) \end{pmatrix} \quad (1)$$

Estimation de la partie moyenne mobile

(1) Méthode de Durbin :

Soit $Y_t = \phi(L)X_t = \theta(L)\epsilon_t$ un processus MA(q) ou $\theta(z) \neq 0$ pour $|z| < 1$ Alors :

$$\epsilon_t = \sum_{k=1}^{+\infty} \psi_k z^k$$

Il existe p_0 , t.q. $\forall p \geq p_0$ l'équation récurrente :

$$\hat{Y}_t = \epsilon + \sum_{k=1}^p \psi_k X_{t-k}$$

définit un AR(p) causal et on a $E[|Y_t - \hat{Y}_t|^2] \rightarrow 0$ où $p \rightarrow \infty$.

D'où l'idée d'approcher un processus MA inversible par un AR causal suffisamment long.

On choisit p suffisamment grand de telle façon que les observations puissent être considérées comme celles d'un AR(p). On en estime les coefficients $\hat{\psi}_1 \dots \hat{\psi}_p$ et partant de l'identité, $\theta(z)\psi(z) = 1$, on calcule $\hat{\theta}_1 \dots \hat{\theta}_q$ en fonction de $\hat{\psi}_1 \dots \hat{\psi}_p$. En conclusion :

- on choisit p d'autant plus grand que possible.

- on estime les p coefficients ψ_1, \dots, ψ_p , du prédicteur linéaire optimal.

- on estime les q coefficients du prédicteur linéaire optimal obtenu à partir des données ψ_1, \dots, ψ_p .

(2) Méthode du maximum de vraisemblance :

Le processus ARIMA s'écrit :

$$\epsilon_t = X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} - \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

Les ϵ_t sont indépendants et de loi normal $N(0, \sigma^2)$, la fonction de densité de probabilité jointe de ϵ est donné par :

$$p(\epsilon|\phi, \theta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{t=1}^n \epsilon_t^2\right]$$

la fonction de densité de probabilité jointe de X_t peut s'écrire :

$$L(\phi, \theta, \sigma^2|X) = g_1(\phi, \theta, \sigma^2) \exp\left[-\frac{1}{2\sigma^2} S(\phi, \theta)\right]$$

où :

g_1 est la fonction des paramètres (ϕ, θ, σ^2)

Et $S(\phi, \theta) = \sum_{t=1-p-q}^n E^2(u_t|X)$

$E(u_t|X) = E^2(u_t|X, \phi, \theta, \sigma^2)$ est l'espérance conditionnelle de u_t sur X, ϕ, θ et σ^2 et

$$u_t = \begin{cases} \epsilon_t & t = 1, 2, \dots, n \\ g_2(\epsilon_*, X_*) & t \leq 0 \end{cases}$$

où :

g_2 est une fonction linéaire des valeurs initiales $a_* = (a_{1-p}, \dots, a_1, a_0)$

Et $X_* = (X_{1-p}, \dots, X_{-1}, X_0)$

2.2.4 Validation

Il s'agit de vérifier notamment que les résidus du modèle ARMA estimé, résidus notés $\hat{\epsilon}_t$, vérifient les propriétés requises pour que l'estimation soit valide, à savoir qu'ils suivent un processus BB, non autocorrélé et de même variance, et qu'ils suivent une loi normale.

Si ces hypothèses ne sont pas rejetées, on peut alors mener des tests sur les paramètres.

Tests sur les résidus

Test de bruité :

(1) Par l'étude de la fonction d'auto-corrélation empirique : Lorsque n est assez grand, les auto-corrélations d'un bruit blanc sont approximativement indépendantes et de loi $N(0, \frac{1}{n})$. Ainsi, 95% des auto-corrélations devraient se trouver dans l'intervalle $[\frac{-1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}}]$, et en traçant les 40 premières auto-corrélations il ne devrait pas y en avoir plus de 2 voir 3 en dehors de ces limites. ie. les autocorrélations simples et partielles. Elles doivent être significativement nulles si les résidus sont un bruit blanc.

(3) Test de normalité : Il s'agit de tester que les résidus estimés $\hat{\epsilon}_t$ suivent une loi normale, c'est-à-dire ne présentent pas d'asymétrie (Skewness) ni d'aplatissement (kurtosis).

Le coefficient de Skewness est donné par :

$$\beta_1^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$$

et le coefficient de kurtosis est donné par :

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

où

$$\mu_k = \frac{1}{T} \sum_{t=1}^T (\hat{\epsilon}_t - \bar{\hat{\epsilon}})^k$$

Si la distribution est normale et le nombre d'observations est grand, alors :

$$\beta_1^{1/2} \sim \mathcal{N}(0, \sqrt{\frac{6}{T}}) \text{ et } \beta_2 \sim \mathcal{N}(3, \sqrt{\frac{24}{T}})$$

On construit alors les statistiques :

$$v_1 = \frac{\beta_1^{1/2}}{\sqrt{\frac{6}{T}}} \text{ et } v_2 = \frac{\beta_2 - 3}{\sqrt{\frac{24}{T}}}$$

qui suivent chacune une $N(0,1)$. Le test de Jarque Bera permet de tester simultanément l'absence d'asymétrie et l'absence d'aplatissement. La statistique de test est donnée par :

$$JB = \frac{T}{6} \beta - 1 + \frac{T}{24} (\beta_2 - 3)^2$$

Cette statistique suit, sous l'hypothèse nulle de normalité, une loi du χ^2_2 .

(3) Test de stationnarité (test ADF) : Dickey et Fuller (1979) proposent un test détectant la non stationnarité d'une série temporelle. En considérant une série temporelle notée y_t , le test de Dickey et Fuller (DF) est un test de racine unitaire qui estime l'hypothèse nulle de racine unitaire (ou de non stationnarité). Le test ADF consiste alors à estimer les modèles en introduisant des variables retardées. Par exemple, le modèle sans constante ni dérive temporelle est le modèle suivant :

$$\Delta y_t = \phi y_{t-1} + \sum_{j=2}^p \beta_j - \Delta y_{t-j+1} + \epsilon_t.$$

L'hypothèse nulle du test ADF est l'hypothèse de racine unitaire (ou non stationnarité) de la variable y_t soit l'hypothèse $H_0 : \phi = 0$.

Le test ADF consiste à comparer la valeur estimée du t de student associé au paramètre ϕ aux valeurs tabulées de cette statistique. L'hypothèse nulle est rejetée au seuil de 5 % lorsque la valeur observée du t de student est inférieure à la valeur critique tabulée où $t < ADF_{0.5}$.

Tests sur les paramètres : On vérifie tout d'abord que les racines des polynômes AR et MA ne sont pas égales à 1. Il s'agit de tester $H_0 : p' = p - 1$ et $q' = q$ ou $H_0 : p' = p$ et $q' = q - 1$. En d'autres termes, on teste : $H_0 = ARMA(p - 1, q)$ vs $H_1 : ARMA(p, q)$. On teste la significativité des retards du modèle ARMA par des tests de Student.

2.2.5 Choix d'un modèle parmi plusieurs :

Comme on l'a vu précédemment l'étude des auto-covariances, auto-corrélations et auto-corrélations partielles peut conduire à certaines hypothèses sur la nature du modèle. Une fois quelques modèles choisis, et leur paramètres estimés, on utilise les critères suivants pour choisir le meilleur modèle :

Critères d'information Pour choisir le modèle qui effectue le meilleur compromis entre :

- ajustement à la série de données.
- complexité du modèle.

Il est en effet très important de prendre en compte ce compromis, car si on ne s'intéressait qu'à coller au mieux aux données, on serait tenté de choisir un modèle ARMA avec un très grand nombre de paramètres. Or, plus il y a de paramètres, plus il faut de données pour les estimer. Et donc pour un nombre d'observations fixé de la série, plus le modèle sera complexe, moins bien seront estimés les paramètres. Les critères de choix de modèles les plus courants sont :

- le critère AIC (Akaike Information Criterion), qui sera généralement préféré si l'objectif de l'étude est de faire de la prévision, et qui est défini par :

$$AIC(p, q) = -2\log L + 2(p + q)$$

- le critère BIC (Bayesian Information Criterion) sera quant à lui généralement préféré si l'objectif de l'étude est de s'ajuster à la série observée, et est défini par :

$$BIC(p, q) = -2\log L + (p + q)\log T$$

où $\log L$ est la log-vraisemblance du modèle ARMA(p,q) estimé et T est le nombre d'observations. Les modèles ayant la plus petite valeur du critère devront être choisis. Ces deux critères conduisent donc à sélectionner des modèles dont la vraisemblance est grande, en la pénalisant par la complexité du modèle.

Principe de parcimonie :

Lorsqu'on veut modéliser une série chronologique, par un processus Stochastique et dans le cas où les critères d'information AIC et BIC de deux ou plusieurs modèles retenus seraient très proches ou contradictoires, nous faisons intervenir ce principe qui cherche à minimiser le nombre de paramètre requis ; il est préférable de conserver un modèle qui est moins bon, mais qui contient moins de paramètres.

2.2.6 Prévision dans les ARMA :

On suppose $X_t \rightsquigarrow ARMA : \phi(L)X_t = \mu + \theta(L)\epsilon_t$ de représentation canonique minimal. On est à la date t et on cherche les meilleurs prévisions possibles pour $(x_{t+1}, \dots, x_{t+h})$, prévision à l'ordre h .

La prévision optimal de x_{t+h} à la date t est $EL(X_{t+h}|\underline{X}_t)$ (la projection de X_{t+h} sur X_t, X_{t-1}, \dots) : On suppose que les coefficients de ϕ et de θ , ainsi que μ sont connus. En pratique, ils seront remplacés par leurs estimateurs.

•Prévision dans un AR(p)

(Prévision optimale). On note : ${}_tX_{t+h}^* = EL(X_{t+h}|\underline{X}_t)$

Soit : $X_t = \mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \epsilon_t$ la représentation canonique.

On a :

(i) Au rang 1 :

$$X_{t+1} = \mu + \varphi_1 X_t + \dots + \varphi_p X_{t+1-p} + \epsilon_{t+1}$$

$${}_tX_{t+1}^* = EL(X_{t+1}|\underline{X}_t) = \mu + \varphi_1 X_t + \dots + \varphi_p X_{t+1-p}$$

(ii) Au rang 2 :

$$X_{t+2} = \mu + \varphi_1 X_{t+1} + \dots + \varphi_p X_{t+2-p} + \epsilon_{t+2}$$

$${}_tX_{t+2}^* = EL(X_{t+2}|\underline{X}_t) = \mu + \varphi_1 X_{t+1} + \dots + \varphi_p X_{t+2-p}$$

(iii) Au rang h :

$$X_{t+h} = \mu + \varphi_1 X_{t+h-1} + \dots + \varphi_p X_{t+h-p} + \epsilon_{t+h}$$

$${}_tX_{t+h}^* = \mu + \sum_{k=1}^{h-1} \varphi_k {}_tX_{t+k}^* + \sum_{k=h}^p \varphi_k X_{t+h-k}$$

Cette prévision s'obtient donc par itérations successives.

On peut aller plus vite si on sait résoudre les équations de récurrence :

$$\phi(L)X_t = \mu + \epsilon_t \Leftrightarrow \phi(L)(X_t - m) = \epsilon_t, \mu = \frac{m}{\phi(1)}$$

On pose $Y_t = (X_t - m)$, et on a : ${}_tY_{t+k}^* = {}_tX_{t+k}^* - m$ $\phi(L)Y_{t+h} = \epsilon_{t+h} \Rightarrow \phi(L){}_tY_{t+k}^* = 0$ avec la notation précédente.

D'où l'équation de récurrence d'ordre p :

$$\phi(L){}_tY_{t+k}^* = {}_tY_{t+k}^* - \sum_{k=1}^p \varphi_k {}_tY_{t+h-k}^* = 0$$

Son polynôme caractéristique est $z^p \phi(\frac{1}{z})$. On obtient ainsi les ${}_tY_{t+k}^*$ à partir de la forme générale de la solution de l'équation de réccurence.

•Prévision dans un MA(q)

$$X_t = m + \theta(L)\epsilon_t = m + \epsilon_t - \theta_1\epsilon_{t-1} - \dots - \theta_q\epsilon_{t-q}$$

Cas $h > q$

$$X_{t+h} = m + \epsilon_{t+h} - \theta_1\epsilon_{t+h-1} - \dots - \theta_q\epsilon_{t+h-q}$$

avec $t+h > t+h-1 > \dots > t+h-q > t$.

D'où :

$${}_tX_{t+h}^* = EL(X_{t+h}|\underline{X}_t) = EL(m + \epsilon_{t+h} - \theta_1\epsilon_{t+h-1} - \dots - \theta_q\epsilon_{t+h-q}|\underline{X}_t)$$

et donc : ${}_tX_{t+h}^* = m$

Cash $\leq q$

$$\begin{aligned} {}_tX_{t+h}^* &= EL(X_{t+h}|\underline{X}_t) \\ &= EL(m + \epsilon_{t+h} - \theta_1\epsilon_{t+h-1} - \dots - \theta_h\epsilon_t \dots - \theta_q\epsilon_{t+h-q}|\underline{X}_t) \\ &= m + 0 - \theta_h\epsilon_t \dots - \theta_q\epsilon_{t+h-q} \end{aligned}$$

Cette forme théorique est inutilisable dans la pratique, faute de pouvoir observer les ϵ_t .

On va donc utiliser la forme $AR(\infty)$:

$$\theta(L)^{-1}(X_t - m) = \epsilon_t \Rightarrow \theta(L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k, a_0 = 1, \sum |a_k| < \infty$$

donc

$$\sum_{k=0}^{+\infty} a_k X_{t-k} - \mu - \epsilon_t \Leftrightarrow X_t = \mu - \sum_{k=1}^{+\infty} a_k X_{t-k} + \epsilon_t; \mu = \frac{m}{\theta(1)}$$

D'où

(i) $h = 1$:

$$X_{t+1} = \mu - \sum_{k=1}^{+\infty} a_k X_{t-k+1} + \epsilon_{t+1}$$

$${}_tX_{t+1}^* = \mu + \sum_{k=1}^{+\infty} a_k X_{t-k+1} + 0$$

(ii) $h = 2$:

$$X_{t+2} = \mu - a_1 X_{t+1} - \sum_{k=2}^{+\infty} a_k X_{t-k+2} + \epsilon_{t+2}$$

$${}_tX_{t+2}^* = \mu - a_1 {}_tX_{t+1}^* - \sum_{k=2}^{+\infty} a_k X_{t-k+2} + 0$$

(iii) Cas général :

$$X_{t+h} = \mu - \sum_{k=1}^{+\infty} a_k X_{t-k+h} + \epsilon_{t+h}$$

$${}_tX_{t+h}^* = \mu + \sum_{k=1}^{+\infty} a_k {}_tX_{t+h-k}^* + 0$$

$${}_tX_{t+h}^* = \mu - \sum_{k=1}^{h-1} a_k {}_tX_{t+h-k}^* - \sum_{k=h}^{+\infty} a_k X_{t-k+h}$$

Prévision pour un ARMA(p,q)

a.Prévision optimale :

On utilise la forme AR(1) :

$$\phi(L)X_t = \mu + \theta(L)\epsilon_t$$

$$\phi(L)(X_t - \mu) = \theta(L)\epsilon_t$$

$$\theta(L)^{-1}\phi(L)(X_t - \mu) = \epsilon_t$$

$$\theta(L)^{-1}\phi(L)(X_t) = \mu_0 + \epsilon_t \quad \mu_0 = \frac{\mu}{\theta(1)}$$

D'où :

$$\sum_{k=0}^{+\infty} a_k X_{t-k} = \mu_0 + \epsilon_t \Leftrightarrow X_t = - \sum_{k=1}^{+\infty} a_k X_{t-k} + \mu_0 + \epsilon_t$$

On va donc faire comme dans le cas d'un MA(q), sauf que les prévisions ne seront pas constantes à partir d'un certain rang :

$${}_tX_{t+h}^* = \mu_0 - \sum_{k=1}^{+\infty} a_k {}_tX_{t+h-k}^* = \mu_0 - \sum_{k=1}^{h-1} a_k {}_tX_{t+h-k}^* - \sum_{k=h}^{+\infty} a_k {}_tX_{t+h-k}$$

La prévision optimale se calcule de manière itérative.

b.Prévision approchée

De même que pour le MA(q), si $t > 0$,

$${}_t\hat{X}_{t+h} = \mu_0 - \sum_{k=1}^{t+h} a_k {}_t\hat{X}_{t+h-k}$$

$${}_t\hat{X}_{t+h} = \mu_0 - \sum_{k=1}^{h-1} a_k {}_t\hat{X}_{t+h-k} - \sum_{k=h}^{t+h} a_k {}_tX_{t+h-k}$$

On le calcule par itérations sur h.

c. Equation de récurrence

$$\phi(L)(X_t - m) = \theta(L)\epsilon_t, \quad m = \frac{\mu}{\theta(1)}$$

$$\phi(L)(Y_t) = \theta(L)\epsilon_t \Leftrightarrow Y_t = -(\varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \epsilon_t - (\theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}))$$

Si $h \geq q$,

$$Y_{t+h} = -\varphi_1 Y_{t+h-1} - \dots - \varphi_p Y_{t+h-p} + \underbrace{\epsilon_{t+h}}_{>t} - \theta_1 \underbrace{\epsilon_{t+h-1}}_{>t} - \dots - \theta_q \underbrace{\epsilon_{t+h-q}}_{>t}$$

Donc,

$$\begin{aligned} {}_t Y_{t+h}^* &= EL(Y_{t+h} | \underline{Y}_t) = EL(X_{t+h} | \underline{X}_t) \\ \Rightarrow {}_t Y_{t+h}^* &= -\varphi_1 {}_t Y_{t+h-1}^* - \dots - \varphi_q {}_t Y_{t+h-q}^* + 0 \end{aligned}$$

Les ${}_t Y_{t+h}^*$ vérifient donc l'équation de récurrence : $\phi(L)({}_t Y_{t+h}^*) = 0$ de polynôme caractéristique $z^p \phi(\frac{1}{z})$. Ils peuvent être explicités avec p conditions initiales (observations ou prévisions) : On a donc ${}_t X_{t+h}^* = {}_t Y_{t+h}^* + m$ vérifient la même équation de récurrence.

2.3 Lissage exponentiel

Introduites par Holt en 1958, Winters en 1960 et popularisées par le livre de 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 temporelle.

2.3.1 Le lissage exponentiel simple (LES)

Modèle considéré : $X_t = a + \epsilon_t$

Soit $0 < \beta < 1$, on cherche la meilleure (au sens des MC pondérés) prévision constante $\hat{X}_T(h)$ i.e. la solution de

$$\min_a \sum_{j=0}^{T-1} \beta^j (X_{T-j} - a)^2$$

Definition :

La prévision de la série à l'horizon h , $\hat{X}_T(h)$, fournie par la méthode de lissage exponentiel simple est donnée par :

$$\hat{X}_T(h) = (1 - \beta) \sum_{j=0}^{T-1} \beta^j X_{T-j}$$

où β est la constante de lissage.

Remarques sur le LES :

1. si β est indépendant de h , on note \hat{X}_T au lieu de $\hat{X}_T(h)$.
2. Cette méthode prend en compte tout le passé de la série chronologique mais en accordant de moins en moins d'importance aux observations les plus éloignées de l'instant T .
3. La valeur de β permet de nuancer la remarque précédente :
 - β proche de 0, prévision souple i.e. fortement influencée par les observations les plus récentes.
 - $\beta = 0$, la prévision est alors égale à la dernière valeur observée.
 - β est proche de 1, prévision rigide i.e. l'influence des observations passées est d'autant

plus importante et remonte loin dans le passé.

- $\beta = 1$, alors toutes les prévisions sont identiques. En pratique, on prend $\beta \in]0, 1[$ afin d'exclure ces deux cas extrêmes dégénérés.

LES - Formules de mise à jour :

A partir de la définition, on obtient la formule de mise à jour :

$$\hat{X}_T(h) = \beta \hat{X}_{T-1}(h) + (1 - \beta) X_T = \hat{X}_{T-1}(h) + (1 - \beta)(X_T - \hat{X}_{T-1}(h))$$

Remarques :

1. Première égalité : $\hat{X}_T(h)$ = barycentre entre $\hat{X}_{T-1}(h)$. La valeur prédite à l'horizon h à partir des $T-1$ premières observations et X_T la dernière observation.
2. Deuxième égalité fait intervenir $(X_T - \hat{X}_{T-1}(h))$ la dernière erreur de prévision.
3. Initialisation de la récurrence par $\hat{X}_1(h) = X_1$ ou la moyenne des observations.

LES - Choix du β :

Un problème important en pratique est le choix du β qui est en général très subjectif et varie selon le contexte de l'étude et/ou le type de prévision souhaitée. En pratique,

.si on veut une prévision rigide, on choisira $\beta \in [0.7; 0.99]$.

.si on veut une prévision souple, on choisira $\beta \in [0.01; 0.3]$.

.si on ne sait pas, une autre solution, dictée par les données, consiste à choisir β comme la solution.

$$\sum_{t=1}^{T-h} (X_{t+h} - \hat{X}_t(h))^2 = (X_{1+h} - \hat{X}_1(h))^2 + \dots + (X_T - \hat{X}_{T-h}(h))^2$$

Afin de ne pas tenir compte de l'initialisation, on minimise seulement :

$$\sum_{t=[(T-h)/2]}^{T-h} (X_{t+h} - \hat{X}_t(h))^2$$

2.3.2 Le lissage exponentiel double (LED)

Modèle considéré : $X_t = a(t - T) + b + \epsilon_t$

On cherche donc une prévision à l'horizon h , $\hat{X}_T(h)$ de la forme

$$\hat{X}_T(h) = \hat{a}_T(h) + \hat{b}_T,$$

où le couple (\hat{a}_T, \hat{b}_T) minimise la fonction

$$\sum_{j=0}^{T-1} \beta^j (X_{T-j} - (aj + b))^2$$

En notant la série lissée S_1 et la série doublement lissée S_2 définies par

$$S_1(t) = (1 - \beta) \sum_{j=0}^{T-1} \beta^j X_{T-j}, \quad S_2(t) = (1 - \beta) \sum_{j=0}^{T-1} \beta^j S_1(T - j)$$

LED - Définition

La prévision de la série à l'horizon h , $\hat{X}_T(h)$, fournie par la méthode de lissage exponentiel double est donnée par

$$\hat{X}_T(h) = \hat{a}_T(h) + \hat{b}_T,$$

où β est la constante de lissage et le couple (\hat{a}_T, \hat{b}_T) est donné par

$$\begin{cases} \hat{a}_T = \frac{1-\beta}{\beta} (S_1(T) - S_2(T)) \\ \hat{b}_T = 2S_1(T) - S_2(T) \end{cases}$$

LED - Formules de mise à jour

Les formules de mise à jour s'obtiennent à partir de ces expressions

$$\hat{a}_T = \hat{a}_{T-1} + (1 - \beta)^2 (X_T - \hat{X}_{T-1}(1)) \quad \text{et} \quad \hat{b}_T = \hat{b}_{T-1} + \hat{a}_{T-1} + (1 - \beta)^2 (X_T - \hat{X}_{T-1}(1))$$

Remarques :

1. Pour l'initialisation, on prend $\hat{a}_2 = X_2 X_1$ et $\hat{b}_2 = X_2$.
2. On peut généraliser cette technique de lissage pour traiter des séries sans saisonnalité présentant une tendance polynomiale de degré supérieur à 2. Les résultats font intervenir, dans ce cas, les opérateurs de lissage d'ordre p $S_p(t)$, $p \in N$, itérées d'ordre p de $S_1(t)$.

2.4 Modèles ARCH et GARCH

les modèles ARCH et GARCH (Autoregressive Conditional Heteroscedasticity and Generalized Autoregressive Conditional Heteroscedasticity) sont des modèles qui jouent un rôle primordial au niveau de la description des séries financières vu le comportement hétéroscédastique de leur variance, chose qui a toujours été mal considérée par les modèles ARIMA dans lesquels on suppose que la variance est inconditionnelle par rapport au temps. En effet, face aux anomalies des représentations ARMA pour les problèmes monétaires et financiers, Engle (1982) [1, 2, 3,6] a mis, à la disposition de l'ensemble des acteurs du marché financier, une nouvelle catégorie de modèles autorégressifs conditionnellement hétéroscédastiques aptes à capter le comportement de la volatilité dans le temps. Cette dernière est un paramètre de mesure du risque du rendement et du prix. La volatilité sert également aux calculs pour optimiser la diversification des portefeuilles d'actifs financiers (MEDAF). Les séries monétaires et financières sont caractérisées par le regroupement (clustering) de la volatilité, à savoir les périodes de forte volatilité alternent avec les périodes de faible volatilité. Ce phénomène, que nous appelons aussi l'hétéroscédasticité conditionnelle.

2.4.1 Modèle ARCH

Commençant tout abord à présenter le modèle ARCH(1) introduit par Engle (1982).

Le processus x_t est un processus ARCH (1) si $x_t = \sqrt{h_t}\epsilon_t$ tel que ϵ_t est un bruit blanc gaussien, $\epsilon_t \sim N(0, 1)$ et $h_t = w_0 + w_1 x_{t-1}^2$ la variance conditionnelle du processus x_t .

Les moments conditionnels se représentent comme suit :

$$E[x_t/h_t] = E[\sqrt{h_t}\epsilon_t/h_t] = \sqrt{h_t}E[\epsilon_t/h_t] = 0$$

La variance conditionnelle s'exprime de la manière suivante :

$$V[x_t/h_t] = V[\sqrt{h_t}\epsilon_t/h_t] = h_t V[\epsilon_t/h_t] = h_t E[\epsilon_t^2/h_t] = h_t$$

On détermine l'espérance non conditionnelle du processus x_t comme suit :

$$E[x_t] = E[E[x_t/h_t]] = E[0] = 0$$

La variance du processus à l'instant t est donnée par :

$$V[x_t] = \frac{w_0}{1 - w_1}$$

De même la covariance de ce processus est définie par :

$$cov[x_t, x_{t+h}/x_{t-m}] = 0 \quad \forall h > 0 \quad \forall m > 0$$

Un processus ARCH (p) est un processus défini par :

$$x_t = \sqrt{h_t} \epsilon_t$$

$$h_t = w_0 + \sum_{i=1}^p w_i x_{t-i}^2$$

L'espérance du processus est donnée par :

$$E[x_t] = 0$$

De ce qui est la variance, il suffit de définir une formule de récurrence de façon à exprimer explicitement cette variance en passant par le calcul de la limite. A ce stade on ne ferait pas les calculs, on se contente de fournir le résultat et de supposer que si $|\sum_{i=1}^p w_i| < 1$ alors la variance du processus existe et elle s'exprime par la formule suivante :

$$V[X_t] = \frac{w_0}{1 - \sum_{i=1}^p w_i}$$

Condition de stationnarité

Un processus ARCH(p) est dit stationnaire si :

$$\sum_{i=1}^p w_i < 1$$

2.4.2 Modèle GARCH

Pour de nombreuses applications, l'introduction d'un grand nombre de retards p dans l'équation de la variance conditionnelle du modèle ARCH (p) est nécessaire pour tenir compte de la longue mémoire de la volatilité qui caractérise certaines séries monétaires et financières . Ce nombre important de paramètres peut conduire à la violation de la contrainte de non-négativité de la variance et poser des problèmes d'estimations. Dans cette perspective, une extension importante, le modèle autorégressif conditionnellement hétéroscédastique généralisé (GARCH), est suggéré par Bollerslev [1986]. Cette approche exige moins de paramètres à estimer que la formulation ARCH(p) pour modéliser les phénomènes de persistance des chocs. La variance conditionnelle de la variable étudiée est déterminée par le carré des p termes d'erreur passés et des q variances conditionnelles retardées.

Un processus GARCH (1,1) s'écrit de la forme :

$$x_t = \sqrt{h_t} \epsilon_t$$

avec $\epsilon_t \sim N(0, 1)$ et $h_t = w_0 + w_1 x_{t-1}^2 + \beta_1 h_{t-1}$ représente la variance conditionnelle du processus.

Ainsi, on fournit les moments conditionnels :

$$E[x_t/F_t] = 0$$

où F_t représente la filtration engendrée par les valeurs passées de x_t, x_t^2 et h_t . De ce fait, la variance conditionnelle du modèle sera alors :

$$V[x_t/F_t] = V[\sqrt{h_t} \epsilon_t/F_t] = h_t V[\epsilon_t/F_t] = h_t$$

Pour ce qui est des moments non conditionnels, l'espérance ainsi que la variance du processus s'obtiennent de la façon suivante :

$$E[x_t] = E[E[x_t/h_t]] = E[0] = 0$$

De même :

$$V[x_t] = \frac{w_0}{1 - (w_1 + \beta_1)}$$

Un modèle GARCH (p, q) s'écrit de la façon suivante :

$$x_t = \sqrt{h_t} \epsilon_t$$
$$h_t = w_0 + \sum_{i=1}^p w_i x_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}$$

L'espérance s'obtient simplement à partir de la loi des espérances itérées :

$$E[x_t] = E[E[x_t/h_t]] = 0$$

Si le processus GARCH (p, q) est stationnaire au second ordre, on aura alors :

$$\sum_{i=1}^p w_i + \sum_{j=1}^q \beta_j < 1$$

Cette condition s'avère nécessaire pour définir la variance non conditionnelle par :

$$V[x_t] = \frac{w_0}{1 - (\sum_{i=1}^p w_i + \sum_{j=1}^q \beta_j)}$$

Conditions de stationnarité

Un processus GARCH (p, q) est dit stationnaire si :

$$\sum_{i=1}^p w_i + \sum_{j=1}^q \beta_j < 1$$

Les conditions de normalité

Pour qu'une série chronologique soit normale, elle doit vérifier les conditions suivantes :
Un Skewness nul (moment centré d'ordre 3). Si ce n'est pas le cas, la distribution sera donc asymétrique.

Un Kurtosis égale 3 (moment centré d'ordre 4). Si cette statistique est supérieure à 3 la distribution sera plus aplatie qu'une gaussienne.

Afin de déterminer la normalité des données, on fait recours au test d'hypothèse suivant :
 H_0 : les données suivent une loi normale VS H_1 : les données ne suivent pas une loi normale avec $JB = \frac{n-k}{6}(S^2 - (\frac{k-3}{4})) \sim \chi^2$ (la statistique de Jarque Bera).

n = Nombre d'observations.

k = Nombre de variables explicatives si les données proviennent des résidus d'une régression linéaire. Sinon, $k = 0$.

S = Skewness : Moment d'ordre 3 d'une variable centrée-réduite.

K = Kurtosis : Moment d'ordre 4 d'une variable centrée-réduite.

Une loi normale a un coefficient d'asymétrie = 0 et un Kurtosis = 3. On saisit alors que si les données suivent une loi normale, le test s'approche alors de 0 et on accepte (on ne rejette pas) H_0 au seuil α .

2.4.3 Estimation et prévision

La méthode de maximum de vraisemblance

Pour comprendre cette approche, nous allons tout d'abord, considérer le cas le plus simple d'un processus ARCH pour Y_t . Nous étudierons ensuite le cas des processus GARCH, et enfin des modèles de régression avec erreur (G)ARCH.

L'estimateur des paramètres de modèle ARCH se base très souvent sur la maximisation de la fonction de vraisemblance. Nous supposons que le processus Y_t est conditionnellement gaussien.

La vraisemblance associée à Y_t conditionnellement au passé I_{t-1} est donc :

$$L(y_t/I_{t-1}) = \frac{1}{h_t\sqrt{2\pi}} \exp\left(-\frac{(y_t - m_t(\theta))^2}{2h_t^2(\theta)}\right), \theta = (\alpha_0, \dots, \alpha_p)$$

La fonction de vraisemblance de $(y_1; y_2; \dots; y_T)$ conditionnelle est par conséquent :

$$L(y_1, y_2, \dots, y_T; \theta) = \prod_{i=1}^T L(y_i/I_{i-1}, \theta)$$

L'estimateur est alors défini comme le vecteur

$$\hat{\theta}_T = (\hat{\alpha}_0; \dots; \hat{\alpha}_p)$$

qui maximise le logarithme de cette fonction vraisemblance :

$$\hat{\theta}_T = \operatorname{argmax} \log L(y_1; y_2; \dots; y_T)$$

Estimation des paramètres du modèle GARCH

La vraisemblance associée Y_t conditionnellement au passé I_{t-1} s'écrit :

$$L(y_t/I_{t-1}) = \frac{1}{h_t\sqrt{2\pi}} \exp\left(-\frac{(y_t - m_t(\theta))^2}{2h_t^2(\theta)}\right)$$

Mais cette fois, la variance h_t^2 suit un processus ARMA et dépend donc des valeurs passées de la variance conditionnelle h_1^2, \dots, h_t^2 . Ces valeurs n'étant pas observées en pratique, la maximisation en direct de la vraisemblance est rendue impossible. En pratique, on estime successivement les valeurs de h_1^2, \dots, h_t^2 avant de calculer la vraisemblance. Ainsi, pour un vecteur $\theta_0 = (\alpha_0^0, \dots, \alpha_p^0, \beta_1^0, \dots, \beta_q^0)$ fixé de paramètres, on calcul récursivement

$$\hat{h}_s^2 = \alpha_0^0 + \sum_{i=1}^p \alpha_i^0 (Y_s - i)^2 + \sum_{j=1}^q \beta_j^0 (\hat{h}_{s-j}^2)^2$$

avec la convention $Y_i = 0$ et $(h_i)^2$ si $i \leq 0$.

On remplace donc la fonction de vraisemblance par

$$L(y_t/I_{t-1}) = \frac{1}{\hat{h}_t\sqrt{2\pi}} \exp\left(-\frac{(y_t - m_t(\theta))^2}{2\hat{h}_t^2(\theta)}\right)$$

et la fonction de vraisemblance totale est :

$$L(y_1, y_2, \dots, y_T; \theta_0) = \prod_{i=1}^T L(y_i/I_{i-1}, \theta_0)$$

Cette fonction de vraisemblance peut être calculée pour différentes valeurs du vecteur θ_0 et sa maximisation livre l'estimateur de maximum de vraisemblance.

Prévision

Modèle avec erreur ARCH

Supposons un processus AR(1) sans constante pour modéliser la moyenne conditionnelle

$$X_t = cX_{t-1} + \epsilon_t$$

avec $\epsilon_t/I_{t-1} \sim N(0, h_t)$. ϵ_t ARCH(1) tel que $h_t = \alpha_0 + \alpha_1(\epsilon_{t-1})^2$

Nous avons calculé précédemment les espérances et variances conditionnelles du processus ARCH. Nous connaissons aussi les formules pour les prévisions X_{T+h} et les erreurs de prévision $(\epsilon_{T+h})^2$ d'un modèle AR(1) supposant un bruit blanc fort des résidus. Plus spécifiquement,

$$E(X_{T+h}/I_T) = c^h X_T$$

on a

$$X_{t+h} = cX_{t+h-1} + \epsilon_{t+h} = c^h X_t + c^{h-1}\epsilon_{t+1} + \dots + c\epsilon_{t+h-1} + \epsilon_{t+h}$$

ainsi

$$V(\epsilon_{T+h}/I_T) = \sigma^2(1 + c^2 + \dots + (c^{h-1})^2)$$

Dans le cas d'un terme d'erreur ARCH(1), les erreurs au carré suivent un processus AR(1) ainsi ,nous avons

$$E[(\epsilon_{T+h})^2/I_T] = \alpha_0 + \alpha_1 E[(\epsilon_{T+h-1})^2/I_T] = \alpha_0 + \alpha_0\alpha_1 + \dots + (\alpha_1)^h(\epsilon_T)^2$$

Il s'agit alors de remplacer les termes appropriés dans l'équation

$$V(\epsilon_{T+h}/I_T) = E[(\epsilon_{T+h})^2/I_T] + c^2 E[(\epsilon_{T+h-1})^2/I_T] + \dots + c^{h-1} E[(\epsilon_{T+1})^2/I_T] + 0$$

Ainsi, les acteurs des marchés financiers peuvent établir leurs prévisions de la volatilité à partir des informations les plus récentes dont ils disposent.

Dans le cas du modèle GARCH(1 ; 1), nous avons

$$(h_{T+1})^2 = h_T(1) = \alpha_0 + \alpha_1(\epsilon_T)^2 + \beta_1(h_T)^2$$

Il est possible décrire la prévision d'une autre manière et en particulier la prévision de plusieurs périodes. En mettant $\epsilon_t = \eta_{tt}$. En élevant au carré nous obtenons

$$(\epsilon_t)^2 = (\eta_t)^2(t)^2.$$

Remplaçons cette expression dans la formulation du modèle GARCH(1,1) on obtient

$$(h_t)^2 = \alpha_0 + (\beta_1 + \alpha_1)(h_T)^2 + \alpha_1(h_T)^2((\eta_T)^2 - 1)$$

Ainsi, en répétant les substitutions, pour la prévision de h périodes, nous avons

$$\hat{h}_t(h) = \frac{\alpha_0[1 - (\beta_1 + \alpha_1)^h - 1]}{1 - (\beta_1 + \alpha_1)} + (\beta_1 + \alpha_1)^{h-1}((h_T)^2)(1)$$

et quand h tend vers l'infini la variance conditionnelle tend vers la valeur d'équilibre

$$\frac{\alpha_0}{1 - (\beta_1 + \alpha_1)}$$

3 Étude Pratique

Cette partie est une application sur l'indice boursier CAC 40 à l'aide du logiciel statistique [EViews 7](#).

3.1 L'indice boursier CAC 40

le CAC 40 ou Cotisation Assistée en Continu est un indice boursier qui traduit la performance des actions des 40 plus grandes entreprises françaises et qui représentent les $\frac{3}{4}$ de la valeur du marché parisien.

Exemples d'entreprises du cac 40 : Total (13% de l'indice) , Renault , Carrefour ,etc ...

3.2 Description de la base de données

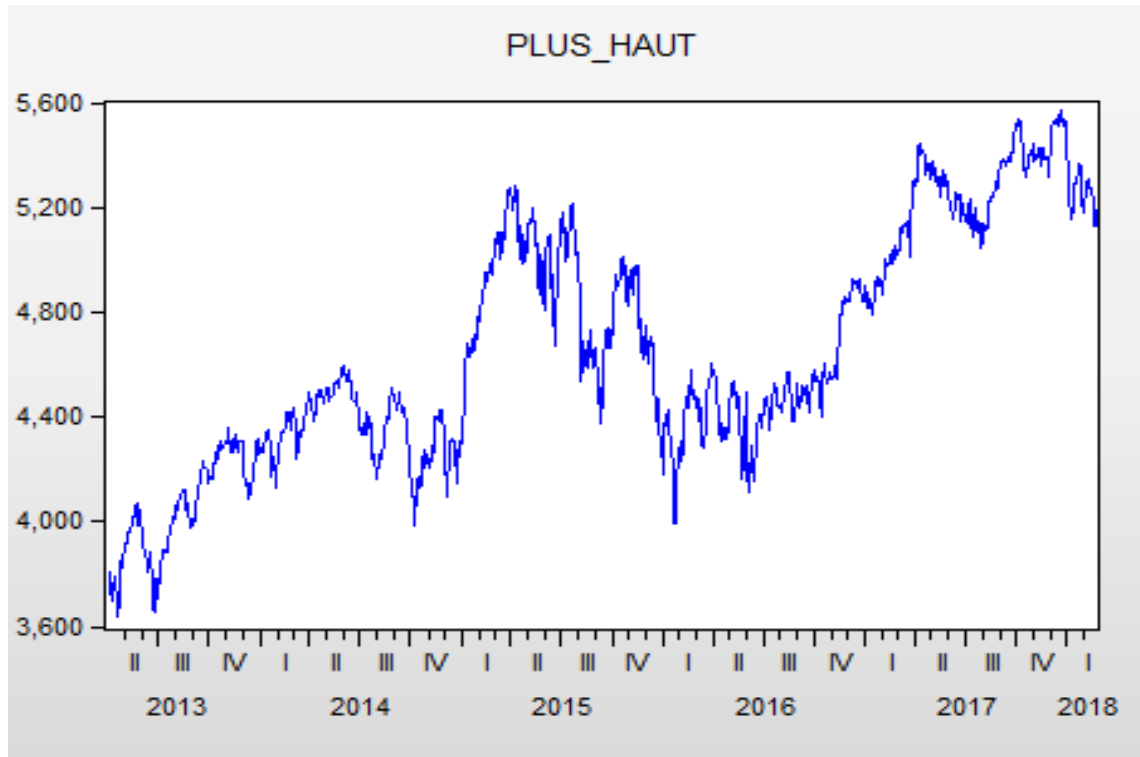
Il s'agit de 1279 observations journalières (5 jours par semaine) s'étalant sur [5 ans](#).
(du 02-04-2013 au 29-03-2018).

La base contient 7 variables tel que :

l'identifiant de la valeur, la date, le cours d'ouverture, le plus haut, le plus bas, le cours de clôture et le volume de titres échangés.

On s'intéresse particulièrement à la variable "[Plus Haut](#)" qui correspond au cours le plus élevée calculé chaque jour de 9h à 17h.30.

3.3 Etude de la stationnarité



Représentation graphique de la variable d'intérêt

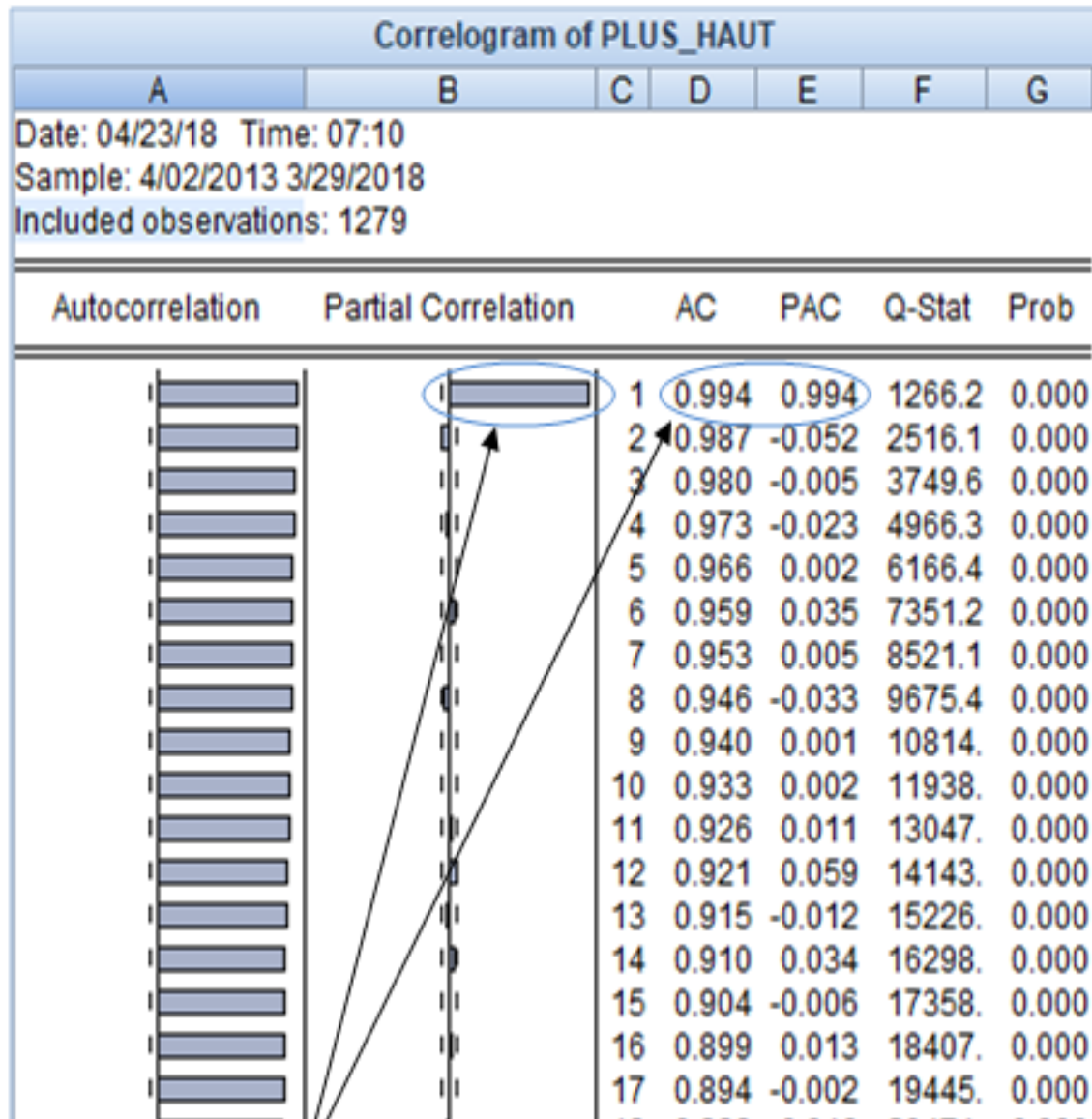
L'allure de ce graphe met en évidence la non-stationnarité puisqu'elle semble présenter :

- La présence d'une tendance.
- Une volatilité qui varie au cours du temps .

Nous allons étudier les corrélogrammes de la série pour plus de précision. A ce niveau, nous allons nous intéresser aux fonctions d'autocorrélations ACF et d'autocorrélations partielles PACF, dans le but d'en tirer des résultats pouvant nous guider à choisir le bon modèle.

Représentation du corrélogramme

La non-stationnarité du processus se confirme par la représentation du corrélogramme ci-dessous.



Présence d'une racine unitaire, processus non stationnaire

Test de racine unitaire (ADF)

Le test augmenté de Dickey-Fuller ou test ADF est un test statistique qui vise à savoir si une série temporelle est stationnaire.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PLUS_HAUT(-1)	-0.012428	0.004177	-2.975103	0.0030
D(PLUS_HAUT(-1))	0.090530	0.027932	3.241024	0.0012
C	51.69035	16.99184	3.042069	0.0024
@TREND(4/02/2013)	0.011043	0.005145	2.146121	0.0321

Les p valeurs de la constante et du Trend sont inférieure à 5%. La constante et le Trend sont significatifs.

Augmented Dickey-Fuller Unit Root Test on PLUS_HAUT		
Null Hypothesis: PLUS_HAUT has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 1 (Automatic - based on SIC, maxlag=22)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.975103	0.1397
Test critical values:		
1% level	-3.965248	
5% level	-3.413334	
10% level	-3.128697	

H_0 : non stationnarité du processus

La p valeur étant supérieure au seuil 5% ($0.1397 > 0.05$)

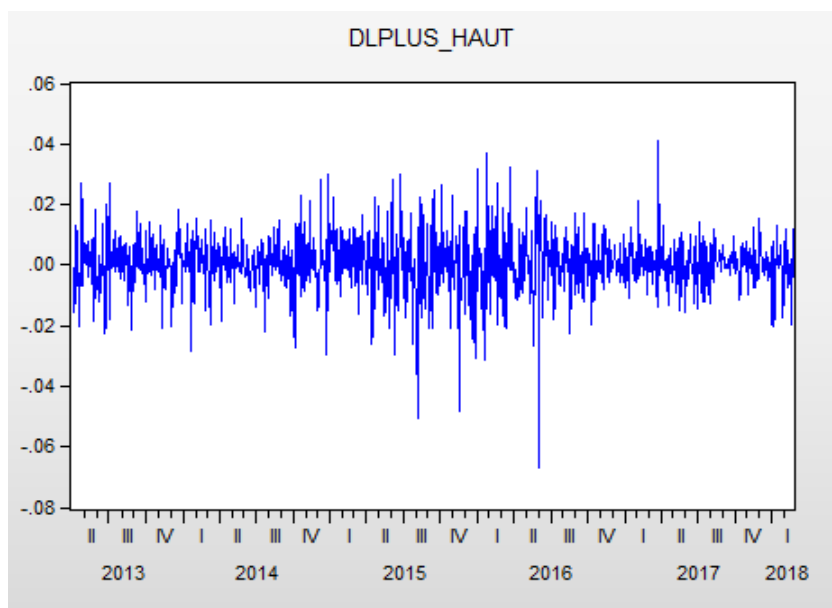
On accepte l'hypothèse de base : la série n'est pas stationnaire.

Conclusion : Le processus est complet et non stationnaire . Il faut le différencier pour le stationnariser.

Différentiation :

Il s'agit de différencier le log de la variable plus haut . L'application du log a pour but de lisser et réduire la tendance alors que la différenciation stationnarise le processus.

Courbe de la différenciation



L'allure de la courbe de différenciation marque une oscillation autour de la moyenne nulle ce qui montre sa stationnarité.

On mène un ADF test pour confirmer ce résultat .

Test de racine unitaire (ADF-test)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLPLUS_HAUT(-1)	-0.913947	0.027932	-32.71992	0.0000
C	0.000392	0.000517	0.757382	0.4490
@TREND(4/02/2013)	-2.63E-07	7.00E-07	-0.375218	0.7076

Les p valeurs de la constante et du Trend sont supérieur à 5%. La constante et le Trend ne sont pas significatifs.

Augmented Dickey-Fuller Unit Root Test on DLPLUS_HAUT		
Null Hypothesis: DLPLUS_HAUT has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic - based on SIC, maxlag=22)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-32.71992	0.0000
Test critical values:		
1% level	-3.965248	
5% level	-3.413334	
10% level	-3.128697	

H_0 : non stationnarité du processus

la p valeur étant nulle (< 0.05)

On rejète l'hypothèse de base : la série différenciée est stationnaire.

Conclusion : Le processus est stationnaire .Le trend et la constante étant non significatifs on refait le test sans trend.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLPLUS_HAUT(-1)	-0.913808	0.027921	-32.72885	0.0000
C	0.000223	0.000258	0.866284	0.3865

La p valeur de la constante est supérieure à 5%. La constante n'est pas significative.

Augmented Dickey-Fuller Unit Root Test on DLPLUS_HAUT		
Null Hypothesis: DLPLUS_HAUT has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on SIC, maxlag=22)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-32.72885	0.0000
Test critical values: 1% level	-3.435259	
5% level	-2.863595	
10% level	-2.567914	

H_0 : non stationnarité du processus.

la p valeur étant nulle (< 0.05)

On rejète l'hypothèse de base : la série différenciée est stationnaire.

Conclusion : Le processus est stationnaire .La constante étant non significative donc on refait le test sans contrainte ni trend .

Augmented Dickey-Fuller Unit Root Test on DLPLUS_HAUT		
Null Hypothesis: DLPLUS_HAUT has a unit root		
Exogenous: None		
Lag Length: 0 (Automatic - based on SIC, maxlag=22)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-32.72061	0.0000
Test critical values: 1% level	-2.566774	
5% level	-1.941071	
10% level	-1.616532	

H_0 : non stationnarité du processus

la p valeur étant nulle (< 0.05)



On rejète l'hypothèse de base : la série différenciée est stationnaire.

Conclusion : Le processus DLPlus-Haut est stationnaire sans tendance ni drift.

3.4 Modélisation

Modélisation ARMA(p,q) :

Pour identifier l'ordre p et q d'un processus ARMA, nous utilisons le corrélogramme de la fonction d'autocorrélation et la fonction d'autocorrélation partiel de la série .

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.086	0.086	9.4904	0.002
		2 -0.002	-0.010	9.4962	0.009
		3 0.010	0.012	9.6354	0.022
		4 -0.034	-0.037	11.161	0.025
		5 -0.081	-0.076	19.654	0.001
		6 -0.001	0.012	19.656	0.003
		7 0.054	0.054	23.433	0.001
		8 -0.010	-0.019	23.574	0.003
		9 -0.009	-0.012	23.688	0.005
		10 -0.018	-0.024	24.112	0.007
		11 -0.097	-0.090	36.192	0.000
		12 -0.040	-0.017	38.235	0.000
		13 -0.046	-0.046	40.977	0.000
		14 -0.003	0.001	40.987	0.000
		15 0.016	0.009	41.306	0.000
		16 0.042	0.026	43.550	0.000
		17 0.018	0.007	43.966	0.000
		18 0.039	0.040	45.906	0.000
		19 0.024	0.017	46.643	0.000

A priori, les éventuels processus de la variation du dlplus-haut sont AR(1), MA(1) et ARMA(1,1).

L'estimation de ce processus par la méthode de moindre carrée ordinaire est représentée dans le tableau suivant et justifie le choix d'un AR(1) avec une p value $(0.0019) < 5\%$.

Dependent Variable: DLPLUS_HAUT
Method: Least Squares
Date: 05/06/18 Time: 21:47
Sample (adjusted): 4/04/2013 3/29/2018
Included observations: 1277 after adjustments
Convergence achieved after 2 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.086802	0.027909	3.110171	0.0019
R-squared	0.006835	Mean dependent var		0.000244
Adjusted R-squared	0.006835	S.D. dependent var		0.009247
S.E. of regression	0.009215	Akaike info criterion		-6.535139
Sum squared resid	0.108358	Schwarz criterion		-6.531105
Log likelihood	4173.686	Hannan-Quinn criter.		-6.533624
Durbin-Watson stat	1.996883			
Inverted AR Roots	.09			


Tests sur les résidus :

a. Test d'absence d'auto corrélation des résidus :

Correlogram of Residuals					
Date: 05/09/18 Time: 20:57					
Sample: 4/04/2013 2/23/2018					
Included observations: 1277					
Q-statistic probabilities adjusted for 1 ARMA term(s)					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.000	0.000	5.E-05	
		2 -0.011	-0.011	0.1501	0.698
		3 0.014	0.014	0.3907	0.823
		4 -0.029	-0.029	1.4507	0.694
		5 -0.079	-0.079	9.5551	0.049
		6 0.001	0.000	9.5568	0.089
		7 0.056	0.055	13.584	0.035
		8 -0.015	-0.013	13.860	0.054
		9 -0.007	-0.011	13.926	0.084
		10 -0.009	-0.017	14.033	0.121
		11 -0.093	-0.091	25.276	0.005
		12 -0.028	-0.021	26.279	0.006
		13 -0.043	-0.048	28.662	0.004
		14 0.000	-0.003	28.662	0.007
		15 0.013	0.007	28.872	0.011
		16 0.039	0.026	30.884	0.009
		17 0.011	0.006	31.042	0.013

les Q-stat sont faibles et les p-values $> 5\%$ on conclue alors que les erreurs ne sont pas corrélées.

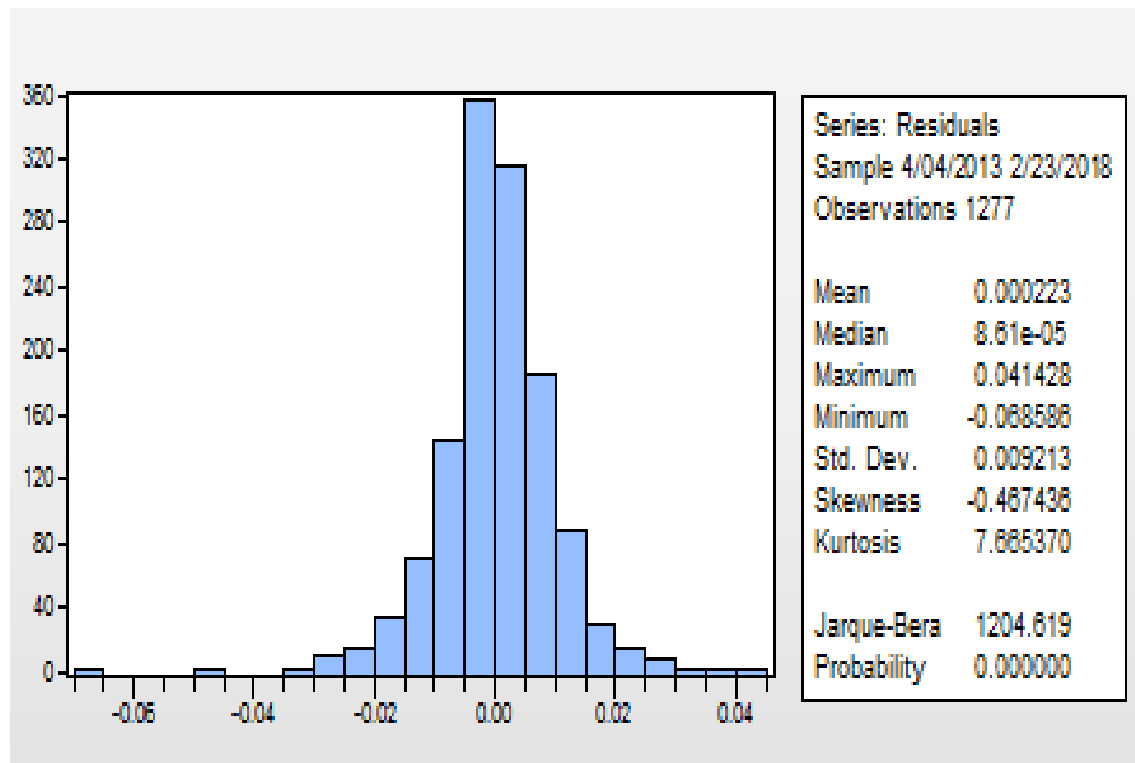
On représente également l'intervalle de confiance résiduel ci dessous :

obs	Actual	Fitted	Residual	Residual Plot
9/30/2013	0.01106	-0.00106	0.01212	
10/01/2013	-0.00377	0.00096	-0.00473	
10/02/2013	-0.00332	-0.00033	-0.00300	
10/03/2013	0.00224	-0.00029	0.00253	
10/04/2013	-0.00256	0.00019	-0.00275	
10/07/2013	0.00158	-0.00022	0.00181	
10/08/2013	-0.00238	0.00014	-0.00252	
10/09/2013	0.01393	-0.00021	0.01413	
10/10/2013	0.00164	0.00121	0.00043	
10/11/2013	-0.00108	0.00014	-0.00123	
10/14/2013	0.00914	-9.4E-05	0.00924	
10/15/2013	-0.00334	0.00079	-0.00413	
10/16/2013	-0.00121	-0.00029	-0.00092	
10/17/2013	0.01023	-0.00010	0.01033	
10/18/2013	2.8E-05	0.00089	-0.00086	
10/21/2013	0.00535	2.4E-06	0.00535	
10/22/2013	-0.00801	0.00046	-0.00848	
10/23/2013	0.00147	-0.00070	0.00216	
10/24/2013	-0.00029	0.00013	-0.00042	
10/25/2013	0.00212	-2.6E-05	0.00215	
10/28/2013	-0.00172	0.00018	-0.00191	
10/29/2013	0.00662	-0.00015	0.00677	
10/30/2013	-0.00142	0.00057	-0.00199	
10/31/2013	0.00033	-0.00012	0.00046	
11/01/2013	-0.00214	2.9E-05	-0.00217	

La majorité des valeurs résiduelles sont à l'intérieur de l'intervalle de confiance d'où L'absence de corrélation résiduelles dans les résidus estimés, ce qui confirme la présence de volatilité.

b. Tests de normalité sur les résidus du modèle optimal AR(1) :

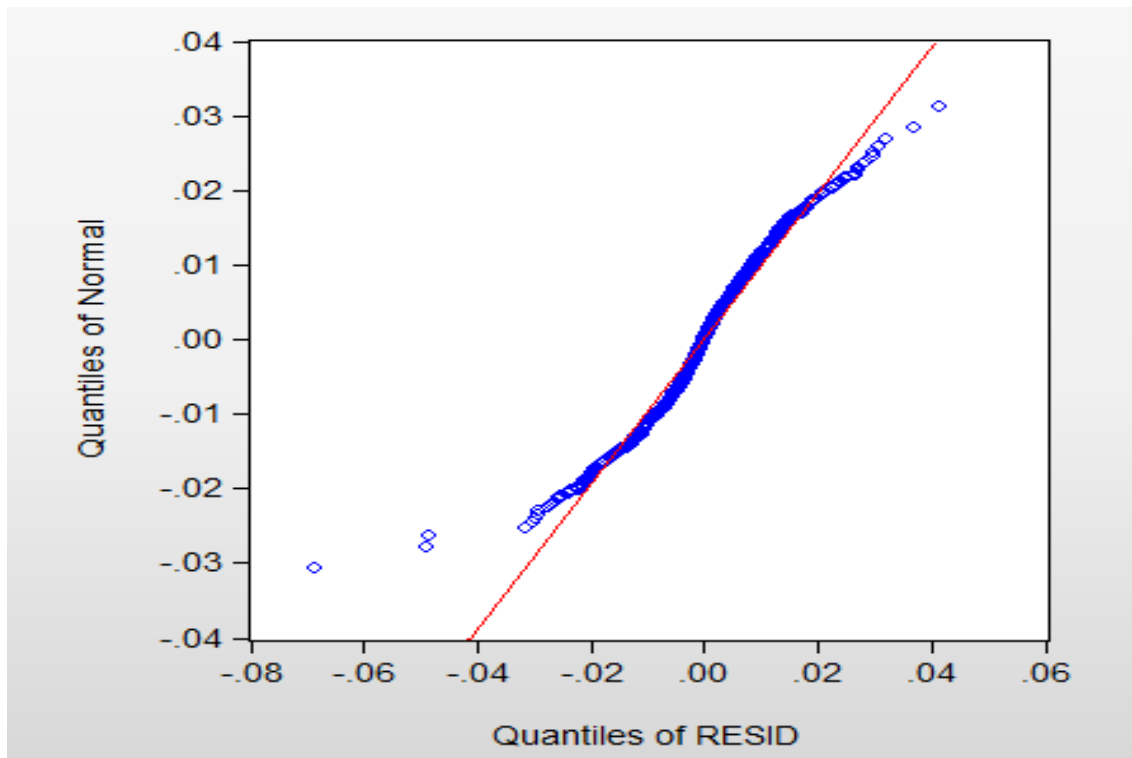
- Test de Skewness (asymétrie) , de Kurtosis (aplatissement) et Test de Jarque et Bera :



Remarquons que le coefficient d'aplatissement Kurtosis est égal à « 7.66537 » est un peu élevé, qui est supérieur à la valeur du kurtosis de la loi normal qui est 3. La valeur de ce coefficient trouvé indique que la courbe de la série est plus aplatie que la courbe de loi normale, cette valeur du coefficient témoigne la forte probabilité d'occurrence de point extrêmes.. De même, le coefficient d'asymétrie Skewness étant de l'ordre de « 0,467436 » n'est pas nul (la valeur théorique du coefficient de skewness pour une la loi normal) , ce coefficient montre la présence de l'asymétrie de la courbe de la série . Le coefficient d'asymétrie est négatif, cela nous permet de dire que la distribution est étalée vers la gauche.

En outre, le test de Jarque Bera donne un résultat de $1204.619 > \chi^2_{0.05}(2) = 5,99$ et une *pvalue* nulle ($< 5\%$) donc on rejette H_0 : les données suivent une loi normale. Ce qui nous conduit à confirmer le rejet de la normalité.

- Test QQ-Plot (méthode graphique)



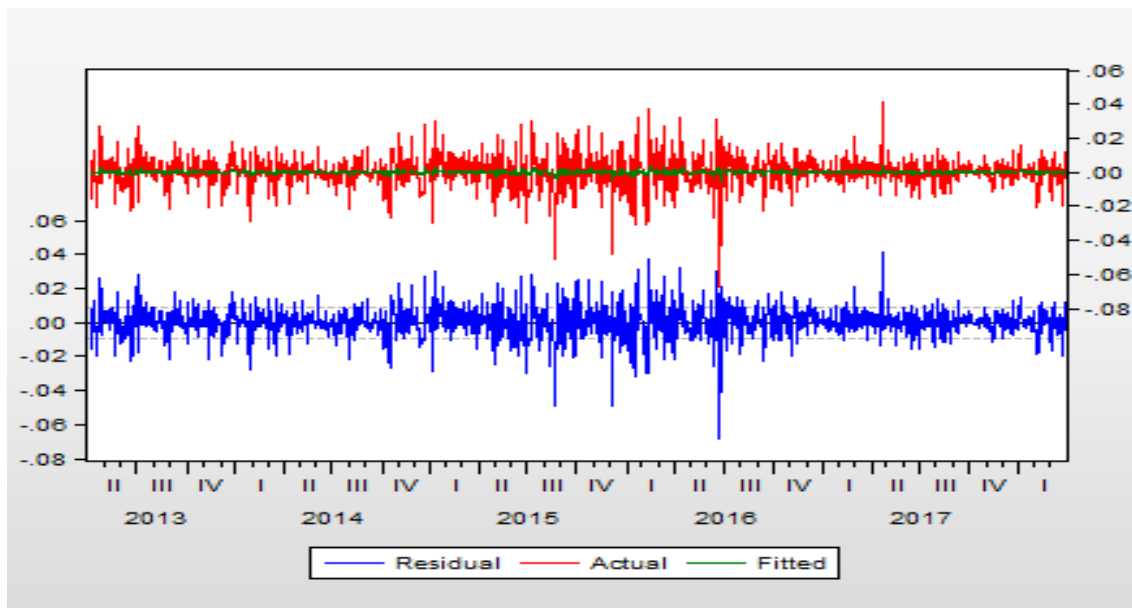
Le nuage de point (en bleu) est formé par (quantiles de $N(0,1)$, quantiles empiriques réduits des résidus), sous H_0 le nuage est rectiligne sur la droite rouge $y = x$)

On remarque que le nuage de point n'est pas rectiligne sur la droite, donc l'hypothèse nulle est rejetée c'est-à-dire les résidus ne suivent pas une loi normale.

On affirme donc que les résidus forment un bruit blanc non gaussien.

c. Test d'effet ARCH et test d'homoscédasticité :



































• Plot des résidus :



Une première observation du graphe des résidus ci-dessous montre que la moyenne de cette série est constante alors que sa variance change au cours du temps. De plus le processus étant non gaussien, on suspecte la présence d'un effet ARCH.

Test ARCH :

- Le corrélogramme des résidus au carré du modèle AR(1) :

Correlogram of Residuals Squared					
Date: 05/06/18 Time: 21:25					
Sample: 4/04/2013 2/23/2018					
Included observations: 1277					
Q-statistic probabilities adjusted for 1 ARMA term(s)					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.083	0.083	8.7557	
		2 0.076	0.070	16.191	0.000
		3 0.060	0.049	20.806	0.000
		4 0.113	0.101	37.226	0.000
		5 0.042	0.020	39.501	0.000
		6 0.012	-0.009	39.690	0.000
		7 0.023	0.009	40.373	0.000
		8 0.085	0.070	49.654	0.000
		9 0.093	0.077	60.828	0.000
		10 0.031	0.008	62.044	0.000
		11 0.031	0.009	63.287	0.000
		12 0.050	0.022	66.459	0.000
		13 0.062	0.035	71.427	0.000
		14 0.047	0.029	74.253	0.000
		15 0.013	-0.005	74.470	0.000
		16 0.030	0.007	75.622	0.000
		17 0.019	-0.009	76.108	0.000

==> On a pic à $p = 4$ et $q = 4$

Il faut préalablement déterminer le nombre de retards q à retenir. Au regard du corrélogramme des résidus au carré du modèle AR(1), on choisit, compte tenu du critère de parcimonie, un nombre de retards égal à 2 comme le montre le test d'hétéroscédasticité suivant .

Heteroskedasticity Test: ARCH

F-statistic	7.541408	Prob. F(2,1272)	0.0006
Obs*R-squared	14.94122	Prob. Chi-Square(2)	0.0006

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 05/06/18 Time: 23:50

Sample (adjusted): 4/08/2013 3/29/2018

Included observations: 1275 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.23E-05	6.88E-06	10.50827	0.0000
RESID^2(-1)	0.077133	0.027966	2.758089	0.0059
RESID^2(-2)	0.069813	0.027966	2.496389	0.0127
R-squared	0.011719	Mean dependent var	8.48E-05	
Adjusted R-squared	0.010165	S.D. dependent var	0.000218	
S.E. of regression	0.000217	Akaike info criterion	-14.02800	
Sum squared resid	6.01E-05	Schwarz criterion	-14.01588	
Log likelihood	8945.852	Hannan-Quinn criter.	-14.02345	
F-statistic	7.541408	Durbin-Watson stat	2.006723	

Les deux statistiques Fisher et Khi-deux donnent deux p-valeurs inférieures au seuil (5 %) : on rejette donc l'hypothèse nulle d'homoscédasticité en faveur de l'alternative d'hétéroscédasticité conditionnelle. Ce qui confirme la présence d'un effet ARCH.

Alors il faut ré-estimer le modèle en tenant compte de cet effet.

Estimation de modèle ARCH et GARCH :

Afin de tenir compte de l'effet d'ARCH, nous allons estimer par la méthode du maximum de vraisemblance l'équation de la variance conditionnelle conjointement à l'équation de la moyenne.

Estimation de modèle ARCH(2) :

L'estimation du modèle ARCH(2) se résume dans le tableau suivant :

Dependent Variable: DLPLUS_HAUT
 Method: ML - ARCH (Marquardt) - Normal distribution
 Date: 05/07/18 Time: 00:19
 Sample (adjusted): 4/04/2013 3/29/2018
 Included observations: 1277 after adjustments
 Convergence achieved after 10 iterations
 Presample variance: backcast (parameter = 0.7)
 GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-2)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.101779	0.029824	3.412673	0.0006
Variance Equation				
C	5.15E-05	2.39E-06	21.55004	0.0000
RESID(-1)^2	0.270997	0.029120	9.306326	0.0000
RESID(-2)^2	0.188692	0.032798	5.753198	0.0000
R-squared	0.006611	Mean dependent var		0.000244
Adjusted R-squared	0.006611	S.D. dependent var		0.009247
S.E. of regression	0.009216	Akaike info criterion		-6.602831
Sum squared resid	0.108383	Schwarz criterion		-6.586693
Log likelihood	4219.908	Hannan-Quinn criter.		-6.596771
Durbin-Watson stat	2.026464			

Le tableau de l'estimation d'ARCH(2) montre que les coefficients de paramètre de l'équation de la variance sont significativement différents de zéro et positifs donc les coefficients vérifient les contraintes assurant la positivité de la variance (que ce n'est pas les cas pour les autres modèles ARCH qu'on a déjà testé). De cela, nous retenons le modèle [ARCH\(2\)](#) comme modèle représentant la variance conditionnelle du processus .

Estimation du modèle GARCH (1,1)

Nous allons résumer l'estimation de ce modèle dans le tableau suivant :

Dependent Variable: DLPLUS_HAUT
 Method: ML - ARCH (Marquardt) - Normal distribution
 Date: 05/07/18 Time: 00:20
 Sample (adjusted): 4/04/2013 3/29/2018
 Included observations: 1277 after adjustments
 Convergence achieved after 17 iterations
 Presample variance: backcast (parameter = 0.7)
 GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.078150	0.027482	2.843718	0.0045
Variance Equation				
C	1.49E-06	3.59E-07	4.135680	0.0000
RESID(-1)^2	0.067717	0.007127	9.500820	0.0000
GARCH(-1)	0.916877	0.008609	106.5048	0.0000
R-squared	0.006760	Mean dependent var	0.000244	
Adjusted R-squared	0.006760	S.D. dependent var	0.009247	
S.E. of regression	0.009216	Akaike info criterion	-6.652149	
Sum squared resid	0.108366	Schwarz criterion	-6.636011	
Log likelihood	4251.397	Hannan-Quinn criter.	-6.646089	
Durbin-Watson stat	1.979757			

Le tableau de l'estimation de GARCH(1,1) illustre que les coefficients de paramètre de l'équation de la variance sont significativement différents de zéro et que les coefficients de l'équation de la variance vérifient les contraintes qui assure la positivité de la variance.

De cela, nous retenons ce modèle [GARCH\(1,1\)](#).

Comme nous avons retenu les modèles ARCH(2) et GARCH(1,1) , la question qui se pose est : Quel modèle doit-on finalement retenir pour la modélisation de la variance conditionnelle ?

Pour répondre à cette question, nous utilisons les critères de choix entre les différents modèles :

- Le coefficient de détermination corrigé R^2
- La valeur de la log-vraisemblance à l'optimum (LL)
- Le critère d'information d'Akaike (AIC)
- Le critère d'information de Schwarz (SIC).

Nous retenons donc le modèle qui a la minimum valeur selon les critères AIC et SIC.

Critères d'informations	ARCH(2)	GARCH(1,1)
R^2	0,00611	0,006760
logLikelihood	4219.908	4251.397
AIC	-6.602831	-6.652149
SIC	-6,586693	-6.636011
HQ	-6.596771	-6.646089

La comparaison des critères de sélection entre les différents modèles retenus nous mène à opter pour le modèle GARCH(1,1) comme meilleur modèle pour la modélisation de la variance conditionnelle du processus.

Notre processus est donc un AR(1)-GARCH(1,1)

Le modèle estimé obtenu s'écrit donc de la façon suivante :

$$\hat{x}_t = 0.078 \hat{x}_{t-1}$$

La variance estimée :

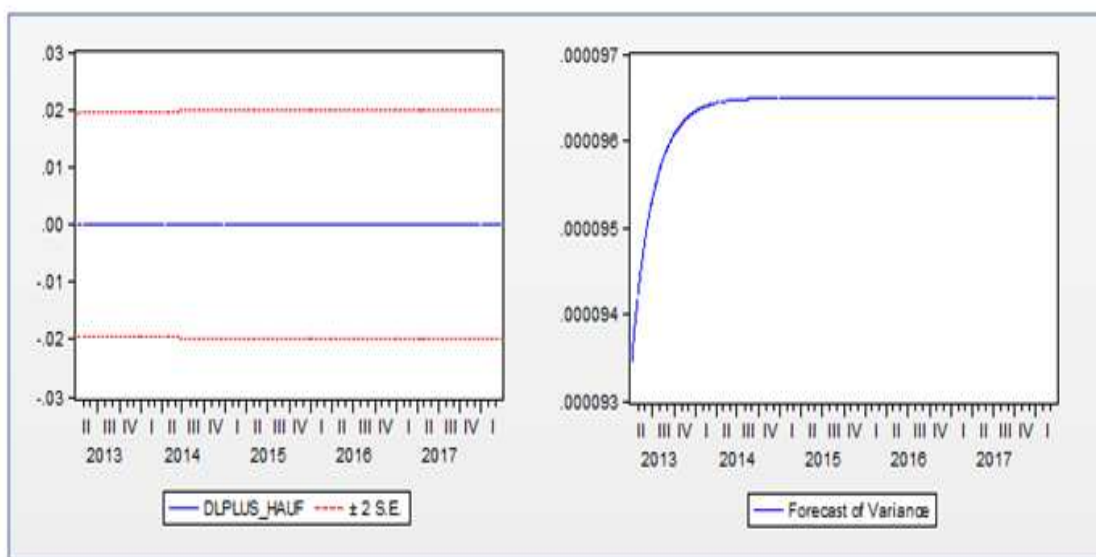
$$\hat{\sigma}_t^2 = 1.4 \cdot 10^{-6} + 0.917 \hat{\sigma}_{t-1}^2 + 0.068 \hat{\epsilon}_{t-1}^2$$

3.5 Prédiction

Après avoir redimensionner la base en augmentant sa taille d'une année, on mène la prévision dynamique et statique pour une année plus tard (2019).

3.5.1 Prédiction dynamique

Calcule les prévisions pour les périodes après la première période de l'échantillon en utilisant les valeurs précédemment prévues de la variable



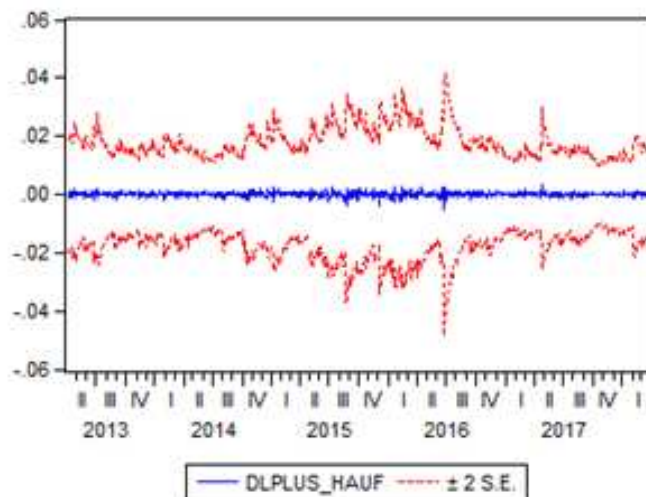
Intervalle de confiance et prévision de la variance

Forecast: DLPLUS_HAUF	
Actual: DLPLUS_HAUT	
Forecast sample: 4/02/2013 3/29/2018	
Adjusted sample: 4/04/2013 3/29/2018	
Included observations: 1277	
Root Mean Squared Error	0.009246
Mean Absolute Error	0.006503
Mean Abs. Percent Error	99.99792
Theil Inequality Coefficient	0.999755
Bias Proportion	0.000694
Variance Proportion	0.998829
Covariance Proportion	0.000476

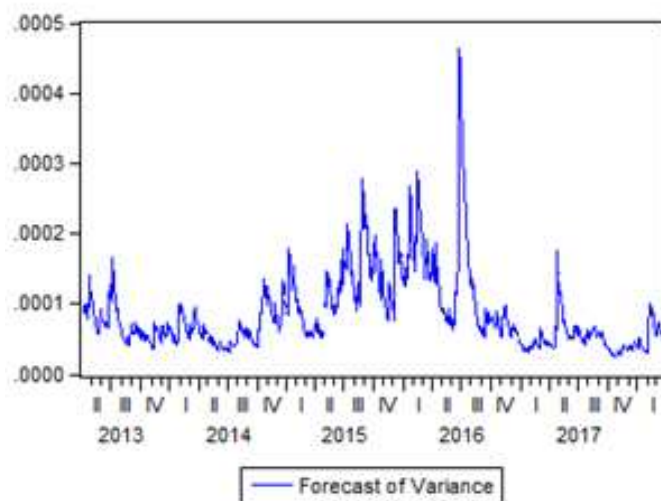
Le Root Mean Squared Error est quasiment nul et est de l'ordre de 0.009 ce qui indique une bonne prévision.

3.5.2 Prédiction statique

Utilise les valeurs réelles plutôt que les valeurs prévues (il ne peut être utilisé que lorsque les données réelles sont disponibles). On les appelle aussi des prévisions progressives ou à un pas.



Forecast: DLPLUS_HAUF	
Actual: DLPLUS_HAUT	
Forecast sample: 4/02/2013 3/29/2018	
Adjusted sample: 4/04/2013 3/29/2018	
Included observations: 1277	
Root Mean Squared Error	0.009212
Mean Absolute Error	0.006498
Mean Abs. Percent Error	119.1306
Theil Inequality Coefficient	0.924102
Bias Proportion	0.000598
Variance Proportion	0.855690
Covariance Proportion	0.143711



- Le Root Mean Squared Error est quasiment nul (0.009212) ce qui indique une bonne qualité de prévision.

- Les valeurs de la prévision sont à l'intérieur de l'intervalle de confiance.

Etant donné que le Root Mean Squared Error de la prévision statique est plus faible que celui de la prévision dynamique ($0.009212 < 0.009246$), on peut conclure que la prévision statique est plus précise que la dynamique.

3.5.3 Prédiction avec lissage exponentielle

Date: 05/09/18 Time: 16:31
Sample: 4/03/2013 2/23/2018
Included observations: 1278
Method: Single Exponential
Original Series: DLPLUS_HAUT
Forecast Series: DLPLUSM

Parameters:	Alpha	0.0010
Sum of Squared Residuals		0.109208
Root Mean Squared Error		0.009244
End of Period Levels:	Mean	0.000225

Le Root Mean Squared Error est presque nul (0,009244)

⇒ il s'agit d'une bonne prédiction.

Conclusion

La prédiction statique est meilleure que la prédiction dynamique et celle avec lissage exponentielle car elle a le Root Mean Squared Error le plus faible.

Visualisation de la continuité du processus

Les figures ci dessous (Figures 1 , 2 et 3) représentent la superposition des courbes de la variable d'intérêt et leurs prévisions.

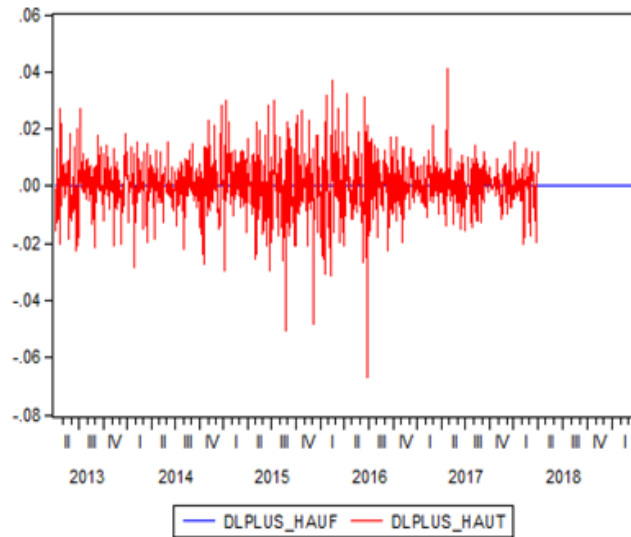


Figure 1

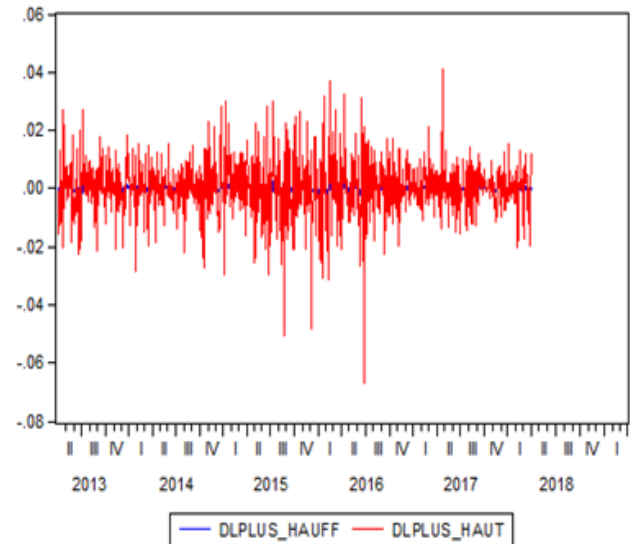


Figure 2

La courbe rouge représente le processus à prévoir alors que les courbes bleu représentent respectivement la prévision dynamique (Figure1) et la prévision statique (Figure 2).

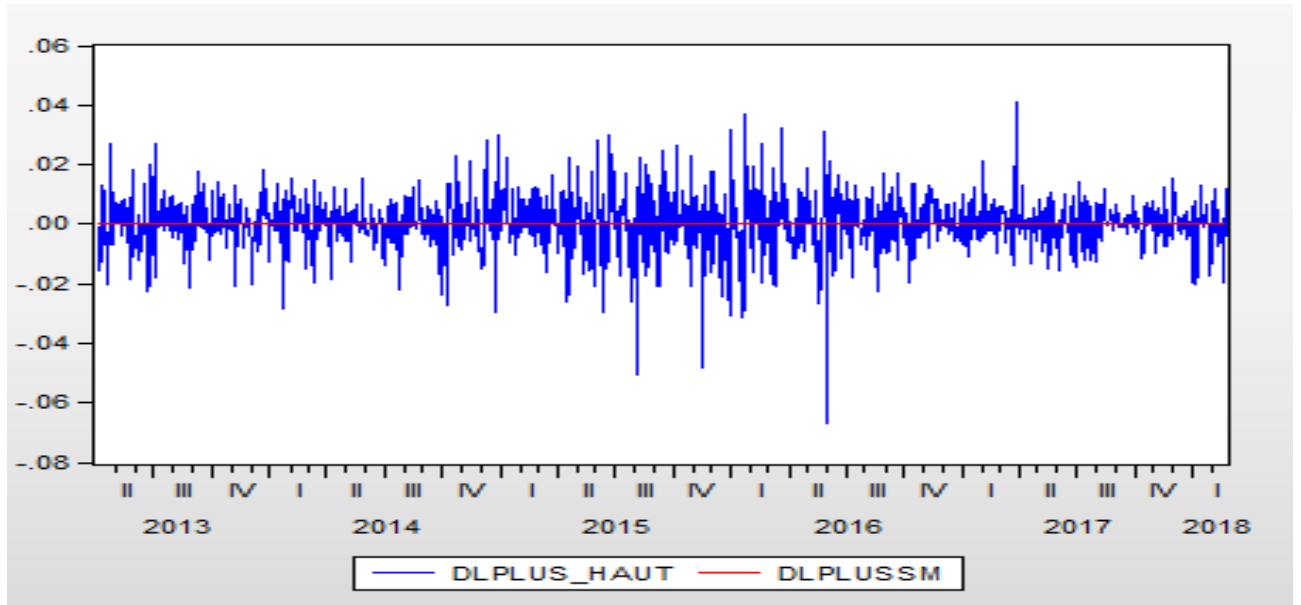


Figure 3

Pour la figure 3 , la variable DLplus-haut est en bleu alors que la prévision avec lissage exponentielle est en rouge.

Conclusion

On remarque que l'allure de chaque prévision garde le même aspect que le processus à prévoir.

On peut aussi affirmer suite à une comparaison entre les trois figures que la prévision statique est la plus favorable . Résultat déjà conclue d'après les valeurs du Root Mean Squared Error.

4 Conclusion

Les modèles ARCH et GARCH évoqués aux termes de cette étude nous ont permis d'exploiter le caractère hétéroscédastique des séries de rendements pour pouvoir calculer la variance conditionnelle de l'indice CAC 40. Cependant le modèle GARCH présente quelques défauts tel est le cas des [résidus qui ne sont pas toujours de loi normale](#) . Les modèles ARCH posent aussi problème lorsque le [nombre de données](#) historiques devient [extrêmement grand](#) auquel cas les variances conditionnelles ont tendance a devenir négatives. En effet, le problème des modèles ARCH vient du fait que la volatilité est prédite par les carrés des innovations. Or, les rentabilités des actifs et la volatilité de ces actifs tendent a être négativement corrélées, phénomène que les modèles ARCH ne peuvent incorporer car ils restreignent la volatilité a être seulement affectée par les changements d'amplitude des innovations.Ce qui fait appel à d'autres modèles comme le modèle [EGARCH](#), qui tente de remédier a cet inconvénient. On cite également d'autres modèles qui étendent le ARCH tel que [ARCH-M](#) et [N-GARCH](#).