



Projet de Gestion des Risques Extrêmes

Groupe :

MOREAU Evan
ROGÉE Marius

Mars 2025

Contents

1	Introduction	2
2	VaR non paramétrique	4
2.1	VaR historique	4
2.2	VaR bootstrap	4
2.3	Vérification de la VaR	4
3	VaR gaussienne	6
3.1	VaR gaussienne journalière	6
3.2	VaR gaussienne 10 jours	6
3.3	VaR Black-Scholes	6
3.4	VaR EWMA	7
3.5	Test de normalité des log rendements	8
4	VaR skew-Student	9
5	Expected Shortfall	10
6	Backtesting	11
6.1	Protocole	11
6.2	Application à nos données	11
7	VaR TVE : Approche Maxima par bloc	13
8	VaR TVE : Approche Peak Over Threshold	15
9	VaR GARCH	17
9.1	Vérification de l'applicabilité du modèle	17
9.2	Vérification de la pertinence du modèle AR(1)	17
9.3	Interprétation des résultats	18
9.4	Analyse des résidus	18
9.5	Tests supplémentaires	19
9.6	Entraînement du modèle AR(1)-GARCH(1,1)	21
9.7	Vérification des résidus du modèle AR(1)-GARCH(1,1)	21
9.8	Test d'Indépendance des Résidus	22
9.9	Modélisation de la dynamique des rendements logarithmiques	24

1 Introduction

L'objectif de cette étude est d'examiner différentes méthodes de calcul de la Value at Risk (VaR), une mesure fondamentale de la gestion des risques, permettant d'estimer la perte maximale probable d'un investissement sur une période donnée, avec un niveau de confiance précis. Ce rapport est accompagné d'une interface dynamique développée avec Dash, offrant la possibilité de moduler les principaux paramètres et d'observer en temps réel l'impact sur les indicateurs.

Les données utilisées proviennent des prix de clôture du CAC40, extraites de Yahoo Finance. Par défaut, la période couverte va du 15 octobre 2008 au 11 juin 2024, mais cette plage peut être modifiée via l'interface. Nous avons effectué un nettoyage des données, notamment pour gérer les valeurs manquantes, et calculé les log-rendements correspondants. Les données ont ensuite été divisées en un échantillon d'apprentissage et un échantillon de test, la séparation étant fixée par défaut au 27 juillet 2022.

Ci-dessous, un graphique montre l'évolution des log-rendements en fonction du temps, ainsi que l'évolution du cours de bourse sur la même période.

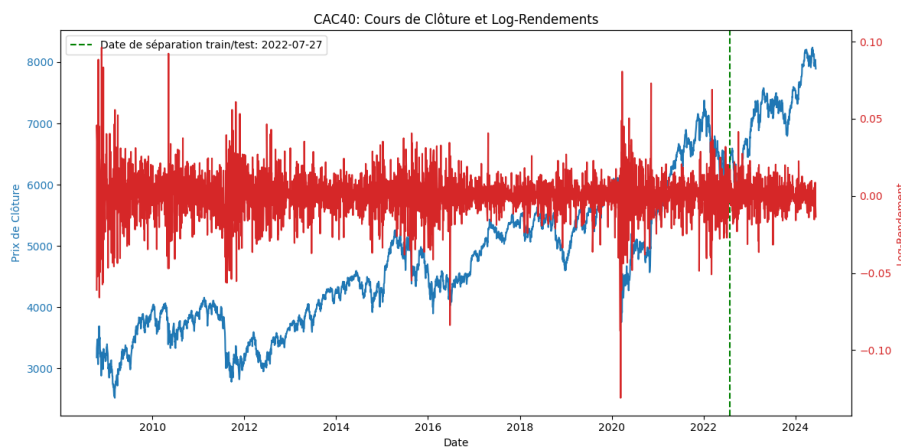


Figure 1: Evolution des rendements logarithmiques et du cours du CAC40 en fonction du temps.

On observe une tendance haussière du cours du CAC 40. De plus, l'analyse des log-rendements met en évidence la formation de clusters : certaines périodes sont relativement stables, tandis que d'autres connaissent une forte volatilité, notamment lors de la crise du Covid-19.

Voici deux tableaux récapitulatifs des statistiques descriptives de nos échantillons:

	Close	Log_Ret
count	3522.000000	3522.000000
mean	4601.646103	0.000173
std	1022.208890	0.013904
min	2519.290039	-0.130983
25%	3798.562439	-0.006095
50%	4463.850098	0.000582
75%	5318.639893	0.006855
max	7376.370117	0.096169

Table 1: Statistiques descriptives de l'échantillon d'entraînement.

	Close	Log_Ret
count	479.000000	479.000000
mean	7178.986868	0.000500
std	588.579420	0.008812
min	5676.870117	-0.036484
25%	6855.710205	-0.004070
50%	7260.729980	0.000767
75%	7487.669922	0.005538
max	8239.990234	0.041504

Table 2: Statistiques descriptives de l'échantillon de test.

Différentes approches de calcul de la VaR seront mises en place, notamment la VaR non paramétrique, la VaR gaussienne, la VaR basée sur la loi skew-Student, la VaR fondée sur l'analyse des extrêmes de la distribution, ainsi que la VaR GARCH. Ces méthodes seront ensuite implémentées et comparées.

Le projet comprend également une validation des modèles à l'aide de tests statistiques, de backtests et du calcul de l'Expected Shortfall (ES) afin d'affiner l'analyse des risques.

2 VaR non paramétrique

2.1 VaR historique

La VaR historique journalière est définie comme suit :

$$\text{VaR}_\alpha = \min\{x \mid P(L \leq x) \geq \alpha\},$$

où L représente les pertes historiques et α est le niveau de confiance.

La VaR historique journalière sur notre échantillon d'apprentissage pour $\alpha = 99\%$ est donnée par :

$$\text{VaR}_\alpha = -0.041.$$

2.2 VaR bootstrap

La VaR historique bootstrap est la moyenne de l'ensemble des VaR générées à partir de sous-échantillons aléatoires. Elle est définie comme suit :

$$\text{VaR}_\alpha^{\text{bootstrap}} = \frac{1}{N} \sum_{i=1}^N \text{VaR}_{\alpha,i},$$

où N est le nombre de sous-échantillons générés et $\text{VaR}_{\alpha,i}$ est la VaR calculée pour le sous-échantillon i .

On peut également donner un intervalle de confiance de niveau α_{IC} défini comme suit :

$$\left[\text{quantile}_{\alpha_{\text{IC}}}(\text{VaR}_\alpha^{\text{bootstrap}}), \text{quantile}_{1-\alpha_{\text{IC}}}(\text{VaR}_\alpha^{\text{bootstrap}}) \right],$$

où les quantiles sont calculés à partir de la distribution des VaR bootstrap.

On obtient les résultats suivants :

- VaR bootstrap à 99% : -0.041
- Intervalle de confiance à 99% pour la VaR : $[-0.044, -0.038]$

2.3 Vérification de la VaR

Nous allons désormais tester si la VaR déterminée à partir de notre échantillon d'entraînement est correcte.

On obtient qu'il n'y a aucun dépassement de la VaR sur l'échantillon de test. Ce résultat s'explique par le fait que la période de test est plutôt stable. Voir le graphique 1.

Ce pourcentage d'exceptions est statistiquement différent de celui attendu. En effet, nous avons généré un intervalle de confiance à partir d'une loi binomiale avec les mêmes paramètres de seuil α . Nous avons donc créé une loi binomiale de paramètres n (taille de l'échantillon de test) et $p = 1 - \alpha$ (probabilité de succès). L'intervalle de confiance binomial à 90% est donné par : $[0.42\%, 1.88\%]$. Cet intervalle représente le pourcentage attendu de pertes inférieures à la VaR. Nos données réelles se trouvent en dehors de cet intervalle.

3 VaR gaussienne

3.1 VaR gaussienne journalière

Pour calculer la VaR gaussienne, nous modélisons nos rendements logarithmiques par une loi normale. Soit R la variable aléatoire représentant les rendements logarithmiques, qui suit une loi normale $R \sim \mathcal{N}(\mu, \sigma^2)$, où μ est la moyenne et σ l'écart-type. La VaR gaussienne à un niveau de confiance α est alors donnée par le quantile de la distribution normale :

$$\text{VaR}_\alpha = \mu + \sigma \cdot \Phi^{-1}(1 - \alpha),$$

où $\Phi^{-1}(1 - \alpha)$ est l'inverse de la fonction de répartition de la loi normale standard, et α est le niveau de confiance.

On trouve ainsi la VaR gaussienne à 99% : -0.032 .

3.2 VaR gaussienne 10 jours

La méthode du scaling permet de calculer la VaR gaussienne sur un horizon de 10 jours. La formule utilisée est la suivante :

$$\text{VaR}_{\alpha,10j} = \text{VaR}_{\alpha,1j} \times \sqrt{10},$$

où $\text{VaR}_{\alpha,1j}$ est la VaR gaussienne à 1 jour, et $\sqrt{10}$ est le facteur de mise à l'échelle pour un horizon de 10 jours.

On obtient ainsi la VaR gaussienne à 99% pour un horizon de 10 jours :

$$\text{VaR}_{\alpha,10j} = -0.102$$

3.3 VaR Black-Scholes

Une autre méthode pour calculer la VaR à 10 jours consiste à utiliser la méthode de diffusion d'un actif. Nous appliquons le modèle de diffusion suivant pour le prix de l'actif S :

$$dS = S \cdot \mu \cdot dt + S \cdot \sigma \cdot Z \cdot \sqrt{dt}$$

où :

- S est le prix de l'actif,
- μ est le rendement attendu de l'actif,
- σ est la volatilité de l'actif,

- Z est une variable aléatoire suivant une distribution normale standard $Z \sim \mathcal{N}(0, 1)$,
- dt est l'intervalle de temps, que nous prendrons comme étant égal à 1 jour.

Pour cette méthode, nous procéderons comme suit :

- S_0 est la valeur du cours de clôture à la dernière date de l'échantillon d'apprentissage.
- μ et σ sont les paramètres estimés dans l'étape précédente.

La simulation de la diffusion du prix sur 10 jours permettra de calculer la VaR gaussienne en simulant plusieurs trajectoires possibles pour le prix de l'actif et en analysant la distribution des pertes associées à ces trajectoires. La Value at Risk (VaR) Gaussienne à 99% pour un horizon de 10 jours, calculée via la méthode de Monte Carlo et en utilisant l'équation du modèle de Black-Scholes, est estimée à :

$$\text{VaR}_{99\%}^{10j} = -0.097.$$

3.4 VaR EWMA

La méthode EWMA (Exponential Weighted Moving Average) attribue un poids plus important aux observations récentes, permettant ainsi une meilleure réactivité aux changements du marché. Le paramètre d'ajustement λ contrôle la décroissance exponentielle des poids au fil du temps.

Les poids sont définis comme suit :

$$\begin{aligned}\omega_i(\lambda) &= \lambda^i (1 - \lambda), \\ \tilde{\omega}_i(\lambda) &= \frac{\omega_i(\lambda)}{\sum_{j=0}^T \omega_j(\lambda)}.\end{aligned}$$

À partir de ces poids, les estimations de l'espérance et de la variance des rendements sont calculées selon :

$$\begin{aligned}\hat{\mu}(\lambda) &= \sum_{i=0}^T \tilde{\omega}_i(\lambda) r_{T-i}, \\ \hat{\sigma}^2(\lambda) &= \sum_{i=0}^T \tilde{\omega}_i(\lambda) [r_{T-i} - \hat{\mu}(\lambda)]^2.\end{aligned}$$

Les résultats de l'estimation de la VaR EWMA à 99% pour différentes valeurs de λ sont présentés dans le tableau ci-dessous :

On observe qu'une augmentation du paramètre λ entraîne une diminution de la VaR ainsi qu'une réduction du nombre d'exceptions.

λ	$\hat{\mu}$	$\hat{\sigma}$	VaR EWMA à 99%	Nombre d'exceptions	% d'exceptions
0.90	0.0019	0.01103	-0.026	4	0.84%
0.95	0.00074	0.01277	-0.030	3	0.63%
0.99	-0.00032	0.01427	-0.033	1	0.21%

Table 3: Résultats de la VaR EWMA à 99% pour différentes valeurs de λ .

3.5 Test de normalité des log rendements

Voici un qq-plot afin d'observer si nos données suivent une distribution gaussienne:

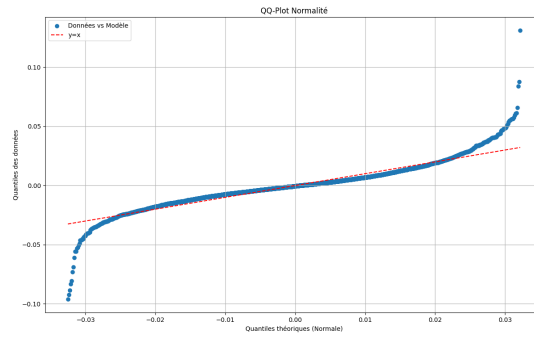


Figure 2: qq-plot gaussien.

La distribution semble normale, à l'exception des valeurs extrêmes qui posent problème. Nous allons compléter cette analyse graphique par un test statistique, en utilisant le test de Shapiro-Wilk.

Hypothèses :

- H_0 : L'échantillon suit une distribution gaussienne.
- H_1 : L'échantillon ne suit pas la distribution gaussienne.

La p-valeur obtenue pour ce test est quasi nulle, ce qui permet de rejeter l'hypothèse nulle (H_0) et de conclure que la distribution n'est pas normale.

Nos données ne s'adaptent pas bien à une modélisation par une distribution gaussienne, notamment en raison de la différence dans l'épaisseur des queues de distribution. Une alternative consiste à modéliser les log-rendements à l'aide d'une loi prenant en compte cette caractéristique. La section suivante est consacrée à la modélisation avec la loi de Skew Student.

4 VaR skew-Student

Nous avons modélisé les log-rendements à l'aide d'une loi Skew-Student. Les paramètres ont été estimés par la méthode du maximum de vraisemblance et sont les suivants :

$$\mu = 1.1644 \times 10^{-3}$$

$$\sigma = 1.5329 \times 10^{-2}$$

$$\nu = 2.9473$$

$$\lambda = -0.1747$$

Afin d'évaluer l'adéquation du modèle aux données, nous avons examiné le QQ-plot et l'histogramme des résidus. Ces visualisations indiquent que la loi Skew-Student ne s'ajuste pas parfaitement aux données observées.

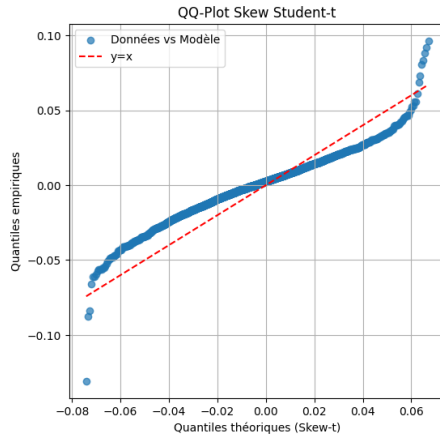


Figure 3: QQ-plot sous une loi Skew-Student.

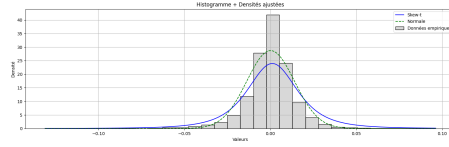


Figure 4: Histogramme sous une loi Skew-Student.

La modélisation Skew-Student accorde davantage d'importance aux valeurs extrêmes, ce qui se traduit par une VaR significativement plus faible que celle obtenue avec le modèle gaussien. Ainsi, la VaR à 99% sous la loi Skew-Student est égale à : -0.074.

5 Expected Shortfall

L'Expected Shortfall empirique est défini par la formule suivante :

$$ES_{\text{emp}} = E[X \mid X \leq \text{VaR}_\alpha]$$

où X représente les pertes, et VaR_α est la Value at Risk d'ordre α .

L'Expected Shortfall théorique est défini par :

$$ES_{\text{théorique}} = \int_{-\infty}^{\text{VaR}_\alpha} x f(x) dx$$

où $f(x)$ est la fonction de densité de la loi de probabilité sous-jacente.

Nous avons calculé l'Expected Shortfall pour différentes approches de la VaR, et les résultats obtenus sont présentés dans le tableau suivant :

Méthode	Expected Shortfall à 99%
Expected Shortfall historique (ES)	-0.055
Expected Shortfall gaussien (ES) empirique	-0.044
Expected Shortfall gaussien (ES) théorique	-0.037
Expected Shortfall Skew-Student (ES) empirique	-0.101
Expected Shortfall Skew-Student (ES) théorique	-0.213

Table 4: Comparaison des valeurs d'Expected Shortfall à 99% pour différents modèles.

6 Backtesting

L'objectif principal du backtesting est de vérifier que la Value at Risk (VaR) estimée sur la période d'apprentissage reste cohérente et efficace lorsqu'elle est appliquée à la période de test. Plus précisément, le but est de montrer que le pourcentage de dépassement de la VaR observé pendant la période de test est conforme à l'ordre de confiance fixé lors de la calibration du modèle. En outre, il est important de vérifier que les dépassements de la VaR sont homogènes dans le temps, c'est-à-dire qu'ils ne présentent pas de variations significatives sur différentes périodes.

6.1 Protocole

Le protocole de backtesting suivant est proposé pour tester la robustesse et l'adéquation du modèle de VaR aux données réelles :

1. **Calibrage initial** : Le modèle de VaR est calibré sur la période d'apprentissage, afin d'estimer les paramètres nécessaires à la prévision de la VaR.
2. **Backtesting ex-post** : Une fois le modèle calibré, il est testé ex-post sur chaque jour suivant la période de calibrage. Pour chaque jour de l'échantillon de test, la VaR est recalculée en utilisant la même méthodologie et les paramètres estimés lors de la période d'apprentissage. Le nombre de dépassements est comparé au niveau de confiance choisi pour la VaR.
3. **Mesure de recalibrage** : Un indicateur est mis en place pour déterminer si le modèle nécessite un recalibrage. Par exemple, un recalibrage peut être envisagé si le pourcentage de dépassements observé durant la période de test dépasse largement le niveau de confiance initial (indiquant que le modèle sous-estime les risques), ou si les dépassements montrent des irrégularités au fil du temps. Un test statistique, tel qu'un test de proportion ou une analyse de la distribution des dépassements sur différentes sous-périodes, peut être utilisé pour évaluer cette homogénéité.

6.2 Application à nos données

Nous avons appliqué cette méthodologie à nos données, en commençant par la VaR à 99% calculée à l'aide de la méthode historique.

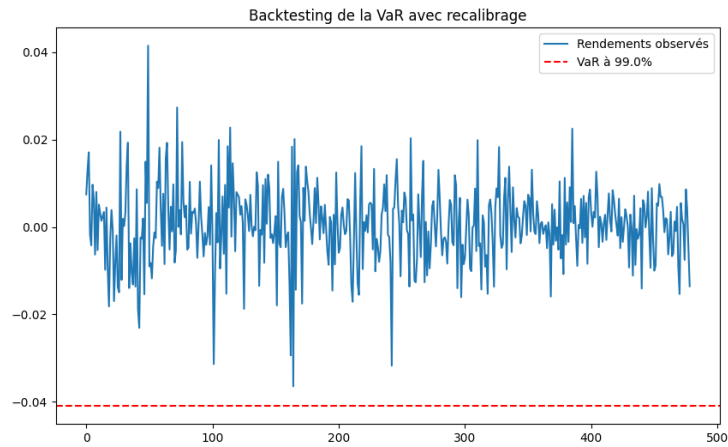


Figure 5: Backtesting de la VaR historique.

Aucun dépassement de la VaR n'a été observé, ce qui confirme les résultats obtenus dans la section sur la VaR non paramétrique.

L'analyse des dépassements de la VaR gaussienne a également été effectuée.

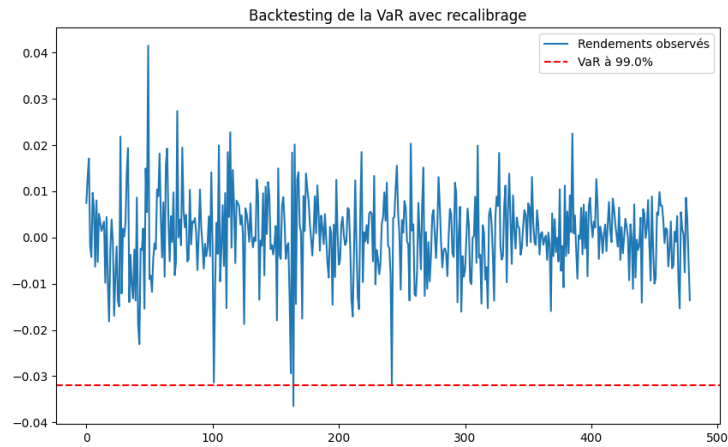


Figure 6: Backtesting de la VaR gaussienne.

On observe un taux de dépassements de 0,2%, bien inférieur au taux attendu. De plus, les dépassements sont espacés et homogènes, ce qui suggère une bonne performance du modèle dans cette configuration.

7 VaR TVE : Approche Maxima par bloc

La Théorie des Valeurs Extrêmes (TVE) s'intéresse aux queues de distribution extrêmes, notamment les maxima et les excès par rapport aux valeurs supérieures. Dans notre étude, nous avons converti nos rendements logarithmiques en pertes logarithmiques.

L'approche des maxima par bloc, que nous avons choisie, consiste à découper nos pertes logarithmiques en blocs mensuels et à prendre le maximum de chaque bloc. Le graphique ci-dessous montre les points sélectionnés.

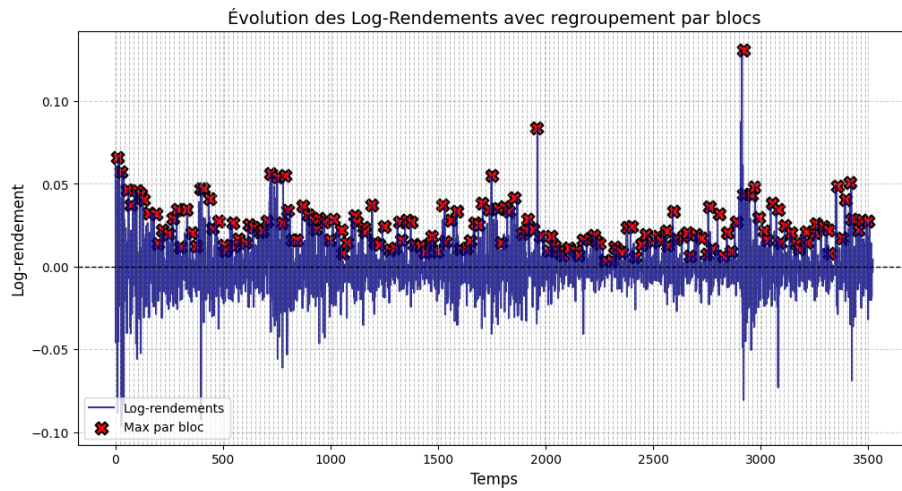


Figure 7: Maximum par bloc mensuel.

Cela nous donne une série de maxima que nous cherchons à modéliser. Une première solution envisagée est la modélisation à l'aide d'une loi de Gumbel, qui est un cas particulier de la loi des valeurs extrêmes généralisée (GEV) avec le paramètre $\xi = 0$.

Voici le graphique du Gumbel plot :

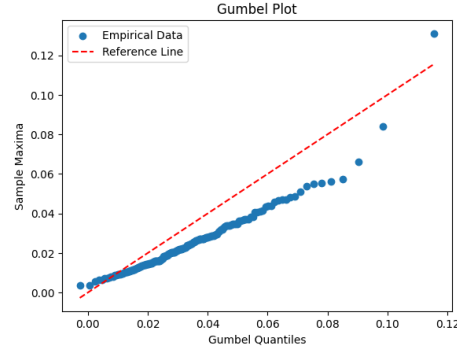


Figure 8: Gumbel plot.

Un test du rapport de vraisemblance est utilisé pour tester l'hypothèse nulle $H_0 : \xi = 0$ (loi de Gumbel). La p-value de ce test est inférieure à 0.05, ce qui nous permet de rejeter l'hypothèse nulle. Cela indique, comme le montre le graphique du Gumbel plot, que nos données ne sont pas adaptées à une loi de Gumbel. Nous préférons donc une loi GEV.

Nous entraînons un modèle GEV, dont les paramètres estimés sont les suivants :

- Shape : -0.14819
- Loc : 0.01766
- Scale : 0.00961

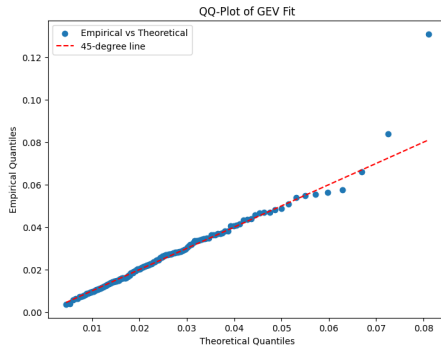


Figure 9: QQ plot de la loi GEV.

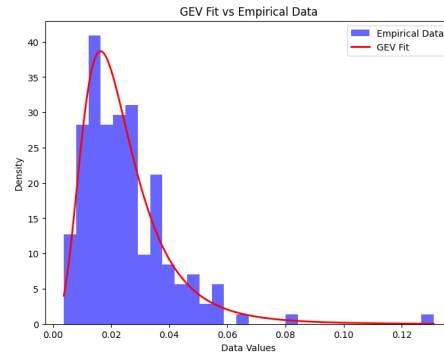


Figure 10: Histogramme ajusté avec la loi GEV.

Nos données extrêmes s'ajustent bien à la distribution GEV. Nous pouvons désormais calculer la Value at Risk TVE par bloc en prenant le quantile d'ordre α de la distribution GEV avec les paramètres estimés. Cela nous donne la VaR TVE à 99% via l'approche des maxima par bloc : -0.081.

8 VaR TVE : Approche Peak Over Threshold

Dans cette section, les pertes extrêmes sont définies comme les pertes supérieures ou égales à un certain seuil u . La première étape consiste à déterminer un tel seuil.

Soit u le seuil, les excès au-dessus de u sont définis par :

$$X_u = X - u \quad \text{pour} \quad X \geq u$$

où X représente les observations des pertes.

Un moyen couramment utilisé pour sélectionner ce seuil est de tracer le Mean Residual Life (MRL). Le MRL est défini comme la moyenne des excès au-dessus du seuil u , c'est-à-dire :

$$MRL(u) = \frac{1}{n_u} \sum_{i=1}^{n_u} (X_i - u)$$

où n_u est le nombre d'excès au-dessus du seuil u .

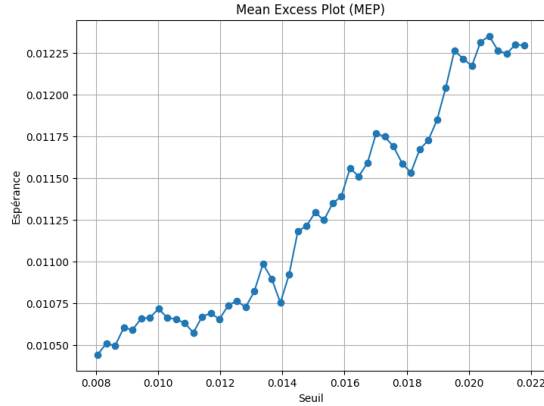


Figure 11: Graphique du Mean Residual Life (MRL).

Sur ce graphique, on peut observer que pour un seuil de 0.014 il y a une hausse rapide de l'espérance. On s'attend à ce que le seuil soit dans cette zone. Le seuil est suffisamment élevé pour respecter les conditions asymptotiques.

Pour calibrer automatiquement le seuil u dans le cadre de l'analyse Peak Over Threshold (POT) et de l'ajustement de la loi de Pareto généralisée (GPD), l'objectif est de choisir un seuil u qui minimise l'erreur d'ajustement tout en garantissant que ce seuil est suffisamment élevé pour capturer les événements extrêmes, sans être trop élevé pour éviter une perte significative d'informations.

Nous avons utilisé un critère basé sur la validité de l'ajustement GPD, à savoir l'AIC, pour déterminer le seuil optimal. Le seuil retenu est $u = 0.015$, qui minimise l'AIC à -2466.863.

Les paramètres optimaux estimés pour la GPD sont les suivants :

- Shape : 0.10662
- Scale : 0.00999
- Location : 0

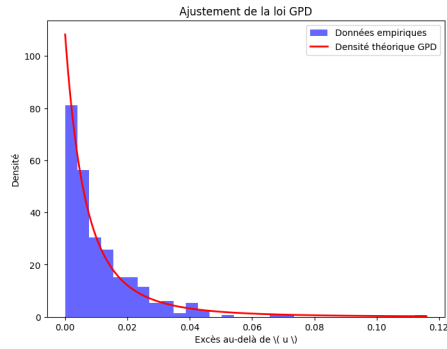


Figure 12: Histogramme des excès ajustés avec la loi GPD.

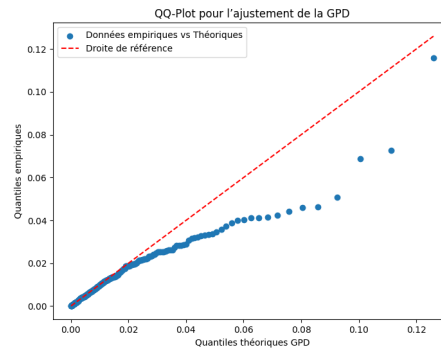


Figure 13: QQ plot de l'ajustement GPD.

Cependant, d'après les graphiques précédents, il semble que la modélisation ne suit pas correctement nos données. Cela peut indiquer la nécessité d'explorer d'autres modèles ou de reconsidérer le choix des données extrêmes.

La VaR TVE selon l'approche Peak Over Threshold (POT) est déterminée en prenant le quantile d'ordre α de la distribution modélisée. En prenant $\alpha = 99\%$, nous obtenons :

$$\text{VaR POT}(\alpha = 99\%) = -0.064$$

9 VaR GARCH

Dans cette section, nous modélisons nos rendements logarithmiques à l'aide d'un modèle AR(1)-GARCH(1,1), puis nous calculerons la VaR à partir de ce modèle.

9.1 Vérification de l'applicabilité du modèle

La première étape consiste à vérifier l'applicabilité d'un tel modèle.

Nous commençons par nous assurer que les rendements logarithmiques sont stationnaires. Pour ce faire, nous utilisons le test ADF (Augmented Dickey-Fuller). La p-value obtenue est quasi nulle, ce qui permet de rejeter l'hypothèse nulle H_0 , qui stipule que la série temporelle possède une racine unitaire, autrement dit qu'elle est non stationnaire. Par conséquent, nous pouvons conclure que la stationnarité est vérifiée.

9.2 Vérification de la pertinence du modèle AR(1)

L'étape suivante consiste à vérifier si un modèle AR(1) (autoregressive d'ordre 1) est une bonne représentation des pertes logarithmiques historiques. Une méthode couramment utilisée pour cela est l'analyse du PACF (Partial Autocorrelation Function), qui permet d'étudier l'autocorrélation partielle d'une série temporelle. Cette méthode mesure la corrélation entre une variable et ses retards tout en contrôlant l'effet des retards intermédiaires. Si des pics significatifs apparaissent, cela suggère l'utilisation d'un modèle autorégressif.

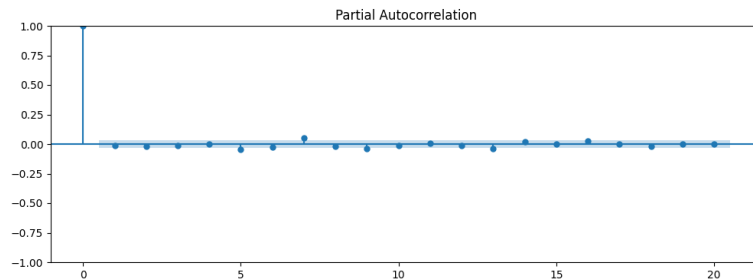


Figure 14: PACF des pertes logarithmiques.

Le graphique du PACF ne montre aucun pic significatif, ce qui suggère que l'utilisation d'un modèle AR(1) n'est pas pertinente. Toutefois, nous allons entraîner ce modèle pour vérifier cette hypothèse.

9.3 Interprétation des résultats

Variable	coef	std err	z	P-value	[0.025	0.975]
const	-0.0002	0.000	-0.743	0.458	-0.001	0.000
ar.L1	-0.0101	0.012	-0.847	0.397	-0.034	0.013
sigma2	0.0002	2.15e-06	89.670	0.000	0.000	0.000

Table 5: Résultats du modèle AR(1)

D'après le tableau précédent, la constante et le coefficient associé au retard 1 (*ar.L1*) sont non significatifs. Le modèle AR(1) ne capture donc pas correctement l'information contenue dans les données. À ce stade, il serait pertinent d'envisager d'autres modèles tels que les modèles MA, ARMA, ARIMA ou ARI-MAX. Cependant, étant donné que l'objectif de cette section est de modéliser un AR(1)-GARCH(1,1), nous poursuivons avec ce modèle.

9.4 Analyse des résidus

Les hypothèses classiques d'un modèle AR stipulent que les résidus doivent être homoscédastiques, centrés et non autocorrélés dans le temps. Nous allons vérifier si ces propriétés sont respectées dans notre modèle.

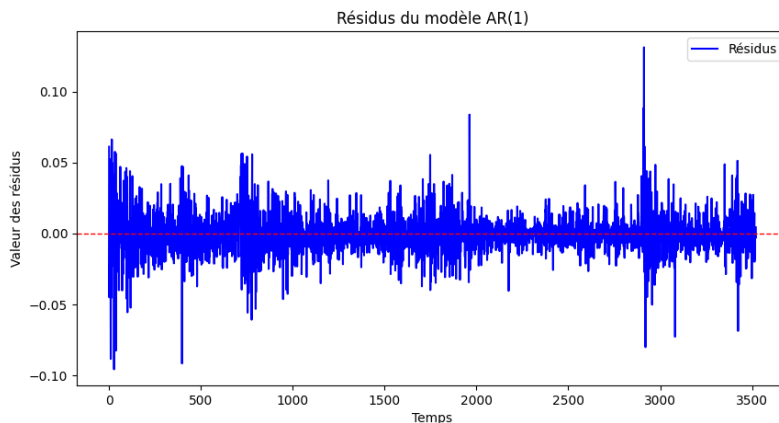


Figure 15: Résidus du modèle AR(1).

D'après l'observation des résidus, ces derniers semblent être centrés, cependant, ils présentent des caractéristiques d'hétéroscédasticité, à savoir une variance non constante, et montrent des autocorrélations temporelles. En effet, on observe des périodes de forte volatilité suivies de phases plus stables, ce qui indique une variabilité dans la dispersion des erreurs.

L'utilisation d'un modèle GARCH est alors pertinente pour modéliser de tels résidus, où la variance est conditionnellement hétéroscédastique. Cette hétéroscédasticité suggère clairement l'utilisation d'un modèle GARCH pour mieux représenter la dynamique des résidus.

9.5 Tests supplémentaires

Pour valider cette hypothèse, nous avons réalisé quelques tests supplémentaires :

- Le test d'Engle ARCH a donné une p-value proche de zéro, ce qui indique la présence d'hétéroscédasticité conditionnelle dans les résidus.
- Nous avons également appliqué un test de Ljung-Box sur les résidus au carré afin de tester la présence d'autocorrélation dans la variance des erreurs. Pour chacun des 20 premiers lags, la p-value est inférieure à 0.05, ce qui nous permet de rejeter l'hypothèse nulle d'absence d'autocorrélation. Cela confirme la présence d'autocorrélation dans la variance des erreurs, suggérant une hétéroscédasticité conditionnelle.

Lag	Statistique Ljung-Box	p-value
1	48.2993	3.66e-12
2	154.6605	2.61e-34
3	378.4322	1.04e-81
4	534.4946	2.31e-114
5	628.6791	1.28e-133
6	737.5090	4.86e-156
7	780.9232	2.43e-164
8	919.9348	2.83e-193
9	1085.2025	7.23e-228
10	1157.7167	1.90e-242
11	1223.7566	1.23e-255
12	1271.7013	6.23e-265
13	1304.3170	6.23e-271
14	1356.4813	3.79e-281
15	1370.5422	3.59e-283
16	1387.3706	8.44e-286
17	1436.6315	2.09e-295
18	1497.4942	1.65e-307
19	1509.7830	3.51e-309
20	1532.6168	0.0000e+00

Table 6: Résultats du test de Ljung-Box sur les résidus au carré

Enfin, pour visualiser l'hétéroscédasticité, nous pouvons également examiner les résidus au carré.

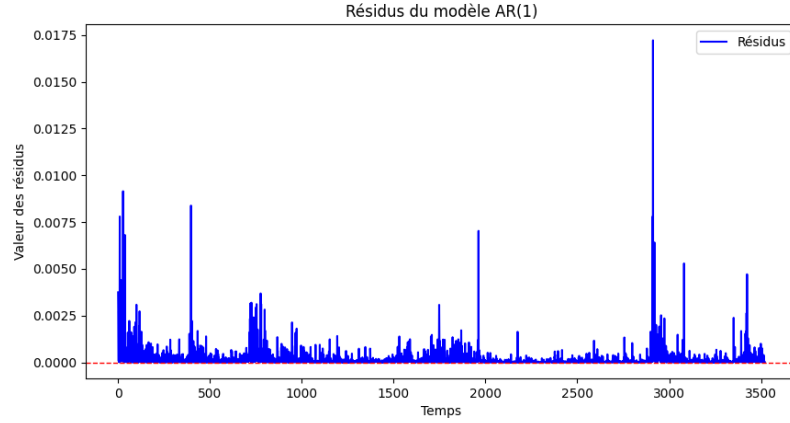


Figure 16: Résidus au carré du modèle AR(1).

Encore une fois, des périodes de forte et faible volatilité observées dans les résidus indiquent une hétéroscédasticité conditionnelle, caractéristique des modèles GARCH.

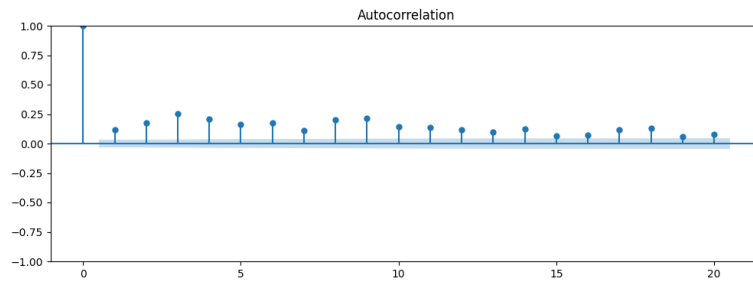


Figure 17: Fonction d'Autocorrélation (ACF) des résidus au carré du modèle AR(1).

La figure 17 présente l'ACF des résidus au carré. On y remarque plusieurs points significatifs, ce qui est typique des modèles GARCH, où l'autocorrélation dans les résidus au carré reflète l'hétéroscédasticité conditionnelle.

L'ensemble des observations et des tests réalisés nous conduit à la conclusion qu'un modèle GARCH est adapté pour cette série temporelle. Par conséquent, nous optons pour un modèle AR(1)-GARCH(1,1) afin de capturer à la fois les dépendances temporelles et la variabilité conditionnelle de la variance.

9.6 Entraînement du modèle AR(1)-GARCH(1,1)

Voici les résultats obtenus lors de l'entraînement du modèle AR(1)-GARCH(1,1). La méthode d'estimation des paramètres et celle du maximum de vraisemblance en supposant des résidus provenant d'un bruit blanc gaussien :

Mean Model			
Variable	coef	std err	P value
Const	-0.0606	0.01881	0.001274
y[1]	-0.0225	0.01939	0.247
Volatility Model			
Variable	coef	std err	P value
omega	0.0442	0.01452	0.002315
alpha[1]	0.1192	0.02610	4.92e-06
beta[1]	0.8579	0.02814	4.35e-204

Table 7: Sortie du modèle AR(1)-GARCH(1,1)

Le tableau ci-dessus présente les résultats du modèle. Le coefficient de retard 1 du modèle AR n'est pas significatif, ce qui indique que le modèle AR(1) ne capte pas suffisamment l'information de la série temporelle. En revanche, les coefficients du modèle de volatilité sont tous significatifs, ce qui confirme la pertinence du modèle GARCH.

9.7 Vérification des résidus du modèle AR(1)-GARCH(1,1)

Afin de valider la pertinence du modèle AR(1)-GARCH(1,1), nous vérifions que ses résidus suivent un comportement de bruit blanc, c'est-à-dire qu'ils sont indépendamment et identiquement distribués (i.i.d.) et suivent une distribution gaussienne.

La figure 18 présente la volatilité calibrée, les résidus et les résidus studentisés du modèle.

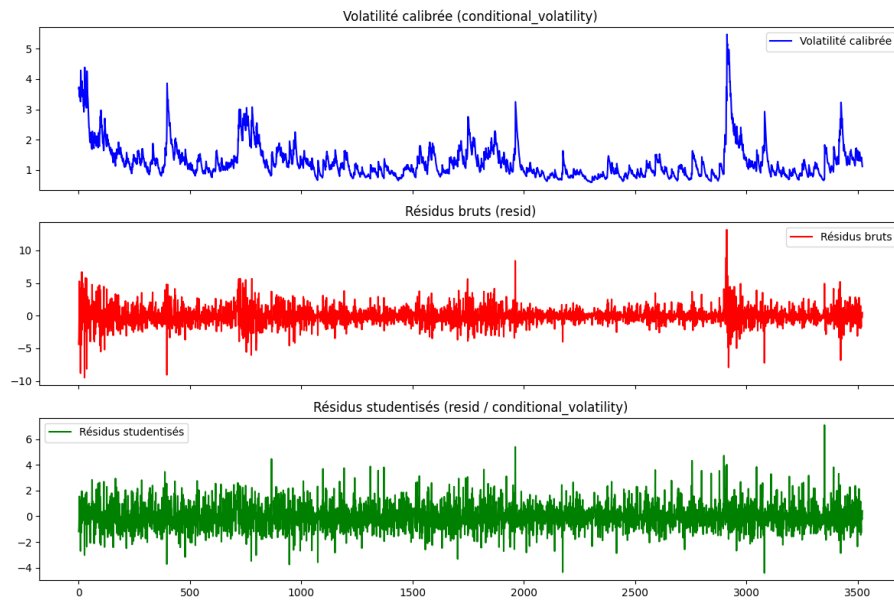


Figure 18: Volatilité calibrée, résidus et résidus studentisés.

Les définitions des différentes séries sont les suivantes :

- **Volatilité calibrée** : correspond aux erreurs brutes du modèle, soit la différence entre la série observée et la prévision, sans prise en compte de la volatilité estimée.
- **Résidus** : écart entre les valeurs observées et prévues, ajusté en fonction de la volatilité conditionnelle estimée, permettant une meilleure représentation des variations temporelles.
- **Résidus studentisés** : résidus normalisés par la volatilité conditionnelle, garantissant une variance constante.

Les analyses suivantes se concentreront sur les résidus studentisés, qui feront l'objet des tests de validation.

9.8 Test d'Indépendance des Résidus

L'indépendance des résidus est testée à l'aide du test de Ljung-Box. Graphiquement, les résidus studentisés semblent centrés et de variance constante, mais une validation statistique est nécessaire.

Les résultats du test de Ljung-Box appliqué aux résidus sont reportés dans le tableau 8. Les p-valeurs étant toutes supérieures à 0.05, l'hypothèse nulle d'absence d'autocorrélation n'est pas rejetée. Cela suggère que les résidus suivent un comportement de bruit blanc.

Table 8: Résultats du test de Ljung-Box sur les résidus studentisés

Lag	p-value
1	0.3287
2	0.4318
3	0.6013
4	0.6451
5	0.5914
10	0.7827
15	0.5800
20	0.7844

Afin de vérifier que le modèle capture correctement la dynamique de la volatilité, nous appliquons également le test de Ljung-Box aux carrés des résidus. L'absence d'autocorrélation dans les résidus au carré confirmerait une bonne modélisation de l'hétéroscédasticité.

Les résultats du test sont présentés dans le tableau 9. Les p-valeurs étant toutes supérieures à 0.05, il n'y a pas d'autocorrélation significative, indiquant que le modèle capture correctement la structure de la volatilité conditionnelle.

Table 9: Résultats du test de Ljung-Box sur les résidus au carré

Lag	p-value
1	0.3614
2	0.5217
3	0.1493
4	0.2478
5	0.2628
10	0.4977
15	0.5597
20	0.6680

Ces résultats confirment que le modèle AR(1)-GARCH(1,1) ajuste correctement les dynamiques de la série temporelle en éliminant les effets d'autocorrélation dans les résidus et dans leur carré.

Un test de normalité est également nécessaire, car l'estimation des paramètres du modèle de volatilité repose sur l'hypothèse que les résidus, notés ϵ_t , suivent une distribution gaussienne.

Pour rappel, le modèle peut être exprimé par les équations suivantes :

$$r_t = \mu_t + \sigma_t \times \epsilon_t$$

avec :

$$\mu_t = \mu + \phi r_{t-1}$$

et :

$$\sigma_t^2 = \omega + \alpha(r_{t-1} - \mu_{t-1})^2 + \beta\sigma_{t-1}^2$$

L'histogramme des résidus studentisés ainsi qu'un QQ plot sont présentés ci-dessous :

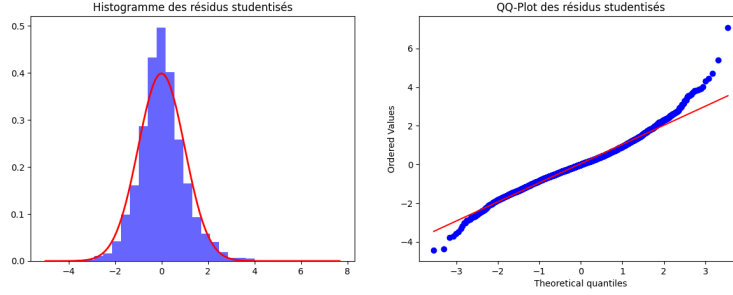


Figure 19: Histogramme des résidus studentisés et QQ plot.

La distribution des résidus suit globalement les quantiles théoriques, à l'exception des valeurs extrêmes. Un test de normalité de Shapiro-Wilk a été effectué, et la p-value étant nulle, l'hypothèse nulle de normalité des résidus est rejetée.

Ainsi, bien que les résidus soient de type bruit blanc, ils ne suivent pas une distribution normale.

9.9 Modélisation de la dynamique des rendements logarithmiques

À partir de notre modèle AR(1)-GARCH(1,1), nous avons estimé les coefficients en utilisant notre base d'entraînement. La connaissance de ces coefficients permet de simuler les trajectoires de μ_t et de σ_t . Nous utilisons pour cela les équations définies précédemment.

La variance à long terme est donnée par :

$$\text{Var}_\infty = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

Les conditions initiales pour la simulation sont choisies comme suit :

$$\mu_0 = \hat{\mu}, \quad \sigma_0 = \sqrt{\text{Var}_\infty}$$

avec μ_0 correspondant à la constante $\hat{\mu}$ du modèle AR(1), et σ_0 étant la racine carrée de la variance à long terme calculée ci-dessus.

Les graphiques ci-dessous montrent les résultats de la modélisation de la moyenne conditionnelle et de la volatilité conditionnelle obtenus à partir de ces conditions initiales.

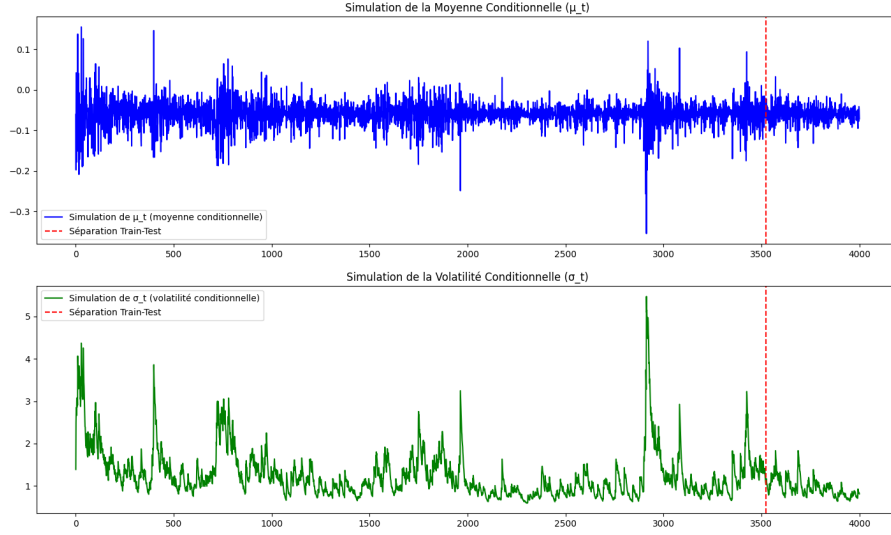


Figure 20: Moyenne conditionnelle et volatilité conditionnelle simulées.

À cette étape, nous avons estimé la moyenne conditionnelle et la volatilité conditionnelle.

À chaque instant t , la Value at Risk peut être estimée à l'aide de la formule suivante :

$$\text{VaR}_t = \hat{\mu}_t + \hat{\sigma}_t \cdot q_{1-\alpha}$$

où :

- $\hat{\mu}_t$ est la moyenne conditionnelle estimée au temps t .
- $\hat{\sigma}_t$ est la volatilité conditionnelle estimée au temps t .
- $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la distribution des rendements (normale, empirique, GEV, etc.).

Cette relation permet d'estimer la perte potentielle maximale avec un niveau de confiance α , en combinant l'information sur la tendance moyenne et la dispersion des rendements.

Nous avons estimé le quantile de trois manières différentes :

- En prenant le quantile d'une loi normale pour la VaR gaussienne.
- En utilisant le quantile historique pour la VaR historique.
- En prenant le quantile de la distribution GEV pour la VaR basée sur l'approche des maximums par bloc.

Les résultats des VaR sur l'échantillon de test sont présentés ci-dessous. La VaR à l'instant $t + 1$ dépend des observations passées et des coefficients estimés par le modèle AR(1)-GARCH(1,1).

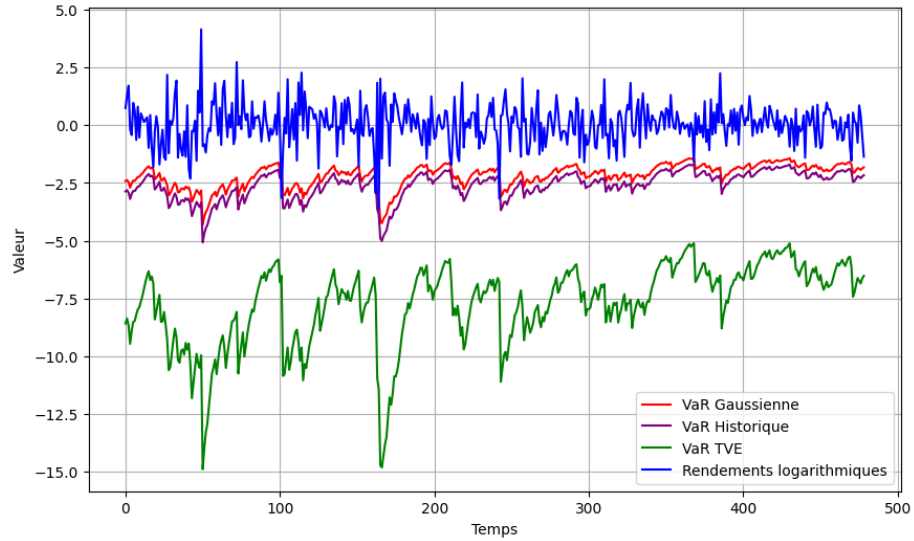


Figure 21: VaR GARCH-Historique, GARCH-Gaussienne et GARCH-TVE (Maximum par Bloc).

Les exceptions observées pour chaque méthode de VaR sont les suivantes :

Méthode	Nombre d'exceptions	Pourcentage d'exceptions
Gaussienne	6	1.25%
Historique	3	0.63%
TVE (MB)	0	0.00%

Table 10: Nombre d'exceptions des différentes méthodes de VaR GARCH.

On constate que la VaR GARCH gaussienne est celle dont le pourcentage d'exceptions observées se rapproche le plus du seuil attendu de 1 %.

Conclusion

Ce projet a permis de comparer plusieurs méthodes de calcul de la VaR, chacune ayant ses avantages et inconvénients. Voici un résumé des principales caractéristiques des différentes approches :

- **VaR historique :**
 - *Avantages* : Simple à mettre en œuvre, ne nécessite pas d'hypothèses sur la distribution des rendements.
 - *Inconvénients* : Ne prend en compte que les données passées, ce qui peut être problématique en cas de changements structurels dans le marché.
- **VaR bootstrap :**
 - *Avantages* : Évite les hypothèses paramétriques sur la distribution des rendements.
 - *Inconvénients* : Peut être sensible à la taille de l'échantillon et aux événements extrêmes qui ne sont pas bien représentés dans les données passées.
- **VaR gaussienne journalière :**
 - *Avantages* : Simple, rapide à calculer, et théoriquement justifiée sous l'hypothèse de rendements normaux.
 - *Inconvénients* : Limité par l'hypothèse de normalité, qui n'est pas toujours valide, surtout en présence de rendements extrêmes ou de queues épaisses.
- **VaR gaussienne scaling 10d :**
 - *Avantages* : Permet d'adapter la VaR à une période plus longue tout en conservant une approche paramétrique simple.
 - *Inconvénients* : L'hypothèse de normalité reste une limitation, et les résultats peuvent être biaisés pour des horizons longs avec des comportements de rendements non gaussiens.
- **VaR Monte Carlo Black-Scholes 10d :**
 - *Avantages* : Très flexible, permet de simuler différentes distributions et de modéliser des scénarios réalistes avec des options complexes.
 - *Inconvénients* : Coûteux en termes de calculs, particulièrement pour des simulations sur de longues périodes et pour un grand nombre de scénarios.

- **VaR EWMA :**

- *Avantages* : Prend en compte les données récentes plus fortement, ce qui peut être pertinent dans des environnements de marché volatile.
- *Inconvénients* : La détermination du paramètre d’amortissement est cruciale et peut affecter significativement les résultats.

- **VaR skew student :**

- *Avantages* : Plus adapté aux rendements qui présentent des asymétries et des queues épaisses, contrairement à la distribution normale.
- *Inconvénients* : Nécessite l’estimation de paramètres supplémentaires et peut être sensible aux erreurs d’estimation pour des échantillons limités.

- **VaR TVE Max par bloc (GEV) :**

- *Avantages* : Permet de modéliser les événements extrêmes de manière robuste et d’obtenir une estimation fiable de la VaR pour des risques rares.
- *Inconvénients* : La méthode peut être sensible au choix des blocs. De plus, elle repose sur des hypothèses asymptotiques.

- **VaR TVE Dépassement seuil (GPD) :**

- *Avantages* : Très efficace pour modéliser les risques extrêmes au-delà d’un certain seuil, avec une estimation paramétrique simple.
- *Inconvénients* : La définition d’un seuil adéquat est complexe et peut avoir un impact important sur les résultats.

- **VaR GARCH :**

- *Avantages* : Prend en compte l’hétéroscédasticité dans les rendements, offrant une meilleure estimation des risques en période de volatilité élevée.
- *Inconvénients* : Plus complexe à mettre en œuvre et à calibrer, et peut être sensible aux spécifications du modèle.

En résumé, les méthodes basées sur les valeurs extrêmes (GEV, GPD) et GARCH se distinguent par leur capacité à traiter les événements rares et la volatilité conditionnelle. Cependant, elles nécessitent une expertise et des choix de paramètres soigneux. Les méthodes plus simples, telles que la VaR historique et la VaR gaussienne, sont rapides à calculer mais peuvent ne pas capturer pleinement les risques extrêmes. En fonction des données et des objectifs, le choix de la méthode doit équilibrer simplicité, robustesse et précision.

Pour résumer les résultats, voici un tableau récapitulatif des différentes valeurs de la VaR calculées au seuil 99% :

Méthode	VaR (99%)
VaR historique	-0.041
VaR bootstrap	-0.041
VaR gaussienne journalière	-0.032
VaR gaussienne 10 jours	-0.102
VaR Monte Carlo 10 jours	-0.097
VaR EWMA, $\lambda = 0.9$	-0.026
VaR EWMA, $\lambda = 0.95$	-0.030
VaR EWMA, $\lambda = 0.99$	-0.033
VaR Skew Student	-0.074
VaR TVE Max par bloc (GEV)	-0.081
VaR TVE Dépassement seuil (GPD)	-0.064

Table 11: Valeurs de la VaR au seuil 99% pour différentes méthodes.