

Modélisation de la volatilité des rendements Bitcoin par les modèles GARCH non paramétriques

Sami Mestiri

Maitre assistant en Méthodes quantitatives

Faculté des Sciences économiques et gestion de Mahdia, Tunisie

mestirisami2007@gmail.com

Résumé

Objectif: L'objectif de cet article est de mettre en évidence l'efficacité du modèle GARCH non paramétrique pour la prédiction des prix futurs du Bitcoin.

Méthodologie: Les modèles GARCH paramétriques pour caractériser la volatilité des rendements Bitcoin sont très utilisés dans la littérature empirique. Alternativement, nous considérons une approche non paramétrique pour modéliser et prévoir la volatilité des rendements Bitcoin.

Résultats: Nous montrons que la prévision de volatilité du modèle GARCH non paramétrique donne des performances supérieures par rapport à une classe étendue de modèles GARCH paramétriques.

Originalité/pertinence : L'amélioration de la précision des prévisions de la volatilité des rendements Bitcoin basée sur le modèle GARCH non paramétrique suggère que cette méthode offre une alternative attrayante et viable par rapport aux modèles paramétriques GARCH couramment utilisés.

Mots clés : Bitcoin ; volatilité; GARCH ; Non-paramétrique, Prévision.

JEL codes: C01

Modeling the volatility of Bitcoin returns using Nonparametric GARCH models

Summary

Objective: The purpose of this paper is to demonstrate the effectiveness of the nonparametric GARCH model for the prediction of future Bitcoin prices.

Methodology: The parametric GARCH models to characterize the volatility of Bitcoin returns are widely used in the empirical literature. Alternatively, we consider a non-parametric approach to model and forecast the volatility of Bitcoin returns.

Results: We show that the volatility forecast of the nonparametric GARCH model yields superior performance compared to an extended class of parametric GARCH models.

Originality / relevance: The improved accuracy of forecasting the volatility of Bitcoin returns based on the nonparametric GARCH model suggests that this method offers an attractive and viable alternative to commonly used GARCH parametric models.

Keywords: Bitcoin; volatility; GARCH; Non-parametric, Forecast.

JEL codes: C01

Introduction

Le marché des crypto-monnaies est une source potentielle d'instabilité financière et son impact sur le marché financier reste encore incertain. Contrairement aux autres actifs financiers qui sont régularisés, il n'y a pas de réglementation formelle pour les crypto-monnaies. Les crypto-monnaies diffèrent également considérablement des autres actifs financiers sur le marché financier et créent ainsi de grandes perspectives pour les investisseurs et les acteurs du marché en termes d'analyse de portefeuille, de gestion des risques et même d'analyse du sentiment des consommateurs. Sur le marché des crypto-monnaies, la modélisation de la volatilité est importante pour mesurer le risque d'un investissement.

La volatilité peut être définie comme une mesure de la dispersion du cours d'un actif financier. Les acteurs du marché et les investisseurs sont donc intéressés par une estimation précise de la volatilité sur le marché des crypto-monnaies. Ceci est le résultat de la corrélation entre la volatilité et les rendements sur investissement. Il est à noter que la volatilité n'est pas directement observable et, par conséquent, il existe un besoin croissant d'un modèle efficace capable de capturer la volatilité des prix sur le marché des crypto-monnaies. Comme Bitcoin a progressivement eu une place sur les marchés financiers et dans la gestion de portefeuille, l'analyse des séries chronologiques est un outil utile pour étudier les caractéristiques des prix et des rendements Bitcoin, et extraire des statistiques significatives afin de prédire les valeurs futures de la série.

L'objectif de cet article est d'appliquer les techniques de séries chronologiques pour modéliser la volatilité des rendements et pour prédire les prix futurs du Bitcoin. En outre, le but est d'examiner l'efficacité du modèle populaire GARCH dans le monde économique et financier. Dans cet article, nous supposons que la forme fonctionnelle du processus de variance non spécifiée et nous tenterons de l'estimer en tant que moyenne additive non paramétrique. Empiriquement, nous vérifions que le modèle non paramétrique peut capturer l'effet de levier des nouvelles négatives et surpasser les modèles paramétriques de la famille GARCH les plus couramment considérés.

Le reste du papier est organisé comme suit. La section 2 est consacrée pour la revue de la littérature. La section 3 décrit les modèles GARCH paramétriques et non paramétriques. La section 4 présente la méthodologie de prévision. La section 5 présente un résumé statistique des données et des résultats empiriques. La section 6 conclut.

1. Revue de la littérature

Quelques études ont déjà été menées sur les caractéristiques financières et statistiques du Bitcoin. Un groupe d'économistes s'est concentré sur la découverte des prix sur le marché Bitcoin, par exemple, Brandvold et al. (2015) et Bouoiyour et al. (2016) révèlent une relation d'avance-retard entre les prix du Bitcoin, l'utilisation des transactions et l'attractivité des investisseurs. D'autres études montrent également que le prix du Bitcoin est soumis à des facteurs spécifiques qui sont considérablement différents de ceux affectant les actifs financiers conventionnels, tels que la recherche sur Internet, les informations sur les tendances de Google et les informations de bouche à oreille sur les réseaux sociaux. En fait, comme Bitcoin est principalement utilisé et considéré comme un actif plutôt que comme une devise, et le marché du Bitcoin est actuellement hautement spéculatif, et plus volatil et sensible aux bulles spéculatives que les autres devises. De plus, la présence de mémoire longue et de volatilité persistante justifie l'application de modèles de type GARCH.

Dans leur article, Tjøstheim et Auestad (1994) ont étudié la possibilité d'identifier des modèles de séries chronologiques non linéaires à l'aide de méthodes non paramétriques. Härdle et Chen (1995) présentent un examen sélectif des approches basées sur la procédure de construction de modèles non paramétriques dans l'analyse des séries chronologiques. Ils soulignent que l'analyse des séries chronologiques non linéaires et non paramétriques est utile pour faire face aux limites des modèles ARMA à moyenne constante.

Härdle et al. (1997) passent en revue certains développements des techniques non paramétriques modernes pour l'analyse des séries chronologiques. Engle et Gonzalez-Rivera (1991) abordent le modèle ARCH semi-paramétrique en introduisant un estimateur plus efficace basé sur une densité estimée non paramétrique. Ils évaluent également la perte d'efficacité de l'estimateur de quasi-maximum de vraisemblance, qui suppose à tort la normalité. Buhlmann et McNeil (2002) ont proposé une approche non paramétrique de la modélisation GARCH. Hou et Suardi (2012) ont examiné l'approche non paramétrique de Buhlmann et McNeil (2002) pour modéliser et prévoir la volatilité des rendements des prix du pétrole brut. Ils utilisent 4845 observations quotidiennes sur les prix au comptant du pétrole brut de West Texas Intermediate, du 6 janvier 1992 au 30 juillet 2010, dans leur application. Selon leurs résultats sur la précision des prévisions, le modèle GARCH non paramétrique a des performances supérieures aux modèles GARCH paramétriques. Ils préfèrent les approches non paramétriques en raison de la non normalité de la distribution des prix du pétrole.

Une autre raison importante dans le développement de modèles non paramétriques est la procédure de sélection du décalage. Les modèles non paramétriques habituels ont des performances moins que satisfaisantes lorsqu'ils traitent plus d'un décalage, en particulier dans le cas de la malédiction de la dimensionnalité. Des critères alternatifs de sélection de décalage ont été étudiés pour les processus autorégressifs non linéaires. Tjøstheim et Auestad (1994) proposent d'utiliser une version non paramétrique de l'erreur de prédiction finale (FPE). Tschernig et Yang (2000) ont dérivé une version non paramétrique de l'erreur de prédiction finale pour la sélection de décalage dans des séries chronologiques autorégressives non linéaires dans des conditions très générales, y compris l'hétéroscédasticité. Yang, et al. (1999) introduisent une nouvelle auto-régression non paramétrique avec volatilité multiplicative et moyenne additive pour obtenir de meilleures estimations. Wang et al. (2012) ont proposé une nouvelle modélisation GARCH semi-paramétrique efficace de la volatilité en tenant compte de la procédure de sélection des décalages.

2. Modèles économétriques

2.1. Le modèle GARCH

Le modèle GARCH de Bollerslev (1986) est le modèle le plus largement utilisé pour l'estimation de la volatilité. Les modèles GARCH ont eu beaucoup de succès dans la littérature en raison de leur spécification simple et de leur facilité d'interprétation. Selon Bera et Higgins (1993), la plupart des travaux financiers appliqués montrent que GARCH (1,1) fournit une approximation flexible et parcimonie de la dynamique de la variance conditionnelle et est capable de représenter la majorité des séries financières. Le modèle GARCH (1,1) s'écrit :

$$Y_t = \mu + \varepsilon_t, \text{ avec } \varepsilon_t = \sigma_t \cdot z_t; z_t \sim N(0,1)$$

L'équation de la variance conditionnelle des résidus est définie comme:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \cdot \varepsilon_{t-1}^2 + \beta_1 \cdot \sigma_{t-1}^2 \quad (1)$$

ou $\alpha_0 > 0, \alpha_1 \geq 0$ et $\beta_1 \geq 0$ sont des constantes et le modèle ARCH(1) correspond à $\beta_1 = 0$. La contrainte $\alpha_1 + \beta_1 < 1$ implique que la variance inconditionnelle de la série de rendements ε est finie et que la variance conditionnelle σ_t^2 évolue dans le temps. Il fournit également la condition nécessaire et suffisante pour le processus stochastique $\sigma_t; t \in Z$ pour être un unique processus strictement stationnaire avec $E(\sigma_t^2) < \infty$.

Deux propriétés clés peuvent être notées à partir de (1). Premièrement, une valeur élevée de ε_{t-1}^2 ou σ_{t-1}^2 donne lieu à une valeur élevée de σ_t^2 et cela génère le groupement de volatilité qui est communément connu dans les séries chronologiques. Deuxièmement, la distribution de la queue est plus épaisse que celle d'une distribution normale.

2.2. Le modèle EGARCH

Malgré leur popularité, les modèles ARCH et GARCH souffrent de plusieurs faiblesses et inconvénients. Nelson (1991) a critiqué les modèles GARCH sous trois aspects : premièrement, les paramètres sont limités à être positifs à chaque instant. Deuxièmement, il ne tient pas compte de l'effet d'asymétrie (ou de l'effet de levier). Troisièmement, il est difficile de mesurer la persistance des chocs sur la volatilité. Nelson (1991) a proposé le GARCH exponentiel (EGARCH) qui tient compte des inconvénients d'un modèle GARCH standard. Le processus EGARCH de premier ordre (ou EGARCH(1,1)) spécifie le modèle comme

$$Y_t = \mu + \varepsilon_t, \text{ avec } \varepsilon_t = \sigma_t \cdot z_t; z_t \sim N(0,1)$$

L'équation de la variance conditionnelle des résidus est définie comme

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1 \cdot g(\varepsilon_{t-1}) + \beta_1 \cdot \log(\sigma_{t-1}^2) \quad (2)$$

Où ε_t suit la loi normale est un bruit blanc faible et la fonction $g(\cdot)$ vérifiée.

$$g(\varepsilon_{t-1}) = \alpha \cdot \varepsilon_{t-1} + \gamma (|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)) \quad (3)$$

Le coefficient γ signifie l'effet de levier des chocs sur la volatilité. Le principal avantage du modèle EGARCH est qu'il n'est pas nécessaire d'imposer les restrictions positives aux coefficients de variance. Les coefficients γ doivent être négatifs pour mettre en évidence des effets asymétriques.

2.3. Le modèle GJR-GARCH

Dans l'approche GARCH (1,1), les bonnes et les mauvaises nouvelles, c'est-à-dire les chocs positifs et négatifs, ont le même impact sur la variance conditionnelle. De nombreuses études ont mis en évidence une asymétrie dans le comportement des cours des actions, c'est-à-dire que les surprises négatives semblent augmenter la volatilité davantage que les surprises positives. Pour permettre des effets asymétriques dans la volatilité, Glosten et al. (1993) ajoutent un terme supplémentaire à la variance conditionnelle et formulent ce que l'on appelle le modèle GJR. Le GJR (1,1) est spécifié comme suit,

$$Y_t = \mu + \varepsilon_t, \text{ avec } \varepsilon_t = \sigma_t \cdot z_t; z_t \sim N(0,1)$$

L'équation de la variance conditionnelle des résidus est définie comme:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \cdot \varepsilon_{t-1}^2 + \gamma \cdot (I_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2) + \beta_1 \cdot \sigma_{t-1}^2 \quad (4)$$

Où z_t désigne un faible bruit blanc de moyenne nulle et de variance constante dans le temps, et les coefficients α_1 , β_1 et γ sont des paramètres réels et $I_{\varepsilon_{t-1} < 0}$ désigne la fonction indicatrice telle que

$$I_{\varepsilon_{t-1} < 0} = \begin{cases} 1 & \text{si } \varepsilon_{t-1} < 0 \\ 0 & \text{sinon.} \end{cases}$$

La structure de ce modèle indique qu'un ε_{t-1} positif contribue $\alpha_1 \cdot \varepsilon_{t-1}^2$ à σ_t , alors qu'un ε_{t-1} négatif a un impact plus important de $(\alpha_1 + \gamma) \cdot \varepsilon_{t-1}^2$ avec $\gamma > 0$. Par conséquent, si les paramètres γ sont significativement positifs, alors les innovations négatives génèrent plus de volatilité que les innovations positives de même ampleur. La principale caractéristique de ce modèle est qu'un choc négatif a un impact plus important qu'un choc positif et, par conséquent, il capture l'effet de levier. Comme le modèle GARCH, le modèle GJR-GARCH capture le clustering de volatilité. En outre, il peut être montré que la distribution inconditionnelle présente un excès de kurtosis même sous la distribution gaussienne.

2.4. Le modèle ARCH non paramétrique

Le point de départ du processus de génération de données d'un processus stochastique à temps discret strictement stationnaire R_t défini sur un espace de probabilité est le modèle général de régression stochastique non linéaire univarié donné par

$$R_t = m(R_{t-1}, \dots, R_{t-p}) + \sigma(R_{t-1}, \dots, R_{t-p})\varepsilon_t, \quad t = 1, \dots, T(5)$$

Avec $m(R_{t-1}, \dots, R_{t-p}) = E(R_t / R_{t-1} = r_1, \dots, R_{t-p} = r_p)$ est une fonction moyenne conditionnelle autorégressive non linéaire, $\sigma^2(R_{t-1}, \dots, R_{t-p}) = \text{Var}(R_t / R_{t-1} = r_1, \dots, R_{t-p} = r_p)$ représente la fonction de variance conditionnelle autorégressive non linéaire et ε_t est une séquence indépendante et identiquement distribuée (i,i,d) de variables aléatoires avec $E(\varepsilon_t / R_{t-1}, \dots, R_{t-p}) = 0$, $\text{Var}(\varepsilon_t / R_{t-1}, \dots, R_{t-p}) = 1$ et indépendant de R_{t-1}, \dots, R_{t-p} .

Le modèle (5) est connu sous le nom de modèle non-linéaire-CHARN autorégressif conditionnel hétéroscédastique ; voir Bossaerts et al. (1996) ou le modèle NARCH non paramétrique autorégressif conditionnel voir Fan et Yao (2005). Dans ce modèle, on n'impose aucune forme particulière (paramétrique) aux fonctions de moyenne conditionnelle et de volatilité. Cependant, en raison du problème de la dimensionnalité, on suppose un certain niveau de structure sur la fonction conditionnelle $m(\cdot)$ et $\sigma(\cdot)$.

Le modèle NARCH (1,1) conditionnel hétéroscédastique non linéaire autorégressif conditionnel de premier ordre s'écrit sous la forme suivante :

$$R_t = m(R_{t-1}) + \sigma(R_{t-1})\varepsilon_t, \quad t = 2, \dots, T$$

où R_t sont observés et dépendent de R_{t-1} avec un décalage de 1, $m(R_{t-1})$ est la fonction de tendance du modèle NARCH, $\sigma(R_{t-1})$ est la fonction hétéroscédastique du modèle NARCH, et ε_t désigne une variable aléatoire dans le terme d'erreur, de moyenne zéro et de variance un. D'après Fan et Yao (1998) si R_t est un processus stationnaire, la fonction de variance conditionnelle peut être décomposée en

$$\sigma^2 = E(R_t^2 / R_{t-1} = r) - (E(R_t / R_{t-1} = r))^2 = g(r) - m(r)^2$$

tel que l'estimation de la variance conditionnelle est basée sur l'estimation non paramétrique de $g(r)$ et $m(r)$ donnée par $\hat{\sigma}^2(r) = \hat{g}(r) - \hat{m}(r)^2$. Une façon d'obtenir des estimations des fonctions $g(r)$ et $m(r)$ consiste à appliquer l'estimateur populaire de Nadaraya-Watson donné par :

$$\hat{m}(R_{t-1}) = \frac{\sum_{t=2}^T K(R_{t-1}-r)/h R_t}{\sum_{t=2}^T K(R_{t-1}-r)/h}$$

$$\hat{g}(R_{t-1}) = \frac{\sum_{t=2}^T K(R_{t-1}-r)/h R_t^2}{\sum_{t=2}^T K(R_{t-1}-r)/h}$$

La fonction $K(\cdot)$ est généralement une densité de probabilité symétrique et des exemples de fonctions de noyau couramment utilisées sont le noyau gaussien $K(t) = (\sqrt{2\pi})^{-1} \exp(-t^2/2)$ et le noyau *Epanechnikov* $K(t) = \max\{\frac{3}{4}(1-t^2), 0\}$. $h > 0$ est le paramètre de bande passante (paramètre de lissage).

2.5. Le modèle GARCH non paramétrique

Le modèle GARCH non paramétrique ne nécessite pas la spécification de la forme fonctionnelle de la volatilité et qui ne tient pas compte de la forme de la distribution de l'innovation. De plus, ce modèle permet à la matrice de covariance conditionnelle des variables dépendantes de suivre une structure dynamique flexible. Le processus stochastique stationnaire $\{\varepsilon_t; 1 < t < n\}$ a la forme non paramétrique GARCH(1,1) donnée dans (Bühlmann and McNeill, 2002) :

$$R_t = \mu + \varepsilon_t, \text{ avec } \varepsilon_t = \sigma_t \cdot z_t; z_t \sim N(0,1)$$

L'équation de la variance conditionnelle des résidus est définie comme:

$$\sigma_t^2 = f(\varepsilon_{t-1}, \sigma_{t-1}^2) \quad (6)$$

Dans l'approche GARCH non paramétrique, la forme exacte de f n'est pas spécifiée et est estimée à l'aide d'une technique de lissage non paramétrique bivariable qui est moins sensible aux erreurs de spécification du modèle telles que la volatilité asymétrique négligée.

En supposant que $\{\varepsilon_t; 1 < t < n\}$ provenant d'un processus satisfaisant (6), l'estimation d'un modèle GARCH non paramétrique est appliquée avec les étapes suivantes telles que proposées dans (Bühlmann et McNeill, 2002) :

1- Premièrement, au pas $m=0$, une estimation de la volatilité $\{\hat{\sigma}_{t,0}^2; 1 < t < n\}$ est obtenue en ajustant un GARCH(1,1) paramétrique ordinaire. Ensuite, les prédictions du modèle GARCH(1,1) sont extraites, ce qui donne les estimations $\{\hat{\sigma}_{t,0}^2\}$ pour le pas $m=0$ de l'algorithme. Comme la première valeur n'est pas estimée dans les rendements, elle est définie comme égale à la moyenne.

2- Dans l'étape $m=1$, ε_t^2 est régressé avec une technique de lissage non paramétrique contre ε_{t-1} et $\hat{\sigma}_{t-1,0}^2$ qui sont obtenus à partir du paramètre GARCH(1,1). Les valeurs au carré des résidus sont obtenues à partir du modèle ARIMA et les valeurs décalées sont le premier décalage des résidus du modèle ARIMA. La variance estimée de la série de rendements est obtenue à partir de l'étape précédente de l'algorithme.

3- A la m 'ième étape, l'algorithme est répété et le $\hat{\sigma}_{t-1,m}^2$ est estimé par ε_{t-1} et $\hat{\sigma}_{t-1,m-1}^2$.

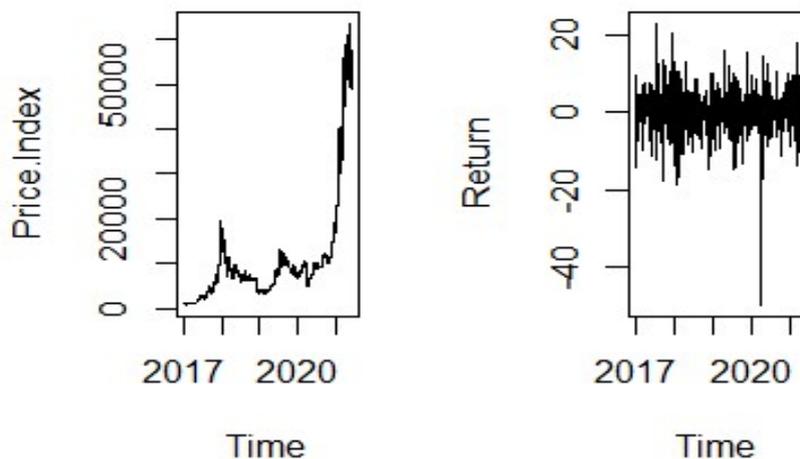
Bien que le modèle de volatilité soit estimé de manière non paramétrique, sa spécification ressemble aux modèles paramétriques GARCH largement utilisés de la volatilité des rendements Bitcoin.

3. Résultats Empiriques

3.1. Description des données

Dans cette étude, nous appliquons les différents modèles GARCH paramétriques et non paramétriques décrits précédemment pour estimer la fonction de variance pour les rendements Bitcoin. Les données utilisées sont le cours de clôture du Bitcoin car il reflétait toutes les activités de Bitcoin pour chaque jour de bourse. Le cours de clôture journalier historique du Bitcoin a été extrait du 01/02/2017 au 30/04/2021 sur <https://finance.yahoo.com/> et comprenait 1580 jours de bourse.

Figure 1: Cours de clôture et séries de rendement de Bitcoin.



Source : L'auteur à partir de l'analyse des données sur R

Afin d'évaluer et de comparer les performances prédictives du modèle GARCH non paramétrique avec divers modèles paramétriques, les données sont ensuite divisées en un groupe dans l'échantillon (du 2 janvier 2017 au 30 avril 2021) et un groupe hors échantillon (du 1er mai 2021 au 10 juin 2021). L'ensemble de l'échantillon compte 1540 observations et les 40 dernières sont utilisées pour des prévisions hors échantillon.

Maintenant, supposons que P_t et P_{t-1} représentent respectivement le prix du jour en cours et le prix de jour précédent de Bitcoin. Soit (R_t) la série de rendement journalier et multiplié par 100 :

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) * 100.$$

Tableau 1: Statistiques descriptives des séries de rendement de Bitcoin

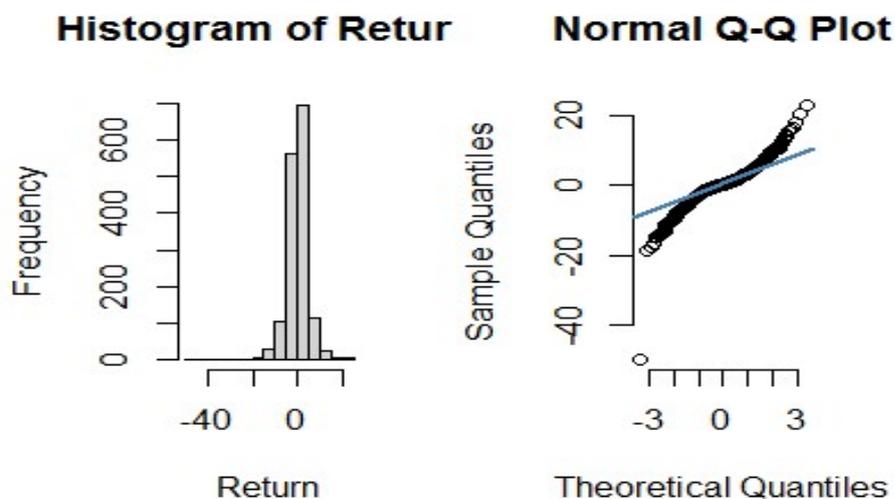
Titre	Moyenne	Min	Max	St.dev	Skew	Kurt	JB
Bitcoin	0.2636	-49.7278	22.7602	4.338	-0.968	16.626	12108.26

Source : L'auteur à partir de l'analyse des données sur R

Le tableau 1 présente les statistiques descriptives de la série de rendements de Bitcoin avec le test de normalité Jarque-Bera qui est calculé comme suit :

$$JB = T \cdot \left(\frac{skew^2}{6} + \frac{(kurt-3)^2}{24} \right)$$

où T est la taille de l'échantillon, skew et kurt sont respectivement l'asymétrie et l'aplatissement de l'échantillon. Sous l'hypothèse nulle les données suivent la loi normale, JB est distribué asymptotiquement en chi carré avec 2 degrés de liberté. Les tests ont rejeté la normalité au niveau de signification de 5%. La série de rendements de Bitcoin est négativement asymétrique. Cela indique que les rendements de Bitcoin sont non symétriques. La valeur négative de l'asymétrie indique que la distribution des séries de rendement Bitcoin est asymétrique vers la gauche. L'excès de kurtosis positif (16,626) indique que les rendements sont leptokurtiques.

Figure 2: Histogramme et tracé q-q normal des séries de rendement de Bitcoin.

Source : L'auteur à partir de l'analyse des données sur R

La figure 1 présente l'évolution de la série chronologique du prix du Bitcoin (figure à gauche) et la série de rendements de Bitcoin pour la période (figure à droite). La figure 2 est l'histogramme et le tracé normal quantile-quantile (q-q) de la série de rendement pour la même période de temps.

Tableau 2: Test de ARCH.

	Ljung box test	LM test	ADF test
P-value	0.0019	0.00152	0.01

Source : L'auteur à partir de l'analyse des données sur logiciel R

Le test Augmented Dickey Fuller (ADF) (Dickey et Fuller, 1979) est utilisé pour tester la stationnarité. D'après le Tableau 2, l'hypothèse nulle de stationnarité est acceptée au niveau de signification de 5%. Par conséquent, il n'est pas nécessaire de différencier les séries de rendement. Pour appliquer les modèles GARCH aux séries de rendements Bitcoin, la présence d'effets de stationnarité et d'ARCH dans la série de rendements résiduels est testée.

Le test de Ljung box (Engle, 2001) est utilisé pour tester la présence d'effets ARCH dans les données. l'hypothèse nulle de « pas d'autocorrélation » dans les carrés des résidus est rejetée au niveau de signification de 5%. C'est-à-dire qu'il existe une dépendance dans la série de rendements au carré de Bitcoin. En utilisant le test Multiplicateur de Lagrange (LM), l'hypothèse nulle de « pas d'effets ARCH » est rejetée au niveau de signification de 5%. D'après le test Ljung box et le test LM, on peut conclure que l'effet ARCH de la volatilité est très présent dans la série de rendement. Par conséquent, les modèles GARCH peuvent être utilisés pour modéliser les données des séries de rendements.

3.2. Résultats de l'estimation

Les statistiques descriptives des données indiquent qu'un modèle approprié de volatilité des rendements Bitcoin devrait tenir compte de sa nature variable dans le temps et de l'écart par rapport à la normalité dans la distribution des rendements Bitcoin. Toutes les estimations et les calculs sont effectués dans R Statistical Environment (R, 2008) en utilisant le package R "rugarch" développé par et Ghalanos (2013) le package R "KernSmooth" développé par Wand et Ripley (2007). Les modèles paramétriques GARCH sont estimés avec la méthode du quasi-maximum de vraisemblance de Bollerslev et Wooldridge (1992) qui donne des erreurs-types robustes.

Tableau 3: Résultats d'estimations des modèles GARCH, EGARCH et GJR

	GARCH		EGARCH		GJR-GARCH	
	Normal	Student-t	Normal	Student-t	Normal	Student-t
μ	0.327 (3.385)	0.264 (3.870)	0.282 (2.848)	0.237 (3.645)	0.257 (2.633)	0.267 (3.916)
ω	1.243 (5.464)	0.316 (1.921)	4.973 (1.107)	0.046 (1.582)	1.402 (5.214)	0.258 (1.474)
α	0.128 (6.290)	0.113 (5.442)	0.114 (4.369)	0.1357 (5.503)	0.084 (4.612)	0.120 (5.190)
β	0.818 (35.613)	0.885 (39.055)	0.7545 (16.155)	0.910 (58.616)	0.802 (31.269)	0.896 (33.643)
γ			0.1858 (1.313)	-0.1402 (-1.574)	0.101 (3.257)	-0.0369 (-1.384)

<i>log. like</i>	-4338.704	-4135.845	-4329.967	-4123.559	-4331.529	-4134.896
<i>AIC</i>	5.661	5.398	5.653	5.385	5.653	5.398
<i>BIC</i>	5.67	5.416	5.674	5.409	5.671	5.419
<i>Q(20)</i>	4.889	5.2496	5.8857	5.176	5.278	4.9664

Source : L'auteur à partir de l'estimation des modèles par logiciel R

Nous avons adapté la série du 01/02/2017 au 30/04/2021 avec le modèle standard GARCH(1,1). Compte tenu de l'existence d'effets d'asymétrie sur les marchés crypto, nous avons ajusté également les données avec les modèles EGARCH et GJR. Pour tous ces modèles, les innovations sont supposées gaussiennes ou distribuées par Student. Les paramètres estimés et les tests de Ljung-Box Q-statistics des résidus standardisés sont présentés dans le tableau 3.

Le tableau 3 résume les résultats de l'estimation du maximum de vraisemblance (MLE) des modèles GARCH(1,1), EGARCH(1,1) et GJR-GARCH(1,1) pour les rendements Bitcoin en utilisant la distribution normale et Student. D'après le tableau, la valeur de log-vraisemblance (-4123,559) est maximale pour le modèle GJR-GARCH(1,1). Les valeurs des deux critères d'information (AIC= 5.385, BIC=5.409) de EGARCH(1,1) sont minimales par rapport à GARCH(1,1)-t et GJR.GARCH(1,1)-t. Ces résultats indiquent que le modèle EGARCH(1,1)-t est le modèle optimal pour décrire la volatilité de la série de rendement de Bitcoin.

Notez que tous les paramètres de la volatilité conditionnelle sont significatifs au niveau de signification de 5%. Le coefficient de variance retardée β montre une persistance de volatilité très élevée. La somme de α et β du modèle GARCH est proche de 1, ce qui soutient la preuve d'un clustering de volatilité. Les valeurs P du test statistique Q de Ljung-Box au décalage de 20 des séries résiduelles standardisées de tous les modèles ne suggèrent pas l'autocorrélation à un niveau de signification de 5%. Ainsi, tous les modèles semblent être adéquats pour décrire la dépendance linéaire dans les séries de rendement et de volatilité.

La valeur estimée des paramètres de levier γ des modèles EGARCH et GJR avec des innovations distribuées gaussiennes et Student sont respectivement: 0,1858 / -0,1402 et 0,10 / -0,0369. Tous ces paramètres sont significatifs au seuil de 5% à l'exception du γ du modèle EGARCH avec erreurs gaussiennes. L'importance des paramètres indique l'existence d'un effet d'asymétrie, c'est-à-dire que les mauvaises nouvelles (choc négatif) ont un impact plus important sur la volatilité des rendements que les bonnes nouvelles (choc positif). Il convient également de noter que l'effet de levier estimé à partir des modèles équipés de t innovation distribuée est plus élevé que ceux avec des innovations distribuées normales. L'existence de l'effet d'asymétrie comme dans d'autres marchés boursiers matures dans le monde peut être un signe positif pour l'efficacité et l'exhaustivité du marché.

3.3. Résultats des prévisions

Bien qu'il existe plusieurs mesures différentes pour évaluer les performances de prévision de la volatilité, l'erreur moyenne absolue (MAE) et l'erreur quadratique moyenne (MSE) (voir Lui et al, 2009) sont utilisées dans cette étude. Lorsque le véritable processus de volatilité sous-jacent est inobservable, nous adoptons la suggestion d'Awartani et Corradi (2005) d'utiliser $(\sigma_t^2 = R_t -$

\bar{R})² comme approximation de la volatilité latente dans ce scénario. Les MAE et MSE pour la prévision n étape sont définis comme suit :

$$MSE = \frac{1}{N} \sum_{t=1}^N ((R_{t+n} - \bar{R})^2 - \hat{h}_t(n))^2$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |(R_{t+n} - \bar{R})^2 - \hat{h}_t(n)|$$

Ou R_{t+n} : le rendement sur l'horizon n avance à l'instant courant t ,

\bar{R} : la moyenne de rendement ,

$\hat{h}_t(n)$: la variance conditionnelle prévue sur l'horizon n avance à l'instant t.

Tableau 4: Qualité de l'ajustement des prévisions hors échantillon

		MSE	MAE
GARCH	Normal	456.52	10.65
	Student-t	525.85	10.69
EGARCH	Normal	433.8	10.63
	Student-t	449.32	10.97
GJR-GARCH	Normal	450.65	10.65
	Student-t	511.81	10.68
N-ARCH	Normal	433.30	10.3
N-GARCH	Normal	430.73	10.29

Source : L'auteur à partir de la validation des modèles par logiciel R

Les performances des prévisions de volatilité hors échantillon de divers modèles sont résumées dans le Tableau 4. Selon les deux mesures de qualité de l'ajustement, il ressort clairement que le modèle N-GARCH est le plus performant en fournissant l'erreur de prévision la plus faible. Par rapport au modèle GARCH, le modèle EGARCH améliore l'estimation de la volatilité en capturant les effets de levier. Pour le modèle GJR, il améliore légèrement le résultat de l'estimation GARCH. Ce n'est peut-être pas surprenant car l'effet asymétrique sur le marché du Bitcoin n'est pas aussi fort. Cependant, cela peut indiquer que le modèle EGARCH peut capturer plus d'effet de levier que le modèle GJR. En examinant le modèle N-ARCH, nous observons une amélioration significative du modèle N-ARCH par rapport au modèle EGARCH avec erreurs gaussiennes. Les résultats empiriques suggèrent que le modèle GARCH non paramétrique a amélioré la précision de prévision par rapport à GARCH paramétriques

4. Conclusion

La modélisation GARCH s'appuie sur les avancées dans la compréhension et la modélisation de la volatilité. Il prend en compte l'excès de kurtosis (c'est-à-dire le comportement de la queue grasse) et le regroupement de la volatilité, deux caractéristiques importantes des séries chronologiques

financières, qui sont également observables dans le cas du Bitcoin. Il est théoriquement capable de fournir des prévisions précises des variances et des covariances des rendements grâce à la modélisation de variances conditionnelles variant dans le temps. En conséquence, les modèles GARCH sont devenus très populaires dans divers domaines tels que la gestion des risques, la gestion de portefeuille et l'allocation d'actifs, la tarification des options, le change et la structure par terme des taux d'intérêt.

Dans cet article, nous avons l'intention de prédire les prix futurs du Bitcoin, l'une des crypto-monnaies les plus utilisées et les plus échangées, et d'étudier le pouvoir prédictif du modèle GARCH sur la série rendement/prix du Bitcoin. À partir des résultats prédits, nous avons réalisé que bien que les modèles GARCH soient utiles dans un large éventail d'applications financières et économiques, ils ne sont pas un modèle tout à fait efficace et approprié pour étudier la série de rendement/prix Bitcoin. L'une des principales raisons est que les modèles GARCH sont des spécifications paramétriques qui fonctionnent le mieux dans des conditions de marché relativement stables. Bien que GARCH soit explicitement conçu pour modéliser des variances conditionnelles variant dans le temps, les modèles GARCH ne parviennent souvent pas à capturer des phénomènes très irréguliers, y compris les fluctuations sauvages du marché (par exemple, les krachs...) et d'autres événements très imprévus qui peuvent conduire à des changements structurels importants, qui sont exactement ce qui s'est passé récemment sur le marché Bitcoin.

Alternativement, nous avons proposé d'appliquer le modèle GARCH non paramétrique pour modéliser la volatilité des rendements Bitcoin. Ainsi, nous avons utilisé la technique de lissage non paramétrique pour estimer le processus de diffusion non observé. Notre application empirique aux prix quotidiens du Bitcoin ont mis en évidence une amélioration significative du pouvoir prédictif du modèle non paramétrique par rapport aux modèles GARCH. La supériorité de la performance prédictive de la volatilité hors échantillon du modèle GARCH non paramétrique est en outre vérifiée. Pour des travaux de recherches futures, nous avons l'intention d'appliquer cette technique aux calculs de la valeur à risque et d'illustrer les avantages potentiels qui peuvent être dérivés de l'application de ce modèle dans le contexte de la couverture contre les fluctuations extrêmes des prix sur le marché des crypto-monnaies.

Références bibliographiques

- Bollerslev, T. 1986. « Generalized autoregressive conditional heteroskedasticity ». *Journal of Econometrics*, 31, 307-327
- Bollerslev, T., Wooldridge, J., 1992. « Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. » *Econ. Rev.* 11, 143-172
- Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. « ARCH models. In Engle, R.F., McFadden, D. (Eds.) », *Handbook of Econometrics*, Amsterdam Elsevier Science, pp. 2959-3038.
- Brandvold, M., Moln, P., Vagstad, K., and Valstad, O. C. A. 2015. « Price discovery on Bitcoin exchanges ». *Journal of International Financial Markets, Institutions and Money* 36: 18- 35.
- Bouoiyour, J., Selmi, R., Tiwari, A.K. and Olayeni, O.R. 2016. « What drives Bitcoin price? ». *Economics Bulletin* 36(2): 843-850.
- Bossaerts P., Hardle W., Hafner C. (1996) « Foreign Exchange Rates Have Surprising Volatility. In: Robinson P.M., Rosenblatt M. (eds) *Athens Conference on Applied Probability and Time Series Analysis*. « *Lecture Notes in Statistics*, vol 115. Springer, New York, NY.
- Bohlmann, P., McNeil, A., 2002. « An algorithm for nonparametric GARCH modelling ». *Computational Statistics and Data Analysis* 40, 665-683.
- Engle R ,2001, Garch 101: « The use of arch/garch models in applied econometrics ». *J Econ Perspect*15: 157-168.
- Engle, R.F., Bollerslev, T., 1986. « Modelling the persistence of conditional variances ». *Econ. Rev.* 5, 1-50.
- Engle, R.F., Gonzalez-Rivera, G. 1991. « Semiparametric ARCH models ». *Journal of Business and Economic Statistics*, 9(4), 345-359.
- Fan, J, Yao, Q. 1998, « Efficient estimation of conditional variance functions in stochastic regression, » *Biometrika*, Volume 85, Issue 3, Pages 645-660.
- Ghalanos. A .2011, « rgarch: A package for flexible garch modelling in R ». Version 1.89.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. « On the relation between the expected value and the volatility of the nominal excess return on stocks ». *J. Finance* 48, 1779-1801.
- Hardle, W., Chen, R. 1995. « Nonparametric Time Series Analysis, a Selective Review with Examples. « *Proceedings of the 50th Session of the ISI*.
- Hardle, W., Lotkepohl, H., Chen, R. 1997. *A Review of Nonparametric Time Series Analysis*. *International Statistical Review/Revue Internationale de Statistique*, 65(1), 49-72.
- Hou, A., Suardi, S., 2012. « A Nonparametric GARCH Model Of Crude Oil Price Return Volatility. « *Energy Economics*, 34(2), 618-626.

Nelson, D.B., 1991. « Conditional heteroskedasticity in asset returns: a new approach ». *Econometrica* 59, 347-1313.

Pagan, A.R., Schwert, G.W., 1990. « Alternative models for conditional stock volatility. » *J. Econ.* 45, 267-290.

Tostheim, D., Auestad, B.H. 1994. » Nonparametric Identification of Nonlinear Time Series: Projections ». *Journal of the American Statistical Association*, 89(428), 1398-1409.

Tschernig, R., Yang, L. 2000. « Nonparametric Lag Selection for Time Series ». *Journal of Time Series Analysis*, 21(4), 457-487.

Wand, M., Ripley, B , 2007 . « KernSmooth : Functions for kernel smoothing for Wand and Jones (1995) ». R package version 2.22-20.

Wang, L., Feng, C., Song, Q., Yang, L. 2012. « Efficient Semiparametric GARCH Modeling Of Financial Volatility ». *Statistica Sinica*, 22, 249-270.