



# MAP 565 MINI-THESIS

April 2024

---

THIAW Mouhamadou Lamine Bara & Yakoub Biram



# CONTENTS

---

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Jeu de Données</b>	<b>4</b>
2.1	Acquisition et Traitement des Données . . . . .	4
<b>3</b>	<b>Analyse ARIMA sur l'Indicateur Alpha</b>	<b>5</b>
3.1	Test de Dickey-Fuller Augmenté . . . . .	5
3.2	Sélection des Paramètres $p$ et $q$ . . . . .	5
3.3	Application du Modèle ARIMA . . . . .	5
3.4	Prédictions et Evaluation . . . . .	6
<b>4</b>	<b>Value at Risk pour l'analyse des extrêmes</b>	<b>12</b>
4.1	Value at Risk: . . . . .	12
4.2	Quantile empirique . . . . .	13
4.3	La méthode du Pick over threshold . . . . .	14

# 1

## INTRODUCTION

---

Dans le domaine de la finance quantitative, la prévision des prix des actifs constitue une composante essentielle pour la prise de décisions d'investissement et de gestion des risques. L'objectif de ce rapport est d'analyser et de prévoir l'évolution des prix de clôture de l'action d'Apple Inc. pour les quatre premiers mois de l'année en cours, en utilisant les données historiques de l'année 2023. Pour atteindre cet objectif, nous appliquerons et comparerons trois méthodes de modélisation différentes : ARIMA, GARCH, et la volatilité rugueuse (Rough Volatility). Chacune de ces méthodes offre une perspective unique sur la modélisation de la volatilité et des tendances des séries temporelles financières, permettant une analyse approfondie et diversifiée.

L'objectif principal de cette étude est de développer un modèle prédictif fiable pour les prix de clôture des actions d'Apple Inc., en se basant sur les données de l'année précédente. Nous visons à déterminer quelle méthode parmi ARIMA, GARCH, et la volatilité rugueuse, utilisée individuellement ou en combinaison, offre la meilleure précision prédictive pour les données à notre disposition. Cette analyse nous permettra non seulement de prévoir l'évolution future des prix de clôture d'Apple mais aussi d'apporter des éclairages sur la dynamique du marché des actions de l'entreprise, contribuant ainsi à une prise de décision éclairée en matière d'investissement et de gestion des risques.

## 2

## JEU DE DONNÉES

## 2.1 ACQUISITION ET TRAITEMENT DES DONNÉES

Les données utilisées pour les premières études proviennent de la bibliothèque ‘yfinance’, qui nous a permis de récupérer les prix d’ouverture et de clôture quotidiens des actions d’Apple Inc. pour l’année 2023. Nous avons ensuite calculé l’indicateur  $\alpha$  pour chaque journée de trading, qui représente la variation relative du prix de l’action durant la journée.

L’indicateur  $\alpha$  est calculé comme suit :

$$\alpha = \frac{\text{Prix de Clôture} - \text{Prix d'Ouverture}}{\text{Prix d'Ouverture}}$$

Cette métrique nous fournit une mesure standardisée de la variation quotidienne du prix des actions, offrant un aperçu de la dynamique intra-journalière du marché.

L’objectif de notre analyse est de prédire les valeurs futures de l’indicateur  $\alpha$ . La prédiction de cet indicateur est cruciale car elle nous permet de déduire le prix de clôture futur à partir du prix d’ouverture, en réarrangeant la formule d’ $\alpha$  :

$$\text{Prix de Clôture} = \text{Prix d'Ouverture} \times (1 + \alpha)$$

Cette approche de prédiction des variations relatives, plutôt que des prix absolus, est préférée car elle nous offre une méthode normalisée et potentiellement plus stable pour analyser et prévoir les mouvements de prix. En se concentrant sur  $\alpha$ , nous réduisons la complexité du modèle en isolant l’effet quotidien des fluctuations de prix, ce qui peut améliorer la précision de nos prévisions.

En résumé, cette méthode de prédiction basée sur  $\alpha$  simplifie notre approche en nous permettant de comprendre et d’anticiper les variations relatives des prix, qui peuvent ensuite être utilisées pour calculer les prix de clôture futurs, à condition que les prix d’ouverture soient connus.

Les données pour l’étude de la VaR peuvent être trouvées sur le lien suivant: **click here**

### 3

## ANALYSE ARIMA SUR L'INDICATEUR ALPHA

Dans cette section, nous détaillons l'application du modèle ARIMA sur l'indicateur Alpha pour les données de l'action d'Apple Inc., en intégrant les étapes mathématiques et les résultats obtenus, y compris les graphiques de la fonction d'autocorrélation (ACF) et de la fonction d'autocorrélation partielle (PACF).

### 3.1 TEST DE DICKEY-FULLER AUGMENTÉ

Pour tester la stationnarité de l'indicateur Alpha, le test de Dickey-Fuller augmenté est utilisé :

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \quad (1)$$

où  $y_t$  est l'indicateur Alpha,  $\Delta$  est l'opérateur de différence, et  $\epsilon_t$  est le terme d'erreur. Le test confirme que l'indicateur Alpha est stationnaire, permettant ainsi l'utilisation du modèle ARIMA.

### 3.2 SÉLECTION DES PARAMÈTRES P ET Q

Les graphiques ACF et PACF sont utilisés pour identifier les ordres  $p$  et  $q$  du modèle ARIMA :

- L'ACF aide à déterminer l'ordre  $q$  du terme MA, indiquant l'ordre auquel les autocorrélations coupent le seuil de signification pour la première fois.
- La PACF identifie l'ordre  $p$  du terme AR, montrant l'ordre auquel les autocorrélations partielles deviennent non significatives.

Ces analyses fournissent les paramètres optimaux pour le modèle ARIMA.

### 3.3 APPLICATION DU MODÈLE ARIMA

Après avoir déterminé les paramètres, le modèle ARIMA est ajusté à l'indicateur Alpha. Les résidus du modèle sont analysés pour assurer l'absence de corrélation et de non-

stationnarité, confirmant la validité du modèle.

### 3.4 PRÉDICTIONS ET EVALUATION

Les prédictions de l'indicateur Alpha sont générées à l'aide du modèle ARIMA, permettant ensuite de déduire les prix de clôture. L'erreur moyenne absolue en pourcentage (MAPE) pour les prix de clôture prédits est calculée, obtenant une valeur de 3.3445, indiquant une précision remarquable du modèle.

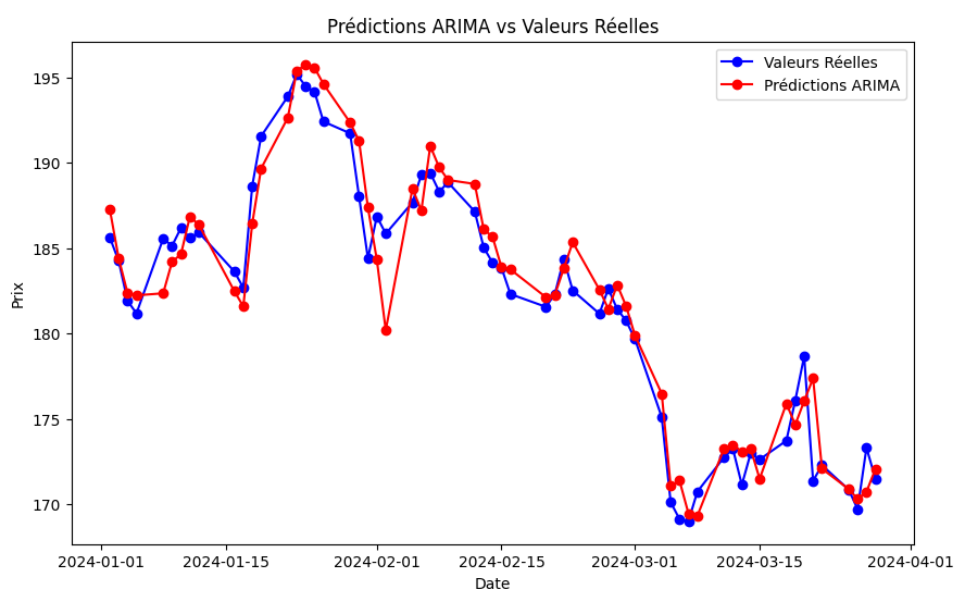
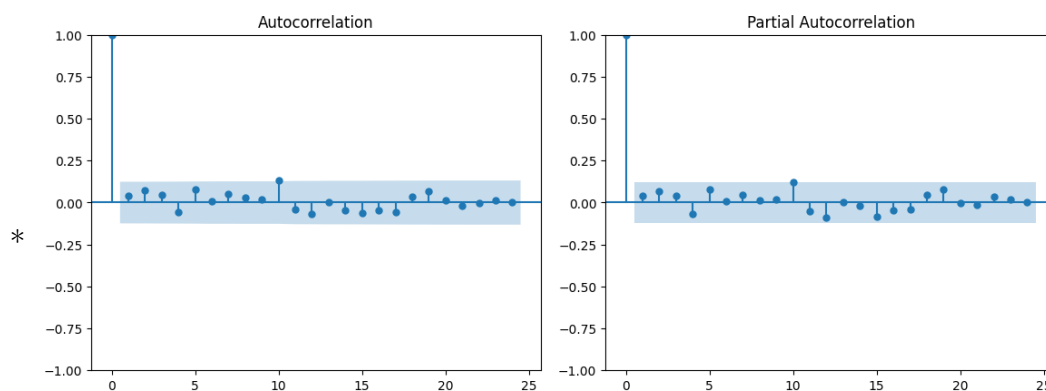


Figure 1: Prédictions ARIMA des prix de clôture .



figureGraphiques ACF pour  
l'indicateur Alpha.

figureGraphiques PACF pour  
l'indicateur Alpha.

# AMÉLIORATION DE LA MODÉLISATION ARIMA AVEC GARCH

## INTRODUCTION AU MODÈLE GARCH

Le modèle GARCH est un outil essentiel en finance quantitative pour analyser et prévoir la volatilité temporelle des rendements d'actifs. En intégrant un modèle GARCH aux résidus d'un modèle ARIMA, nous capturons non seulement la structure autocorrélative des données mais aussi les clusters de volatilité, où les périodes de forte volatilité sont souvent suivies par des périodes de forte volatilité et vice versa.

## APPLICATION DE GARCH SUR LES RÉSIDUS D'ARIMA

Après avoir estimé un modèle ARIMA pour notre série temporelle, nous observons les résidus pour déceler une éventuelle autocorrélation conditionnelle. Les résidus, s'ils ne sont pas complètement aléatoires, peuvent indiquer une volatilité et une structure temporelle que le modèle GARCH peut capturer efficacement.

Le modèle GARCH est défini comme suit :

$$\epsilon_t = \sigma_t z_t, \quad (2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (3)$$

où  $\epsilon_t$  sont les résidus du modèle ARIMA,  $\sigma_t^2$  est la variance conditionnelle (volatilité),  $z_t$  est un terme d'erreur iid normalisé,  $\alpha_0$  est un terme constant, et les  $\alpha_i$  et  $\beta_j$  sont les coefficients qui capturent respectivement l'effet des chocs passés et la persistance de la volatilité.

## OPTIMISATION ET VALIDATION DU MODÈLE ARIMA+GARCH

Après l'ajustement de notre modèle ARIMA, nous avons appliqué un modèle GARCH sur les résidus pour affiner notre analyse de la volatilité. Notre objectif était de trouver les paramètres qui minimisent la fonction de perte, en l'occurrence la log-vraisemblance négative, tout en garantissant la validité statistique de notre modèle.

## • UTILISATION DE LA FONCTION DE LOG-VRAISEMBLANCE

Nous avons utilisé la fonction de log-vraisemblance suivante pour estimer les paramètres de notre modèle GARCH(1,1) :

$$\mathcal{L}(\alpha_0, \alpha_1, \beta_1) = \sum_{t=1}^T \left[ \log \sigma_t^2 + \frac{\epsilon_t^2}{\sigma_t^2} \right], \quad (4)$$

où  $\epsilon_t$  représente les résidus de notre modèle ARIMA et  $\sigma_t^2$  est la variance conditionnelle estimée par le modèle GARCH. Nous avons ajusté les paramètres  $\alpha_0$ ,  $\alpha_1$ , et  $\beta_1$  pour maximiser cette fonction.

## • PROCESSUS DE VALIDATION

Pour valider notre modèle GARCH, nous avons suivi plusieurs étapes cruciales :

1. **Vérification de la Significativité des Coefficients** : Nous nous sommes assurés que les coefficients estimés étaient statistiquement significatifs, ce qui a été vérifié à l'aide des p-valeurs issues des tests de Student.

2. **Analyse des Résidus** : Nous avons analysé les résidus standardisés pour confirmer leur comportement de bruit blanc, s'assurant ainsi qu'ils étaient indépendants et normalement distribués avec une moyenne nulle et une variance unitaire.

3. **Application du Test de Ljung-Box** : Nous avons employé le test de Ljung-Box sur les résidus standardisés et leurs carrés pour détecter toute autocorrélation résiduelle, vérifiant ainsi que notre modèle GARCH capturait adéquatement la volatilité.

$$Q = n(n+2) \sum_{k=1}^h \frac{r_k^2}{n-k}, \quad (5)$$

où  $n$  est la taille de l'échantillon et  $r_k$  représente les autocorrélations aux décalages  $k$ .

## • OPTIMISATION DES PARAMÈTRES

Pour l'optimisation des paramètres, nous avons recours à des méthodes numériques, notamment la maximisation de la vraisemblance, afin de maximiser la fonction de log-vraisemblance mentionnée ci-dessus.

Grâce à ces méthodes, nous avons pu optimiser et valider notre modèle ARIMA+GARCH.



## RÉSULTATS ET VISUALISATIONS

Les graphiques suivants montrent les résultats de l'application du modèle ARIMA+GARCH :

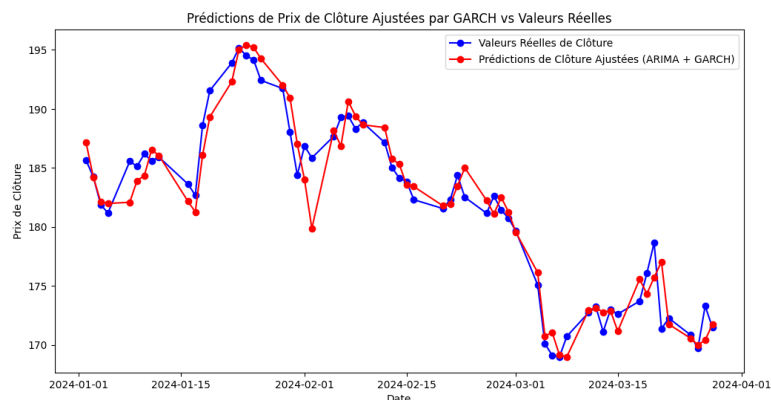


Figure 2: Prédications de Prix de Clôture Ajustées par GARCH vs Valeurs Réelles.

En intégrant GARCH avec ARIMA, nous améliorons notre capacité à modéliser et prévoir la volatilité des séries temporelles financières. L'erreur moyenne absolue en pourcentage (MAPE) pour les prix de clôture prédits est calculée, obtenant une valeur de 3.26

## AMÉLIORATION DES PRÉDICTIONS D'ALPHA EN UTILISANT LA VOLATILITÉ RUGUEUSE SUR LES RÉSIDUS D'ARIMA

Après avoir ajusté un modèle ARIMA pour prédire Alpha directement à partir des données de prix, nous extrayons les résidus du modèle pour capturer les erreurs de prédiction. Ensuite, nous appliquons la volatilité rugueuse pour modéliser ces résidus, permettant ainsi d'améliorer la précision des prédictions d'Alpha en incorporant une meilleure estimation de la volatilité inhérente aux prédictions.

## ESTIMATION DU PARAMÈTRE DE HURST

L'estimation du paramètre de Hurst,  $H$ , est cruciale pour caractériser la nature de la volatilité dans les modèles financiers. Nous appliquons la méthode des ponts discrétisés, une

technique robuste pour analyser la série temporelle et déduire la présence de mémoire à long terme.

## • PONTS DISCRÉTISÉS

Pour une série temporelle  $X_t$  de longueur  $N$ , nous divisons la série en segments de différentes tailles  $n$ . Pour chaque segment, nous transformons les données en un pont discrétisé en soustrayant la tendance linéaire, ce qui nous donne:

$$Y_t = X_t - (a + bt), \quad (6)$$

où  $a$  et  $b$  sont les coefficients de la tendance linéaire du segment. Le pont discrétisé,  $Y_t$ , nous permet d'isoler la structure de dépendance temporelle inhérente à la série.

## • CALCUL DES VARIANCES

Pour chaque taille de segment  $n$ , nous calculons la variance des ponts discrétisés,  $Var(Y_t)$ , pour obtenir une mesure de la variabilité qui n'est pas affectée par la tendance linéaire. Ces variances sont ensuite moyennées sur tous les segments de même taille.

## • RÉGRESSION LINÉAIRE SUR LE GRAPHIQUE LOG-LOG

La relation entre la taille des segments et les variances correspondantes est analysée en utilisant une régression linéaire sur un graphique log-log :

$$\log(Var(Y_t)) = A + 2H \log(n), \quad (7)$$

où  $A$  est une constante, et  $2H$  est la pente de la droite de régression. Le paramètre de Hurst,  $H$ , est alors dérivé comme la moitié de la pente obtenue dans la régression linéaire.

$$H = \frac{\text{slope}}{2}. \quad (8)$$

Cette estimation de  $H$  permet de quantifier la persistance ou l'anti-persistance dans la série temporelle et d'inférer la nature rugueuse de la volatilité.

# PRÉDICTION D'ALPHA AVEC LA VOLATILITÉ RUGUEUSE

Après avoir estimé le paramètre de Hurst,  $H$ , sur les résidus du modèle ARIMA, nous appliquons un modèle de volatilité rugueuse pour simuler des chemins futurs de volatilité des résidus. Ces chemins sont générés en utilisant un mouvement brownien fractionnaire, qui est caractérisé par  $H$  et capture la dynamique de la volatilité des erreurs de prédiction d'Alpha.

## • SIMULATION DES CHEMINS DE VOLATILITÉ DES RÉSIDUS

Soit  $\epsilon_t$  les résidus obtenus de la prédiction d'Alpha par le modèle ARIMA, la volatilité rugueuse des résidus,  $\sigma_{\epsilon,t}$ , est modélisée comme suit :

$$\sigma_{\epsilon,t} = \sigma_{\epsilon,0} \exp \left( -\frac{1}{2} \int_0^t |t-s|^{2H-2} dW_s \right), \quad (9)$$

où  $\sigma_{\epsilon,0}$  est la volatilité initiale des résidus,  $H$  est le paramètre de Hurst estimé, et  $dW_s$  est l'incrément d'un mouvement brownien.

## • AJUSTEMENT DES PRÉDICTIONS D'ALPHA

Nous ajustons les prédictions d'Alpha ( $\hat{\alpha}_t$ ) obtenues par le modèle ARIMA en multipliant ces prédictions par les prédictions de volatilité des résidus :

$$\tilde{\alpha}_t = \hat{\alpha}_t \times \sigma_{\epsilon,t}, \quad (10)$$

où  $\hat{\alpha}_t$  est la prédiction d'Alpha du modèle ARIMA et  $\sigma_{\epsilon,t}$  représente la volatilité estimée des résidus au temps  $t$ . Cette méthode ajuste les prédictions d'Alpha pour refléter l'impact de la volatilité des résidus, fournissant ainsi une estimation plus précise d'Alpha qui tient compte des fluctuations imprévues de la volatilité.

## • ÉVALUATION DES PRÉDICTIONS

Pour évaluer l'efficacité de notre méthode améliorée, nous comparons l'erreur quadratique moyenne (MSE) des prédictions ajustées  $\tilde{\alpha}_t$  avec celles obtenues par le modèle ARIMA seul. Une réduction du MSE indique que l'intégration de la volatilité rugueuse apporte une amélioration significative.

$$MSE = \frac{1}{n} \sum_{t=1}^n (\tilde{\alpha}_t - \alpha_{t,\text{réel}})^2, \quad (11)$$

où  $\alpha_{t,\text{réel}}$  est la valeur réelle d'Alpha au temps  $t$  et  $n$  est le nombre total de points de données.

En intégrant la volatilité rugueuse dans l'analyse des résidus d'ARIMA, nous améliorons la précision des prédictions d'Alpha, offrant une méthode plus raffinée pour anticiper les variations des prix sur les marchés financiers.

Modèle	MSE
ARIMA seul	3.3445
ARIMA + GARCH	3.2663
ARIMA + Volatilité Rugueuse (FBM)	3.2658

Table 1: Comparaison des erreurs quadratiques moyennes pour différents modèles.

## ANALYSE DES RÉSULTATS DE LA VOLATILITÉ RUGUEUSE

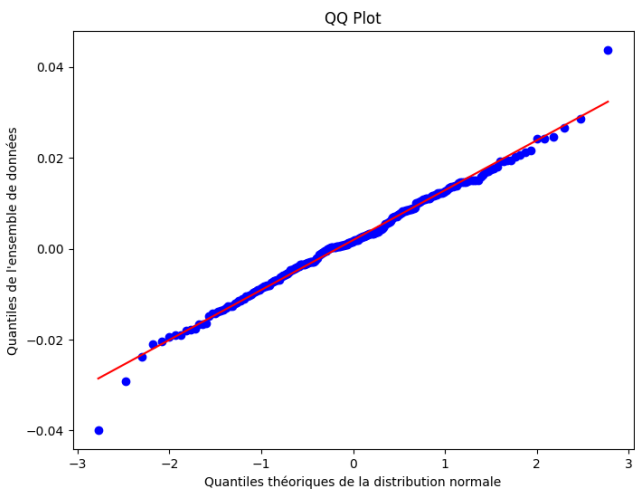


Figure 3: QQ Plot issu de la modélisation.

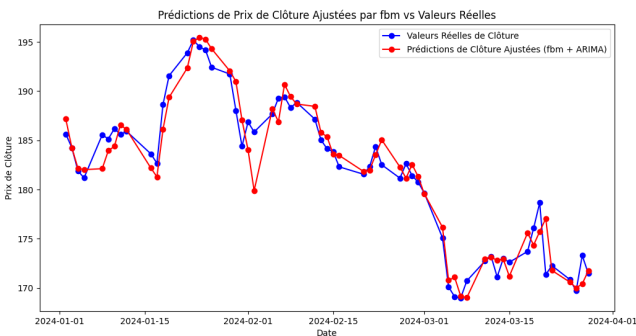


Figure 4: Prédictions de Prix de Clôture Ajustées par GARCH vs Valeurs Réelles.

### 4

## VALUE AT RISK POUR L'ANALYSE DES EXTRÊMES

### 4.1 VALUE AT RISK:

La "Value at Risk" (VaR) est un outil statistique permettant de mesurer le risque d'un actif financier. Il peut être utilisé pour estimer la perte potentielle de valeur d'un portefeuille d'actifs sur une période de temps spécifiée et à un niveau de confiance donné. En substance,

la VaR fournit une estimation de combien un ensemble d'investissements pourrait perdre (ou gagner) sur une période spécifiée, dans des conditions de marché normales, à l'intérieur d'un certain intervalle de confiance. La VaR joue un rôle crucial dans la prise de décision financière en aidant les investisseurs et les institutions financières à comprendre et à atténuer les risques potentiels liés à leurs investissements. Calculer la VaR à un niveau de confiance de 95% revient à trouver le quantile de 95% de la perte quotidienne (daily loss) :  $\frac{P_{t-1}-P_t}{P_{t-1}}$ .

Nous étudierons la VaR d'un portefeuille du CAC40 (FCHI) de janvier 2023 à mars 2024. Au début, nous essayons de comparer la perte quotidienne avec une distribution gaussienne. Voici le résultat.

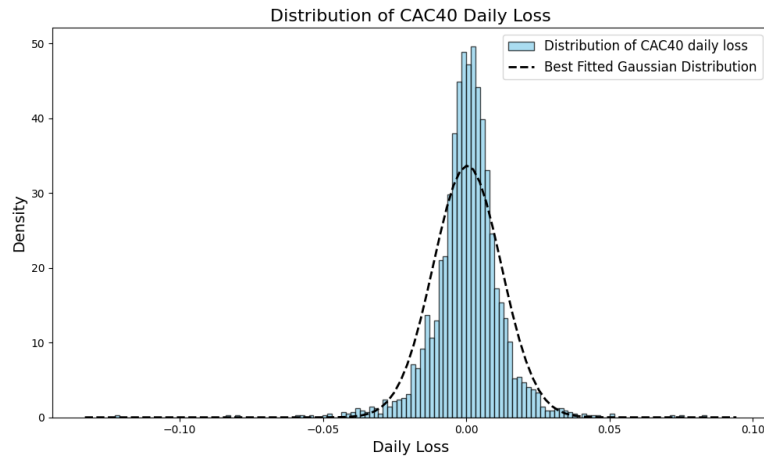


Figure 5: Distribution of CAC40 daily loss

Nous pouvons voir que la perte quotidienne du CAC40 ne suit pas une distribution normale. Pour la suite, nous essayerons d'autres méthodes.

## 4.2 QUANTILE EMPIRIQUE

Pour trouver le quantile de niveau  $\alpha$  de  $(X_1, \dots, X_n)$  en utilisant la formule du quantile empirique, nous calculons  $\hat{q}_\alpha$  donnée par:  $\hat{q}_\alpha = X_{([\alpha n], n)}$  où  $X_{p,n}$  est la  $p$ -ième plus petite donnée parmi  $(X_1, \dots, X_n)$ . Vous pouvez trouver les résultats dans la figure suivante.

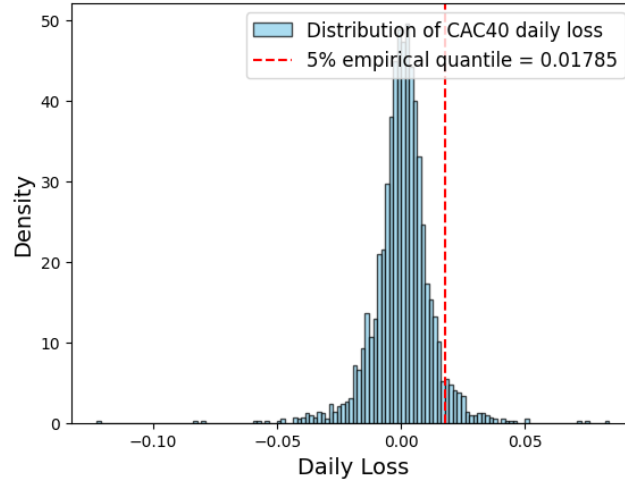


Figure 6: Quantile empirique de niveau 95%

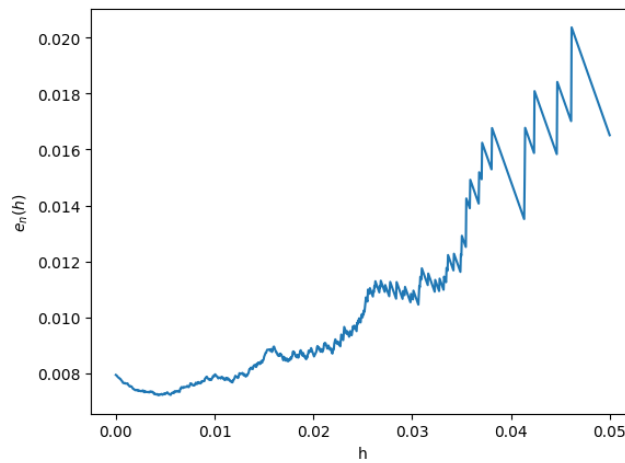
### 4.3 LA MÉTHODE DU PICK OVER THRESHOLD

L'objectif de cette méthode est de trouver le quantile au niveau  $\alpha$  de  $(X_1, \dots, X_n)$ . Pour cette méthode, nous procéderons en 3 étapes.

**Première étape:**

Nous choisissons le seuil  $h$ . Posons:  $N_h = |\{i = 1, \dots, n \mid X_i > h\}|$ .

Soit  $e_n(h) = \frac{1}{N_h} \sum_{i, X_i > h} (X_i - h)$ . Nous avons réalisé un graphique de  $e_n(h)$  et avons finalement choisi un seuil de 0,015 de telle sorte que  $e_n$  soit essentiellement linéaire pour des valeurs supérieures à  $h$ .

Figure 7: Tracé de  $e_n(h)$

### Deuxième étape: Modélisation par la distribution de Pareto :

Modélisons la variable aléatoire  $Z_i = X_i - h | X_i > h$  par la distribution de Pareto généralisée

$$G_{\xi, \beta}(y) = \frac{\xi}{\beta} \left(1 + \frac{\xi y}{\beta}\right)^{-\frac{1}{\xi}-1}.$$

Par maximum de vraisemblance, nous pouvons obtenir les meilleurs paramètres pour notre problème:  $\tilde{\xi} = 1.206e - 01$  and  $\tilde{\beta} = 7.519e - 03$

```
message: CONVERGENCE: REL_REDUCTION_OF_F_<=_FACTR*EPSMCH
success: True
status: 0
  fun: -686.0975401394453
   x: [ 7.519e-03  1.206e-01]
  nit: 8
  jac: [ 7.844e-04  4.547e-05]
 nfev: 36
 njev: 12
hess_inv: <2x2 LbfgsInvHessProduct with dtype=float64>
```

Figure 8: Fitting

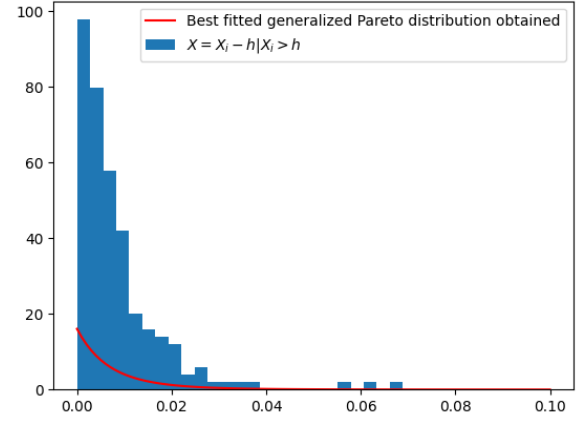


Figure 9: Distribution de Pareto de  $X_i - h | X_i > h$

### Dernière étape: La distribution cumulative

Le théorème de Pickands–Balkema–De Haan donne:  $\hat{F}(x) = 1 - \frac{N_h}{n} \left(1 + \frac{\hat{\xi}(x-h)}{\tilde{\beta}}\right)^{-\frac{1}{\tilde{\xi}}}$ . On peut l'utiliser pour trouver le quantile 95%.

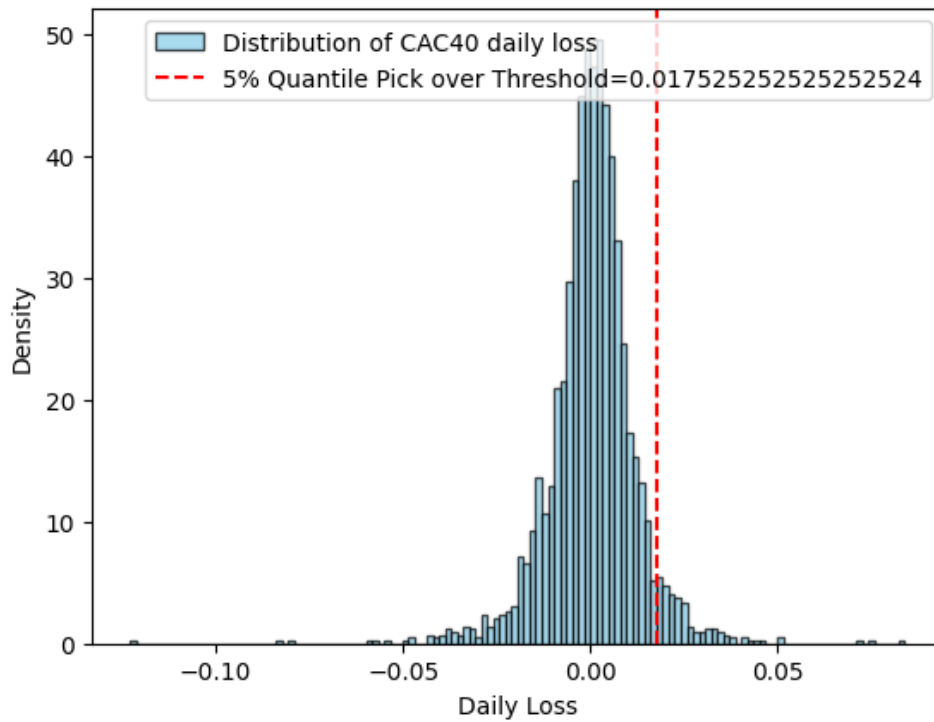


Figure 10: VaR avec la méthode Pick Over Threshold

Nous terminerons en comparant les résultats de ces deux méthodes.

Table 2: Comparaison de la VaR estimée avec un niveau de confiance de 95%.

Méthode	VaR Estimée
Quantile Empirique	0.017854733165897068
Méthode du Pick over threshold	0.017525252525252524