THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Sciences Économique

École doctorale : Ecole Doctorale Economie Gestion (EDEG, ED 231)

Unité de recherche Montpellier Recherche en Economie (MRE, UR 209)

Une contribution sur l'allocation ou la prévision d'actifs d'un portefeuille

Présentée par Frédy Valé Manuel POKOU
le 23 décembre 2022

Sous la direction de François BENHMAD et Jules SADEFO-KAMDEM

Devant le jury composé de

Pascal NGUYEN, Professeur des universités, Université de Montpellier	Examinateur
Bernard KAMSU-FOGUEM, Professeur des universités, Institut National Polytechnique de Toulouse	Examinateur
Fredj JAWADI, Professeur des universités, Université de Lille	Rapporteur
Duc Khuong NGUYEN, Professeur des universités, IPAG Business School	Rapporteur
François BENHMAD, Maitre de conférences HDR, Université de Montpellier	Directeur de thèse
Jules SADEFO-KAMDEM, Professeur des universités, Université de Montpellier	Co-directeur de thèse

UNIVERSITÉ DE MONTPELLIER

Je dédie cette thèse à mes parents : Pokou Konan Léon et N'Dja Nguessan Odette pour leur soutien indéfectible.

Remerciements

Je tiens en premier lieu à exprimer toute ma gratitude à M. François BENHMAD et au Professeur Jules SADEFO-KAMDEM qui ont dirigé mes travaux de thèse, pour leur grande disponibilité de tous les instants et leur soutien au cours de ces quatre années de thèse. Leur exigence et leur ouverture scientifique, leur rigueur et leur enthousiasme très communicatif, leurs critiques et leurs encouragements m'ont immensément aidé à améliorer la qualité de mon travail de recherche.

Mes sincères remerciements au Professeur Pascal NGUYEN qui m'a honoré en acceptant d'être membre du jury. Ma grande gratitude au Professeur Fredj JAWADI et au Professeur Duc Khuong NGUYEN qui m'ont fait le privilège d'accepter d'être les rapporteurs de ma thèse. Mon immense reconnaissance au Professeur Bernard KAMSU-FOGUEM qui a également accepté d'être membre du jury de ma thèse. J'exprime par ailleurs ma reconnaissance au laboratoire d'accueil (MRE), l'école doctorale (EDEG) et la faculté d'économie d'avoir gardé les portes ouvertes pour y trouver un espace dans lequel j'ai pu me sentir à l'aise et être productif, à l'instar de Sabine, s'assurant avec diplomatie du bon déroulement des procédures administratives.

Je souhaite aussi adresser mes chaleureux remerciements à tous les membres de ma famille qui ont été toujours à mes côtés. Il m'est également impossible d'oublier mes amis et estimés collègues pour leurs suggestion, soutien, compréhension et gentillesse durant toute cette période. Que ce modeste travail témoigne de ma reconnaissance et de mes sentiments les plus profonds envers tous ceux et celles qui ont de près ou de loin contribué à sa réalisation.

« L'Université n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse : ces opinions doivent être considérées comme propres à leur auteur. »

Plan de thèse La thèse comporte trois articles écrits en anglais :

1. Empirical assessment of the contribution to the performance of an ESG asset portfolio : Case of American market
2. Empirical assessment of the contribution to the performance of an ESG asset portfolio : Case of Eurozone markets
3. Hybridization of ARIMA with learning models for better forecasting of financial time series

Résumé

Les observations du comportement des investisseurs sur les marchés financiers suggèrent que les investisseurs sont de plus en plus sensibles aux critères extra-financiers dans le processus de construction des portefeuilles. Nous menons donc 2 études empiriques sur l'impact de l'adoption d'une stratégie d'investissement intégrant les enjeux environnementaux, sociaux et de gouvernance dans la sélection d'actifs (ESG) sur la performance des portefeuilles. La première étude nous a permis de comparer la performance de trois portefeuilles composés d'actions du marché américain. La composition de ces portefeuilles a été faite via des filtres ESG. D'une part les best-in-class c'est-à-dire les actions ayant les meilleures notations ESG, puis celles avec les notations moyennes et pour finir les actions les moins bien notées. La mise en œuvre de l'optimisation s'effectue dans un cadre théorique mean-LPM qui met en évidence notre modélisation spécifique de la structure de dépendance du portefeuille. La deuxième étude porte cette fois sur deux portefeuilles constitués d'actions du marché de la zone euro. Un portefeuille avec les best-in-class et le second composé d'actions ayant des notations homogènes. À la différence de la première étude, l'optimisation est faite dans un cadre théorique worst case worst case mixture copula mean-CvaR. De ces études nous sommes arrivés à la conclusion que ce critère n'est pas incompatible avec le profit mais il faudrait réduire ces exigences en terme de préférence pour les actions avec les notes ESG élevées. En vue d'une potentielle extension avec des modèles plus performants, une troisième étude empirique a été réalisée pour mettre en évidence la puissance et la pertinence des modèles d'apprentissage automatique pour la prévision de séries temporelles financières. Ou comme alternative une hybridation entre modèle de réseaux de neurones artificiels et le modèle ARIMA pour prédire la tendance de la série temporelle.

Mots clés : conditional value-at-risk ; esg ; financial time series ; forecasting ; hybrid models ; portfolio selection ; marckov regime switching mixture copula ; robust optimization vine copula

Abstract

Observations of investor behavior in financial markets suggest that investors are increasingly sensitive to non-financial criteria in the portfolio construction process. We therefore conduct two empirical studies on the impact of adopting an investment strategy that integrates environmental, social and governance (ESG) issues into asset selection on portfolio performance. The first study allowed us to compare the performance of three portfolios composed of U.S. market stocks. The composition of these portfolios was made via ESG filters. First, the best-in-class stocks, i.e. those with the highest ESG ratings, then those with average ratings and finally the lowest rated stocks and finally the lowest rated stocks. The implementation of the optimization is carried out in a mean-LPM theoretical framework that highlights our specific modeling of the portfolio dependency structure. The second study focuses on two portfolios of Eurozone stocks. One portfolio with the best-in-class and the second one composed of stocks with homogeneous ratings. Unlike the first study, the optimization is done in a worst case worst case mixture copula mean-CvaR theoretical framework. From these studies we concluded that this criterion is not incompatible with profit, but it would be necessary to reduce these requirements in terms of preference for stocks with high ESG ratings. In view of a potential extension with more efficient models, a third empirical study was conducted to highlight the power and relevance of machine learning models for forecasting financial time series. Or as an alternative a hybridization between artificial neural network model and ARIMA model to predict the trend of the time series.

Keywords : esg ; portfolio selection ; conditional value-at-risk ; marckov regime switching mixture copula ; robust optimization vine copula

Table des matières

Introduction **6**

Contexte . . . 6

Les questions de recherche . . . 8

Structure de la thèse . . . 9

1 Revue de littérature **10**

1.1 État de l'art sur les méthodes d'optimisation de portefeuille . . . 10

2 Évaluation empirique de la contribution à la performance d'un portefeuille d'actifs ESG : Cas du marché américain **17**

2.1 Introduction . . . 17

2.2 Cadre théorique . . . 22

2.2.1 Modélisation univariée . . . 22

2.2.2 Modélisation multi-variée . . . 24

2.2.3 Modèle d'optimisation . . . 29

2.2.4 Intégration des notations ESG d'optimisation d'un portefeuille . . . 32

2.3 Résultats empiriques . . . 33

2.4 Discussions . . . 50

2.5 Conclusion . . . 53

3 Évaluation empirique de la contribution à la performance d'un portefeuille d'actifs ESG : Cas des marchés de la zone Euro **55**

3.1 Introduction . . . 55

3.2 Cadre théorique . . . 60

3.2.1 Minimisation de la CVaR . . . 60

3.2.2 Minimisation de la Worst-Case CVaR . . . 63

3.2.3 Minimisation de la Worst Case Copula CVaR . . . 65

3.3 Applications et résultats . . . 69

3.4 Conclusion . . . 77

4 Hybridation d'ARIMA avec des modèles d'apprentissage pour une meilleure prévision des séries temporelles financières **78**
4.1 Introduction . . . 78
4.2 Théorie . . . 83
4.2.1 Les modèles de prévision traditionnels . . . 83
4.2.2 Les modèles d'apprentissage avancés . . . 85
4.2.3 Les modèles hybrides . . . 96
4.2.4 Les métriques d'évaluation . . . 97
4.3 Résultats empiriques . . . 97
4.4 Discussions . . . 106
4.5 Conclusion . . . 108

Conclusion **112**
Discussion et conclusion . . . 112
Contribution . . . 114
Limites . . . 114
Travaux futurs . . . 114

A Annexe du chapitre 2 **116**
A.1 Model performance . . . 116

B Annexe du chapitre 3 **125**
B.1 Plots of weights . . . 126

C Annexe du chapitre 4 **128**
C.1 Model performance . . . 128

Bibliographie **132**

Table des figures

2.1 Scores ESG selon MSCI . . . 33
2.2 Boxplot de 2 indices de performance des différents des portefeuilles . . . 51

3.1 Les scores ESG selon MSCI . . . 69
3.2 Les différents modèles . . . 75
3.3 Les performances des portefeuilles selon les approch . . . 76

4.1 Réseau neuronal avec un biais en entrée avec $x_0 = +1$ et poids $\omega_{k0} = b_k$ . . . 89
4.2 Réseau de neurone Feedforward . . . 89
4.3 Réseau de neurones récurrents . . . 90
4.4 Deux versions de RNN . . . 92
4.5 Architecture du CNN . . . 94
4.6 Architecture du TCN . . . 95
4.7 Les séries brutes . . . 100
4.8 Les séries des rendements . . . 101
4.9 Les diagrammes quantile-quantile avec une loi normale . . . 102
4.10 Les diagrammes quantile-quantile avec une loi de Student . . . 103
4.11 Les meilleurs modèles pour la prévision des séries SP500 et TTE.PA . . . 110
4.12 Les meilleurs modèles pour la prévision des séries EURUSD et JNUG . . . 111

A.1 Model performances (1) . . . 117
A.2 Model performances (2) . . . 118
A.3 Model performances (3) . . . 119
A.4 Model performances (4) . . . 120
A.5 Model performances (5) . . . 121
A.6 Model performances (6) . . . 122
A.7 Model performances (7) . . . 123
A.8 Model performances (8) . . . 124

B.1 Les pondérations . . . 126

B.2 Les pondérations (suite) 127

Liste des tableaux

2.1 Les copules bivariées utilisées . . . 29
2.2 Les composants du portefeuille 1 . . . 34
2.3 Les composants du portefeuille 2 . . . 35
2.4 Les composants du portefeuille 3 . . . 36
2.5 Statistiques descriptives des actions du portefeuille 1 . . . 37
2.6 Statistiques descriptives des actions du portefeuille 2 . . . 38
2.7 Statistiques descriptives des actions du portefeuille 3 . . . 39
2.8 Les estimations des paramètres GARCH du portefeuille 1 . . . 40
2.9 Les estimations des paramètres GARCH du portefeuille 2 . . . 41
2.10 Les estimations des paramètres GARCH du portefeuille 3 . . . 42
2.11 Les estimations des paramètres GARCH avec changement de régime du portefeuille 1 . . . 43
2.12 Les estimations des paramètres GARCH avec changement de régime du portefeuille 2 . . . 44
2.13 Les estimations des paramètres GARCH avec changement de régime du portefeuille 3 . . . 45
2.14 Les estimations des paramètres du modèle copule-GARCH . . . 46
2.15 Resumé de la sélection de copules bivariées pour le modèle C-vine avec des marginales GJR-GARCH et RS GJR-GARCH . . . 47
2.16 Les performances des différents modèles sur le portefeuille 1 . . . 48
2.17 Les performances des différents modèles sur le portefeuille 2 . . . 48
2.18 Les performances des différents modèles sur le portefeuille 3 . . . 49
2.19 Les performances des différents modèles sur le portefeuille Benchmark . . . 49
2.20 Tests statistique sur 2 indices de performance des différents portefeuilles . . . 51
2.21 Test de comparaison des valeurs des Omega ratio des différents portefeuilles . . . 52
2.22 Test de comparaison des valeurs des Max Drawdown des différents portefeuilles . . . 52

3.1 La composition du portefeuille 1 . . . 70
3.2 La composition du portefeuille 2 . . . 71
3.3 Statistiques descriptives des actions du portefeuille 1 . . . 72
3.4 Statistiques descriptives des actions du portefeuille 2 . . . 73
3.5 Les performances du portefeuille 1 . . . 74

3.6 Les performances du portefeuille 2 74

4.1 Les actifs financiers sélectionnés 98
4.2 Statistiques descriptives des prix 98
4.3 Statistiques descriptives des rendements 99
4.4 test de non linéarité de Terasvirta 104
4.5 Les 10 meilleurs modèles de prédiction sur la période 1 104
4.6 Les 10 meilleurs modèles de prédiction sur la période 2 105
4.7 Les 10 meilleurs modèles de prédiction sur la période 3 105
4.8 Test de Wilcoxon unilatéral des RMSE entre les modèles ARIMA N et ARIMA T 106
4.9 Test de Wilcoxon unilatéral des RMSE entre les modèles hybrides et machine learning 106

C.1 L'ensemble des modèles de prédiction sur la période 1 129
C.2 L'ensemble des modèles de prédiction sur la période 2 130
C.3 L'ensemble des modèles de prédiction sur la période 3 131

Introduction

Contexte

L'observation du comportement des investisseurs sur les marchés financiers montre une sensibilité croissante à des critères extra-financiers dans le processus de construction de leurs portefeuilles. Cette nouvelle tendance dans les choix d'investissement a pour objectif de remettre en cause la primauté des critères financiers qui guidaient de ces choix. Ce qui a constitué un terreau favorable à l'émergence du concept de responsabilité sociale des entreprises. En conséquence, la communauté financière a entamé une réflexion sur l'impact des entreprises qui bénéficient de l'allocation de fonds d'investissement dans le cadre global des Objectifs de Développement Durable. Cet impact des entreprises peut être décomposé en trois dimensions : sociale, environnementale et une dimension de gouvernance. Ce qui a permis l'émergence des critères ESG qui sont de nature extra-financière. Loin d'être une obligation, la RSE est plutôt une incitation. Cette incitation s'est accentuée avec l'avènement de l'Investissement Socialement Responsable (ISR), une forme d'investissement qui sélectionne prioritairement des entreprises soucieuses de respecter les principes du développement durable.

Cela signifie que l'ISR est, pour l'essentiel, une méthode d'investissement qui applique un filtre négatif pour exclure certaines entreprises. Cette philosophie d'investissement n'est pas nouvelle. Selon des auteurs tels que Schueth (2003); Arjaliès (2010); Viviers and Eccles (2012), nous pouvons voir que la stratégie d'investissement ISR est née dans les années 1700 et a traversé plusieurs phases. La première phase concerne les questions morales et religieuses. Par la suite, le concept SRI est entré dans la phase radicale des mouvements de protestation anti-esclavagistes, anti-guerre du Vietnam, anti-apartheid et anti-nucléaire. Puis, à partir des années 1950, sous l'influence de théories des auteurs comme Bowen (1953, 2013), Sethi (1975); Carroll (1979) et Freeman (1984), cette philosophie d'investissement s'est progressivement transformée en une préoccupation éthique. L'intérêt croissant pour l'ISR au cours des 20 dernières années s'explique par plusieurs raisons. La principale raison de cet intérêt réside dans les inquiétudes concernant le comportement environnemental inapproprié de certaines entreprises. En effet, dans notre société de consommation, la quête des entreprises pour augmenter la productivité et les progrès technologiques ont non seulement conduit à la surexploitation des ressources naturelles, mais aussi à une détérioration dramatique de la qualité des sols, de l'eau et de l'air, entraînant des niveaux de pollution sans précédent.

De cette façon, nous avons assisté à une explosion de catastrophes écologiques. On pourrait citer la marée noire qui a profané le golfe du Mexique par British Petroleum (BP) en 2010 (voir Freudenburg and Gramling, 2011; Leifer et al., 2012; Bishop et al., 2017), l'incident nucléaire de Fukushima (voir Holt et al., 2012; Funabashi and Kitazawa, 2012), ou encore des catastrophes environnementales pouvant avoir des répercussions humanitaires (sociales), comme l'explosion d'un port de Beyrouth en 2020 (voir Bauer and Hann, 2010; Gordon, 2013). Ensuite, une raison supplémentaire qui a suscité l'intérêt de l'ISR serait la crainte grandissante de nouvelles réglementations issues des politiques environnementales des différents gouvernements. (Chava, 2014; Hirschl, 2009; Michalena and Hills, 2012). L'une des premières politiques publiques à aborder la question de la durabilité environnementale est la proposition de tarification du carbone de l'Union européenne. Cet instrument politique devrait permettre de lutter contre le changement climatique selon Bohl et al. (2013), Kaplan and Ramanna (2021), Panibratov et al. (2022). Ceci est conforme aux observations de Meinshausen et al. (2009) et Heede and Oreskes (2016), qui proposaient de ne pas utiliser les réserves fossiles inexploitées pour atteindre l'objectif des 2°C, au-delà duquel on ne pourrait plus revenir à une forme de normalité.

Cette politique de tarification du carbone encourage le désinvestissement dans les énergies fossiles et l'investissement dans les énergies renouvelables, comme c'est le cas en Allemagne (Ansar et al., 2013). Cela n'est pas sans conséquence, compte tenu de la forte dépendance de nos économies aux énergies fossiles. Enfin, la dernière raison qui pourrait expliquer l'appétit croissant des investisseurs pour l'investissement responsable serait liée à la tentative de l'industrie de la gestion d'actifs de s'aligner sur les Principes d'investissement responsable (PRI) des Nations Unies. Cette industrie a exprimé une réelle volonté d'intégrer et de promouvoir l'intégration des critères ESG dans ses choix d'allocation de portefeuille sur les marchés financiers mondiaux (Della Croce et al., 2011). L'Investissement Socialement Responsable (ISR), après plusieurs phases d'évolution, a atteint sa maturité et peut être divisé en trois grandes catégories :

- L'Investissement Socialement Responsable (ISR), qui vise à exclure les entreprises qui ne répondent pas aux exigences extra-financières de leur choix d'investissement
- L'investissement d'impact, qui consiste à investir pour générer un impact social et environnemental positif et mesurable avec un rendement financier nul ou positif [1].
- Les investissements environnementaux, sociaux et de gouvernance (ESG), qui consistent à sélectionner les entreprises dans lesquelles nous souhaitons investir en fonction de leurs pratiques environnementales, sociales et de gestion.

Bien que cela représente une opportunité pour les investisseurs, de nombreuses questions subsistent quant à la performance financière de l'ISR. Parmi ces questions, on peut citer :

- L'ISR est-il condamné à être moins performant qu'un investissement traditionnel ?
- L'ISR rapporte-t-il autant que les autres investissements ?
- L'ISR est-il compatible avec le profit ?

Cet intérêt pour les performances des ISR s'explique par la volonté de les promouvoir, de les développer et d'assurer leur légitimité sur les marchés financiers. Une telle légitimité ne pouvait être apportée que par la communauté scientifique à travers la recherche académique sur le sujet. Cette recherche académique sur l'efficacité de l'ISR peut être divisée en 3 courants de pensée.

1. https://thegiin.org/impact-investing/need-to-know/#what-is-impact-investing

Le premier courant de pensée soutient l'opinion selon laquelle l'ISR n'est pas rentable. Ce fut le cas de Friedman (2007), qui soutenait qu'il était impossible d'obtenir des rendements financiers souhaitables en investissant dans ce qu'il considérait comme la "voie directe" des ISR. Cependant, afin d'honorer leurs valeurs éthiques, il est judicieux de privilégier un portefeuille diversifié et d'allouer une partie des rendements excédentaires à des projets qui adhèrent aux valeurs recherchées. C'est ce qu'il appelle "la voie indirecte".

Rudd (1981) et Clow (1999), quant à eux , expliquent la sous-performance de l'ISR par la restriction du nombre de secteurs d'activités dans lesquels on pourrait investir. Cette restriction des choix d'investissement va à l'encontre des hypothèse du modèle de Markowitz (1952a) notamment celles relative à la rationalité de l'investisseur et à la nécessité de diversification pour réduire le risque global du portefeuille.

Contrairement au premier courant de pensée, le deuxième courant de pensée suppose que la performance ISR est neutre, c'est-à-dire que les rendements de l'ISR sont similaires à ceux des autres investissements. La faiblesse de ce courant de pensée est qu'il n'a aucun fondement théorique, car il repose entièrement sur des recherches empiriques comme celles de Abramson and Chung (2000); D'Antonio et al. (2000); Le Maux and Le Saout (2004); Sauer (1997).

Cependant, contrairement aux deux premiers, le troisième courant de pensée est issu des travaux de chercheurs qui soutiennent la thèse selon laquelle les performances financières des lSR surpassent celles des investissements traditionnels. Ainsi, Porter (1991) a introduit une théorie selon laquelle l'amélioration des performances environnementales d'une entreprise augmenterait sa productivité, permettant ainsi l'innovation et la différenciation par rapport à la concurrence. Une deuxième théorie soutenant la performance supérieure des ISR est celle de Mallin et al. (1995) soutiennent que la demande croissante d'actifs financiers conformes à l'ISR fera grimper les prix de ces actifs. Kurtz (2008) soutient la théorie selon laquelle intégrer les entreprises les mieux notées (ESG) dans la composition d'un portefeuille réduit et anticipe les risques extra-financiers car ces entreprises plus stables génèrent des revenus à long terme sur le marché.

Il est à noter que l'intégration des critères ESG peut se faire selon 2 approches bien connues dans la littérature existante : approche de Markowitz revisitée (Gasser et al., 2017; Schmidt, 2022; Escobar-Anel, 2022) et l'approche multi-objectifs ou Multicriteria Decision Making (Ballestero et al., 2012; Dobrovolskienė and Tamošiūnienė, 2016; Garc1́a et al., 2019; Hirschberger et al., 2013; Utz et al., 2014, 2015).

Les questions de recherche

Notre article vise à analyser l'impact de l'intégration des enjeux environnementaux, sociaux et de gouvernance (ESG) sur l'allocation et la gestion des portefeuilles.

Pour atteindre cet objectif, cette thèse tente de répondre aux questions de recherche suivantes :

1. Adopter une stratégie d'investissement intégrant des critères ESG affecte-t-il la performance du portefeuille ?
2. Les portefeuilles construits à l'aide de la stratégie best-in-class sont-ils plus performants que les portefeuilles traditionnels ?

3. Les portefeuilles construits à l'aide de la stratégie best-in-class présentent-ils un risque de perte plus élevé que les autres stratégies ?
4. Les portefeuilles construits avec la stratégie best-in-class offrent-ils le meilleur compromis risque-rendement ?
5. Le choix du modèle de structure de dépendance du portefeuille affecte-t-il la performance ?
6. Les portefeuilles constitués d'actifs américains et d'actifs européens ont-ils des performances similaires, sachant que les deux marchés ont des caractéristiques différentes ?

Structure de la thèse

Cette thèse est composée de 4 chapitres et structurée de la manière suivante.

- Le chapitre 1 passe en revue la littérature pertinente sur l'optimisation de portefeuille basée sur différentes mesures de risque et nécessaire pour les chapitres suivants.
- Le chapitre 2 est notre premier essai, explorant l'intégration des critères ESG pour des portefeuilles composés d'actifs du marché américain. Une étude empirique utilisant la vine copula et Markov Switching dans le cadre mean-LPM.
- Le chapitre 3 est notre deuxième essai, examinant l'intégration des critères ESG pour des portefeuilles constitués d'actifs des marchés de la zone euro. Une étude empirique menée dans le cadre worst case mixture copula mean-CVaR.
- Le chapitre 4 est le troisième dans lequel nous étudions la pertinence des modèles de machine learning pour la prévision de séries temporelles financières. Ce chapitre vise à établir un lien entre nos travaux de recherche et des pistes d'extension possibles au-delà des contraintes rencontrées.

On termine par une conclusion générale qui synthétise nos principaux résultats empiriques et présente les propositions de recherche future.

CHAPITRE 1

Revue de littérature

1.1 État de l'art sur les méthodes d'optimisation de portefeuille

L'intérêt des chercheurs et des praticiens financiers pour les méthodes efficaces d'investissement des actifs financiers remonte aux années 1950. L'une des méthodes les plus connues est celle développée par Markowitz (1952b,c). En effet, Markowitz (1952b) a posé les bases de la théorie moderne du portefeuille en se basant sur certaines hypothèses, telles que l'hypothèse de rationalité de l'investisseur et l'hypothèse d'une distribution normale des rendements des actifs financiers. La formulation mathématique de l'approche de Markowitz (1952b) suggère une optimisation basée sur deux critères, qui sont la moyenne et la variance. Cette optimisation de la variance moyenne repose sur la prise en compte du caractère multi-objectif essentiel entre deux objectifs potentiellement conflictuels lors de la sélection d'un portefeuille. Selon l'hypothèse de rationalité de l'investisseur, deux formulations différentes peuvent être formulées. Ainsi, parmi des portefeuilles à niveau de risque identique, un investisseur rationnel choisira celui dont le rendement attendu (moyen) est le plus élevé. A l'inverse, cet investisseur choisira le portefeuille le moins risqué parmi des portefeuilles aux rendements attendus identiques. La deuxième hypothèse, quant à elle, résulte de la recommandation de Markowitz (1952b), qui a suggéré que les deux premiers moments de la distribution de probabilité des rendements sont largement suffisants pour optimiser le portefeuille. C'est l'hypothèse de normalité. Cette hypothèse, rarement satisfaite, semble irréaliste.

Pour cette raison, le modèle de Markowitz (1952b) n'a pas rencontré de critiques dans la communauté scientifique, notamment au niveau de ses hypothèses, qui génèrent une incertitude quant à l'estimation des paramètres d'entrée. C'est pourquoi David Li dans l'article "*How a formula ignited market that burned some big investors*" de Whitehouse (2005) met en garde contre le risque d'utiliser le modèle de Markowitz (1952b) sans tenir compte des hypothèses sous-jacentes. Comme Fabozzi, Focardi, Kolm, and Pachamanova (2007), qui s'opposent à l'hypothèse de normalité en se basant sur leur connaissance des spécificités des séries temporelles financières. Le non-respect de l'hypothèse de normalité de la distribution des bénéfices a conduit Michaud (1989) à qualifier ce modèle d'optimisation d'erroné.

En effet, en supposant la normalité, les estimateurs des rendements attendus correspondants et la matrice de covariance sont les estimateurs du maximum de vraisemblance. Bien qu'ils aient des propriétés intéressantes, ces estimateurs peuvent être biaisés pour des échantillons finis. Chopra and Ziemba (2013) ont réussi à mettre en évidence l'impact négatif des erreurs d'estimation sur la performance du portefeuille. Une interprétation de leurs résultats est que l'information obtenue à partir des données historiques n'est pas suffisante.

Cela conduit à des allocations sous-optimales avec une concentration excessive sur quelques actifs, ce qui compromet les avantages de la diversification. Il est important de faire la distinction entre les erreurs d'estimation liées aux rendements attendus et les erreurs liées à la matrice de covariance, car elles ne reçoivent pas la même attention dans la littérature. L'une des formulations les plus utilisées du modèle de Markowitz (1952b) est la variance minimale globale (GMV). L'idée est de minimiser la variance du portefeuille tout en limitant le rendement cible. Il formule mathématiquement le concept de diversification de la sélection d'actifs afin que le risque du portefeuille soit inférieur à la somme des risques individuels. Cette formulation du GMV montre le rôle central joué par la matrice de covariance, ce qui peut aussi expliquer l'intérêt de la communauté scientifique à proposer des méthodes appropriées pour l'estimer par rapport à l'espérance de rendements.

Comme mentionné précédemment, le chargement de l'estimateur du maximum de vraisemblance sous l'hypothèse de normalité a été expliqué par Bengtsson and Holst (2002), Ledoit and Wolf (2003) par la nécessité de disposer de suffisamment de données. Cependant, le fait de disposer de données historiques sur une trop longue série chronologique comporte des risques qui contiennent beaucoup de bruit et sont même trop anciens pour saisir les conditions actuelles du marché. La solution immédiate à cette affirmation consiste à utiliser une moyenne mobile pondérée exponentielle (Exponential Weighting Moving Average) pour calculer la matrice variance-covariance. Pour relever ce défi, de nombreuses méthodes plus sophistiquées ont été introduites, comme le shrinkage proposée par Ledoit and Wolf (2017) et d'autres méthodes qui ont déjà fait leurs preuves dans des domaines éloignés de la finance, comme la théorie des matrices aléatoires empruntée à la physique et développé par Wigner (1958). La théorie des matrices aléatoires fournit un meilleur estimateur de matrice de variance-covariance qui réduit le bruit et capture les caractéristiques hétéroscédastiques des séries inverses. Ces améliorations ont été bien appréciées par les chercheurs en finance à l'instar de Laloux et al. (2000), Plerou et al. (2000), Pafka et al. (2004), Bai et al. (2009).

Alors que certains modèles proviennent d'autres disciplines, les chercheurs en modélisation financière ne sont pas en marge de ces avancées scientifiques. Ils ont le mérite de proposer des modèles assez cohérents avec les caractéristiques empiriques des données financières. Ainsi, en parcourant la littérature financière sur les méthodes empiriques, on peut trouver des méthodes qui pourraient résoudre des problèmes liés à l'estimation du vecteur des moyennes ou l'espérance de rendements. On peut citer EWMA (Exponential Weighted Moving Average)(Longerstaey and Spencer, 1996), Bayes-SteinJorion (1986), la méthode de Martinelli(Martellini, 2009) et bien d'autres. Cependant, l'un des concepts les plus importants dans le cadre de l'optimisation de portefeuille est la structure de dépendance pour modéliser les interdépendances entre les actifs.

Dans la théorie de Markowitz (1952b), l'hypothèse de normalité réduit la structure de dépendance à une matrice de covariance. Ainsi, pour réduire la variance et donc le risque d'un portefeuille d'actifs financiers, la sélection de ces actifs doit se faire en tirant parti des corrélations linéaires négatives pouvant exister entre ces actifs. C'est pourquoi identifier, mesurer et modéliser correctement la structure de dépendance est un enjeu central dans l'analyse de la contagion financière, la gestion des risques et la diversification des portefeuilles. Afin de modéliser la nature complexe de la structure de dépendance, les premières conceptions de solutions reposaient sur une classe de modèles multidimensionnels GARCH (MGARCH) qui capturent la dynamique de la variance et de la covariance dans le temps. Ainsi, l'extension de la modélisation, qui était réservée aux variables unidimensionnelles, à un modèle multidimensionnel semble être une voie très prometteuse.

De tels modèles multidimensionnels doivent non seulement être parcimonieux, mais aussi avoir une flexibilité suffisante. La première tentative de modèle GARCH multidimensionnel a été le travail de Bollerslev (1990), qui a introduit le modèle CCC-GARCH ou Constant Conditional Correlation-GARCH. La méthode du modèle CCC-GARCH consiste à décomposer la matrice de covariance conditionnelle en écart-type conditionnel et en matrice de corrélation conditionnelle. Longin and Solnik (1995, 2001), Ang and Bekaert (2002); Ang and Chen (2002) a également pu fournir des résultats empiriques qui mettent l'accent sur la nature dynamique des corrélations dans le temps, entraînant une augmentation des corrélations entre les actifs financiers pendant les périodes de baisse des marchés. Cette dernière preuve empirique invalide l'hypothèse de la corrélation conditionnelle constante du modèle CCC-GARCH précédent. En conséquence, Engle III and Sheppard (2001) ont introduit le modèle DCC-GARCH (Dynamic Conditional Correlation-GARCH), qui propose une matrice de corrélation conditionnelle qui évolue dans le temps, prenant ainsi en compte la dynamique nature de la corrélation. Le modèle DCC-GARCH a été affiné par Cappiello, Engle, and Sheppard (2006), qui ont introduit le modèle A-DCC-GARCH pour tenir compte de la nature asymétrique des corrélations conditionnelles. Cependant, l'analyse de la corrélation entre les actifs du portefeuille permet de se focaliser sur l'interdépendance entre ces actifs.

Cependant, limiter la mesure des interdépendances et des co-mouvements entre ces actifs au coefficient de corrélation de Pearson peut être très réducteur et entraîner une allocation de portefeuille sous-optimale. Si l'on tient compte du fait que le coefficient de corrélation de Pearson est basé sur la distribution gaussienne des variables, alors on a affaire à l'hypothèse de l'impossibilité de satisfaire la normalité (Mandelbrot et al., 1963; Fama, 1965; Embrechts et al., 1999) ont tenté de dépasser l'hypothèse de normalité en la généralisant aux distributions elliptiques. Toutefois, pour une modélisation correcte de la structure de dépendance, le modèle idéal doit répondre à certaines propriétés. Tout d'abord, ce modèle doit pouvoir prendre en compte l'asymétrie, l'épaisseur des queues de distribution, le regroupement de la volatilité individuellement dans chaque série temporelle de rendements des actifs financiers inclus dans le portefeuille. Ensuite, ce modèle doit avoir une structure de dépendance capable de capturer au-delà de la dépendance linéaire, la dépendance asymétrique et la dépendance aux événements extrêmes aux extrémités des distributions. Ces propriétés peuvent être satisfaites par les dômes introduits par Sklar (1959). La copula, qui est un outil statistique qui prend en compte plusieurs propriétés théoriques souhaitables, il peut donc être un bon modèle de dépendance. Elle a l'avantage d'être flexible car elle donne la possibilité de modéliser une distribution multivariée en deux étapes indépendantes.

Dans la première étape, la copule a modélisé les distributions marginales et dans la seconde la structure des dépendances. Parmi les différentes classes de copules, les copules elliptiques sont les plus couramment utilisés, le premier étant le dôme gaussien, qui n'est rien d'autre qu'une distribution gaussienne multidimensionnelle. La copule gaussienne est très utile pour comprendre les problèmes à grande échelle, comme lorsque les portefeuilles sont constitués d'un grand nombre d'actifs. Mais, il conserve la structure de covariance, qui est le concept de dépendance symétrique. Ce qui lui a valu de vives critiques, également motivées par l'absence de dépendance dans les queues, qui signifie que les événements extrêmes sont asymptotiquement indépendants. Cette inefficacité de la copule gaussienne a été mise en évidence lors de la crise financière de 2007-2008 par des auteurs tels que MacKenzie and Spears (2014), et surtout Felix Salmon dans son article "*The Formula that Killed Wall Street*" (Salmon, 2012).

Cependant, les meilleurs résultats de la distribution univariée de student suggèrent que sa généralisation multivariée serait un meilleur candidat. Cette distribution multidimensionnelle correspond à une copule de student. Cela permet de réfuter l'hypothèse de l'absence d'occurrence simultanée d'événements extrêmes. Cela dépasse l'hypothèse de normalité multidimensionnelle suggérée dans la littérature financière par Frey and McNeil (2001), Schönbucher (2002). D'après Fabozzi et al. (2005), Mittnik et al. (2007), compte tenu des propriétés que devrait avoir un très bon modèle, il est évident que la part des copules elliptiques sont limitées. Une caractéristique précieuse de la copule serait la capacité de modéliser conjointement l'occurrence d'événements extrêmes négatifs et positifs. Cela correspond à l'opinion selon laquelle les actifs financiers sont plus susceptibles se crasher ensemble que de prospérer ensemble. Les copules elliptiques n'ont pas cette caractéristique et ne peuvent donc pas modéliser une dépendance asymétrique. Ainsi, sortir du cadre des distributions elliptiques revient à abandonner la dépendance linéaire mesurer par le coefficient de corrélation de Pearson au profit du tau de Kendall (τ) ou le rho de Spearman (ρ). Elles permettent de prendre en compte la non linéarité et la propriété d'invariance aux transformations non linéaires strictement croissantes (Chen, Smithson, Popovich, ter Y, et al., 2002). La deuxième classe de copules répond au besoin de modèles capables de capturer des structures de dépendance de queue asymétriques, telles que les copules de Clayton, Gumbel et Joe.

Introduit par Clayton (1978), la copule de Clayton est devenue très populaire dans la littérature sur les risques financiers en raison de sa capacité à modéliser le regroupement des risques et donc les dépendances dans la queue inférieure. Quant aux copules de Gumbel(Gumbel, 1960) et Joe(Joe, 1997), elles permettent de modéliser une distribution multivariée en attribuant un poids de probabilité plus important aux événements extrêmes simultanés positifs. De plus, il convient de noter que dans la famille des copules archimédiennes, il existe des copules qui, comme la copule gaussienne, sont symétriques et indépendants de la queue, mais sont capables de capturer des relations non linéaires au centre des distributions, telles que la copule de Frank.
La copule Joe-Clayton, capable de capturer une dépendance asymétrique avec une dépendance dans la queue inférieure plus accentuée que dans la queue supérieure, serait une copule idéale mais elle ne peut pas être implémentée pour des dimensions au delà de 2 en raison de difficultés de calcul. La copule Joe-Clayton est capable de capturer une dépendance asymétrique avec une dépendance dans la queue inférieure plus accentuée que la queue supérieure., serait une copule idéale. Mais elle ne peut pas être mise en œuvre pour des dimensions supérieures à 2 en raison de difficultés de calcul.

Malgré les avantages offerts par les copules archimédiennes, la flexibilité de la structure de dépendance est limitée car ils n'ont qu'un seul paramètre pour contrôler les dépendances de queue entre toutes les paires de variables. Le choix de copules de dimension supérieure ($n > 2$) est donc assez limité, et la construction de copules de cette dimension semble être une tâche difficile. Pour surmonter les limites des copules archimédiennes, la classe des copules de vines et les copules mixtes ont été introduites. Ainsi, Joe (1996), à travers ses travaux sur la construction de copules multidimensionnels, est considéré comme un précurseur de l'idée de construire des paires de copules (Pair Copula Construction), principe de qui consiste à décomposer des densités multidimensionnelles en une cascade de paires de copules. Cette décomposition est une manière plus flexible et intuitive qui permet d'étendre les copules bivariées à des dimensions supérieures. La recherche a été poursuivie par Bedford and Cooke (2001, 2002,?), Kurowicka and Cooke (2006) et Aas et al. (2009) afin d'établir une base probabiliste et inférentielle solide pour obtenir une forme analytique de la distribution de densité multidimensionnelle. Cette augmentation de la flexibilité n'est pas sans inconvénients, principalement dus au nombre [1] important d'éventuelles structures de vine différentes lorsque la dimension de n variables est élevée. Aas et al. (2009) ont ainsi pu isoler des cas particuliers, en l'occurrence les Canonical vines et les Drawable vines deux cas particuliers des Regular vines avec un nombre beaucoup plus restreint de structures différentes possibles.

Il a fallu attendre les travaux de Dißmann (2010); Dissmann et al. (2013), Czado et al. (2013) pour fournir des algorithmes de sélection séquentielle et d'estimation de la structure optimale de la vine ainsi que de ses copules bivariées puisqu'il n'y avait aucune procédure claire pour le faire. Des progrès qui ont motivé de nombreux chercheurs sur la dépendance entre les actifs financiers à travers divers aspects tels que l'allocation d'actifs, la diversification de portefeuille, la gestion des risques, le risque de crédit, la contagion financière, et récemment les questions ESG, sans épuiser le sujet, nous pouvons citer par exemple Schirmacher and Schirmacher (2008), Heinen and Valdesogo (2008), Berg and Aas (2009), Chollete et al. (2009), Min and Czado (2010), Min and Czado (2011), Hernandez (2015), Arreola Hernandez et al. (2017), Mejdoub and Mounira (2017), Low (2018), BenSaïda (2018), Bax et al. (2021), Zhang and Zhang (2021), Muteba Mwamba and Mwambi (2021), Rikhotso and Simo-Kengne (2022).

Une alternative aux copules de vines est les copules mixtes (mixture copula) proposées par Schwarz (1978) et Vuong (1989). Les copules mixtes permettent de modéliser la structure de dépendance de façon plus intuitive que sur le critère d'information bayésien (BIC)(Chollete et al., 2005) ou le test de Vuong(Hu, 2006). Ce sont des combinaisons linéaires de copules multivariées qui permettent le partitionnement de la dépendance dans la queue inférieure, la queue supérieure et même le centre de la distribution à l'aide de copules. Hu (2006), Trivedi et al. (2007) et Kakouris and Rustem (2014) ont proposé une copule archimédienne mixte avec la copule de Gumbel et sa copule de survie ou Clayton, Frank et Gumbel. Malgré toutes ces innovations empiriques, ces modèles de copules restent statiques. Ainsi, l'observation de Longin and Solnik (1995) qui permet de rendre compte des corrélations dynamiques dans les modèles GARCH multivariés reste valable et soutient l'utilisation de copules pouvant varier dans le temps. Ces copules dynamiques peuvent offrir la possibilité de modéliser correctement des dépendances entre actifs financiers qui ne sont pas constantes dans le temps.

1. $C_n^2 \times (n-2)\,! \times 2^{C_{n-2}^2}$ Morales-Napoles (2010)

Ces progrès sont principalement dus aux travaux de Patton (2006) sur l'extension du théorème de Sklar et l'introduction de fonctions de copules conditionnelles pour modéliser les dépendances conditionnelles variant dans le temps. Bien qu'utilisant un modèle GARCH multivarié (MGARCH) pour étudier le degré de corrélation en fonction des tendances du marché (voir Grundke and Polle, 2012; Adams et al., 2017) et les modèles de copule pour capturer les dépendances non linéaires, nous ne parvenons toujours pas à capturer la dynamique complexe qui suit les rendements financiers. Afin d'y remédier, Lee and Long (2009), et So and Yeung (2014) ont poussé la réflexion plus loin en construisant des modèles MGARCH basés sur des copules avec des erreurs dépendantes non corrélées appelées C-MGARCH. Dans le même esprit mais avec une approche différente, Jondeau and Rockinger (2006a); Ausin and Lopes (2010) ont proposé le modèle copule-GARCH dans lequel les marginales sont estimées à partir de modèles GARCH univariés et les résidus standardisés uniformes avec une fonction copule. Babaei et al. (2015) et Yan et al. (2020) soutient que le modèle C-MGARCH est particulièrement efficace pour résoudre les problèmes d'allocation de portefeuille.

Il existe d'autres esquisses de solutions rendues possibles grâce à des modèles déjà bien connus dans l'environnement de la modélisation financière. Ainsi, Hafner and Reznikova (2010), Liu, Long, Zhang, and Li (2019), Yang, Cai, Hafner, Liu, et al. (2021) combine des copules pour proposer des modèles de copules mixtes variant dans le temps (time-varying mixture copula), où les poids et les paramètres dépendants sont des fonctions déterministes et non spécifiées du temps. Nous avons aussi la possibilité de combiner les modèles de dépendance multivariée et les modèles de Markov-Switching(MS), un réel défi sachant que ces modèles à changement de régime ne sont pas faciles à étendre pour des dimensions élevées. Mais certains chercheurs comme Chollete, Heinen, and Valdesogo (2009), Almeida and Czado (2012), Stöber and Czado (2014), Fink et al. (2017) sont parvenus en adoptant une procédure approximative d'Expectation-Maximization (EM) dans un cadre de maximum de vraisemblance (ML) ou de Monte-Carlo par chaînes de Markov [2].

Une revue de la littérature empirique sur les méthodes visant à bien modéliser les structures de dépendance pour minimiser la variance du portefeuille nous incite à les évaluer de manière critique. De notre point de vue, le choix de la variance pose un vrai problème de modélisation. Par définition, la variance mesure la variation autour de la moyenne, il n'y a donc pas de distinction entre la variation au-dessus et en dessous de la moyenne. D'un point de vue pratique, ce problème peut s'expliquer par la sensibilité différente des investisseurs aux changements à la baisse et à la hausse. C'est pourquoi Markowitz (1959) soutient lui-même que la semi-variance peut être une mesure plus appropriée du risque pour l'optimisation de portefeuille.
Les modifications apportées à la mesure du risque permettront en outre d'assouplir l'hypothèse de distribution normale requise lors de l'utilisation de la variance comme mesure du risque. On parle de mesure de risque de baisse (downside risk) et il existe de nombreuses autres mesures de risque qui vont dans ce sens. Nous parlons de mesures de risque de baisse (downside risk), et il existe de nombreuses autres mesures de risque qui vont dans cette direction. Parmi lesquelles, nous avons la Value-at-Risk (VaR)(Jorion, 2000), la Conditional Value-at-Risk (CvaR)(Rockafellar et al., 2000), (Rockafellar and Uryasev, 2002), (Tasche, 2002) et le Lower Partial Moments (LPM) (Bawa, 1975; Fishburn, 1977; Bawa and Lindenberg, 1977). On distingue deux types de mesures de risque, celles basées sur des moments partiels inférieurs (LPM) et celles utilisant des quantiles (VaR et CVaR).

2. Markov Chain Monte Carlo (MCMC)

La première catégorie est conforme à la proposition de Markowitz, la semivariance. Le choix des moments partiels inférieurs (LPM) comme mesure de risque assouplit l'hypothèse de distribution normale requise. Elles ont été introduites pour la première fois dans la théorie du portefeuille par Bawa (1975), Fishburn (1977) et Nawrocki (1991, 1992). L'avantage qu'ils offrent est qu'ils nous libèrent des contraintes d'utilisation de la variance, ce qui implique une fonction d'utilité quadratique, il n'est donc pas possible de prendre en compte les 3ème et 4ème moments (skewness et kurtosis). On peut dénombrer plusieurs auteurs qui utilisent ces mesures de risque dans leurs travaux (Chow and Denning, 1994; Foo and Eng, 2000; Cumova and Nawrocki, 2011; Sadefo Kamdem et al., 2011) et, de plus, ce sont des mesures de risque cohérentes au sens de Artzner et al. (1999). La deuxième catégorie utilise des distributions de probabilité. Cependant, dans la mesure où nous ne disposons pas d'informations précises sur la distribution du vecteur de récompense, nous devons recourir à des méthodes d'optimisation robustes.

On parle de problèmes d'optimisation sous incertitude. Au lieu de supposer une distribution sous-jacente pour le vecteur de retour, nous supposons qu'il appartient à une famille de distributions appelées ensembles d'incertitude. Ce faisant, nous obtenons des portefeuilles optimaux avec des poids stables et très sensibles aux paramètres d'entrée. L'optimisation robuste est à la hausse alors que la communauté scientifique commence à s'intéresser vivement à l'optimisation robuste. On peut le voir à travers les travaux de Zhu and Fukushima (2009) sur la CVaR du pire (worst-case ou WCVaR). Leur approche robuste est conçue pour permettre aux investisseurs volontaires de se couvrir contre le pire scénario possible en utilisant un ensemble d'incertitude ellipsoïdale. Kakouris and Rustem (2014) ont proposé une extension du cadre worst-case CVaR (WCVaR) grâce à l'utilisation des copules mixtes (multidimensional mixture Archimedean copulas). Il existe très peu d'applications de ces modèles sans doute à cause de leur complexité, à l'exception de celle de B Sabino da Silva and Ziegelman (2017), Deng and Liang (2021), Dhara et al. (2021).

CHAPITRE 2

Évaluation empirique de la contribution à la performance d'un portefeuille d'actifs ESG : Cas du marché américain

2.1 Introduction

L'intérêt pour la prise en compte de critères non financiers par le biais de l'investissement socialement responsable (ISR), des décisions éthiques et des questions environnementales, sociales et de gouvernance d'entreprise (ESG), qui sont plus à la mode, n'est pas récent. Il remonte aux années 1700 avec les mouvements religieux pour s'étendre progressivement aux mouvements anti-esclavagistes, anti-guerre, anti-apartheid et anti-nucléaires. Ils s'imposaient comme des contraintes d'exclure certaines entreprises de leurs paniers d'investissement en fonction de leurs valeurs religieuses, personnelles ou politiques (Sparkes, 2003; Hussein and Omran, 2005; Renneboog et al., 2008; Viviers and Eccles, 2012).

L'ISR a beaucoup changé depuis sa création, passant de filtres conçus pour exclure les investissements dans les actions du péché ("sin stocks") à une vision plus large. Cette forme d'investissement s'est enrichie davantage avec l'intégration du concept de responsabilité sociale des entreprises (RSE) (impacts environnementaux et sociaux positifs) qui permet d'évaluer en toute transparence les performances environnementales, sociales et de gouvernance (ESG) des entreprises. La prise en compte des facteurs ESG offre la possibilité d'orienter les investissements vers des entreprises qui s'efforcent en permanence de générer des impacts environnementaux et sociaux positifs, jusqu'alors exclus.

De plus, elle améliore l'analyse financière traditionnelle en identifiant les facteurs extra-financiers pour mieux évaluer la performance financière et les risques futurs. Ce profond changement dans les préférences des investisseurs est loin d'être une tendance passagère. D'un type d'investisseurs qui privilégient la performance financière à un autre type d'investisseurs qui sont prêts à sacrifier une partie de la performance financière pour respecter leurs valeurs éthiques.

La pandémie de la COVID-19 nous a également permis d'illustrer l'importance de tels critères, devenues un préalable, ce qui renforce l'idée que l'investissement ESG soit un investissement à long terme. Cependant, des questions subsistent quant à la performance de l'investissement ESG par rapport à l'investissement traditionnel, et les théories pour répondre à cette question varient. L'un des principaux arguments en faveur de la surperformance de l'investissement ESG est la sous-réaction du marché boursier aux rapports sur la responsabilité sociale des entreprises (RSE)(corporate social responsibility) (voir Gompers et al., 2003; Servaes and Tamayo, 2013; Ioannou and Serafeim, 2015; Amel-Zadeh and Serafeim, 2018). Hormis les événements extrêmes tels que la marée noire de British Petroleum en avril 2010, l'accident nucléaire de Fukushima avec Tokyo Electric Power Company en mars 2011, les graves problèmes sociaux et de gouvernance qui ont entraîné des pertes de vie pendant la grève des mineurs de Lonmin Corporation Marikana du 1er août 2012 et le Scandale du Libor en 2012 avec Barclays.

Des auteurs comme Capelle-Blancard and Petit (2019), Hoepner et al. (2018), Serafeim and Yoon (2022) à travers leurs travaux de recherches ont apporté des preuves empiriques sur le fait que les marchés boursiers ne soient pas suffisamment sensibles à l'impact sur la valeur des événements ESG ce qui renforce davantage l'hypothèse d'efficacité du marché. Le deuxième argument en faveur de la surperformance de l'investissement ESG est sans aucun doute sa popularisation (voir Bolton, 2016; Hartzmark and Sussman, 2019; Dźaz et al., 2021), qui conduit à une demande croissante pour les actions et les obligations émises par ces entreprises. L'offre d'actifs ESG restant limitée, cette demande supplémentaire pourrait entraîner une hausse des prix. De plus, l'investissement ESG présente certains désavantages en raison de la restriction des offres qui sont liées aux critères supplémentaires pour la sélection des entreprises. Cela restreint le choix des investisseurs, en particulier pour ceux qui souhaitent diversifier leur portefeuille ou investir dans des secteurs spécifiques. Alors qu'une large gamme d'actifs serait plus probable d'offrir le meilleur compromis risque-rendement car la diversification permet de réduire le risque sans réduire les rendements attendus (Markowitz, 1991).

C'est ce que montrent les stratégies d'intégration ESG les plus couramment utilisées par les gérants d'actifs, telles que le Screening, le Best-in-class et le Thematic, y compris par l'exposition des actions à certains facteurs ESG, leur score ESG ou certaines thématiques. Notre attention sera portée sur l'intégration des critères ESG dans l'optimisation du portefeuille et travaillerons dans ce sens. Il s'agit d'une préoccupation pertinente pour les investisseurs et les solutions potentielles présentées enrichiraient également la littérature académique et élargiraient le modèle de sélection de portefeuille de Markowitz afin d'inclure un troisième critère au-delà du rendement et du risque attendus. Cette volonté d'intégrer des critères supplémentaires (contraintes) dans la sélection du portefeuille est de plus en plus fréquente dans la littérature (voir Anagnostopoulos and Mamanis, 2010; Steuer et al., 2013; Hadi et al., 2016; Qi et al., 2017) pour des raisons diverses comme la prise en compte de la liquidité, les coûts de transaction, les moments d'ordre supérieur à 2 (voir Fang et al., 2006; Lo et al., 2006; Jana et al., 2009; Usta and Kantar, 2011; Zhang et al., 2012; Yue and Wang, 2017) et dernièrement les facteurs ESG.

La grande majorité des articles traitant de la prise en compte des critères ESG choisissent une approche multi-objectifs (Multicriteria Decision Making (MCDM)) passant ainsi du bi-critère return-risk (Schandl et al., 2001; Ehrgott et al., 2004), à un tri-critère (ou multi-critère) return-risk-ESG ou une méthode avec une fonction d'utilité générale de Von Neumann and Morgenstern (2007). Non seulement ils aident à la gestion des risques pour des rendements plus élevés, mais ils respectent également les contraintes ESG. Dans les travaux liés aux méthodes multi-objectifs on peut citer les travaux de Hallerbach, Ning, Soppe, and Spronk (2004) qui utilisent une optimisation multi-objectifs et un processus de décision étape par étape (stepwise decision process) pour construire un ensemble de portefeuilles réalisables correspondant aux contraintes des investisseurs. Lundström and Svensson (2014) choisissent de résoudre un problème d'optimisation multi-objectifs avec trois fonctions objectives : rendement, risque et la somme pondérée des scores ESG.

Bilbao-Terol et al. (2012) utilisent des méthodes d'optimisation multi-objectifs et de "fuzzy logic" en raison de la nature multidimensionnelle du problème. Quant à l'approche avec une fonction d'utilité, des auteurs comme Utz et al. (2014, 2015) ont utilisé une approche d'optimisation tri-critère avec un objectif supplémentaire au-delà du rendement et du risque permettant d'étendre le modèle de portefeuille de Markowitz et de passer de frontière non dominée à surface non dominée. Ballestero et al. (2012) ont adopté une approche différente en créant deux sous-ensembles d'actifs constitués d'actifs conformes à l'éthique et d'autres actifs non conformes à l'éthique, puis en appliquant leur modèle, liant les objectifs financiers et éthiques au coefficient d'aversion au risque absolu et les objectifs sont liés selon le profil de l'investisseur. Schmidt (2020) propose dans son article d'étendre le modèle de Markowitz avec le *preferences for the portfolio ESG value* (PESGV) (préférences des investisseurs pour la valeur ESG du portefeuille) qui est proportionnel à la somme pondérés des scores ESG des actifs du portefeuille contrôlable par l'intermédiaire d'un hyper-paramètre.

Contrairement aux approches précédentes, des auteurs tels que Brandstetter and Lehner (2015) ont pris une autre direction, proposant une approche qui permet la quantification des facteurs sociaux et d'influence, puis applique le modèle d'optimisation du portefeuille Black-Litterman. Nous n'avons pas procédé à un examen exhaustif des travaux réalisés, ce qui n'était pas notre objectif, mais une tentative de sélection représentative. Entre autres, l'hypothèse de normalité des rendements, le grand nombre de paramètres d'entrée nécessaires pour estimer dans ce cas le rendement attendu, la variance de chaque actif ainsi que les estimations de la corrélation entre chaque paire d'actifs.

Cela engendre des erreurs d'estimation car, dans la pratique traditionnelle, l'estimation par maximum de vraisemblance est utilisée, ce qui entraîne une sous-performance du portefeuille, une concentration sur quelques actifs avec des allocations extrêmes, ce qui viole le principe de diversification. C'est à partir de ce constat que Michaud (1989) affirme que les modèles maximisent l'erreur d'estimation et sont souvent peu utiles en pratique. Dans le même sens, DeMiguel et al. (2009) montre que le portefeuille à poids égal ($1/N$) est capable de surpasser plusieurs modèles de Markowitz avec des améliorations. Dans cette volonté de réduire les conséquences des erreurs d'estimation suite à l'utilisation d'échantillons de rendements historiques, de nombreux travaux ont été menés pour améliorer l'estimation du vecteur d'espérance des rendements (voir Jorion, 1986; Longerstaey and Spencer, 1996; Martellini, 2009; Green et al., 2013) et la matrice variance-covariance (Laloux et al., 1999; Plerou et al., 2000; Laloux et al., 2000; Ledoit and Wolf, 2003).

En raison de son principe de diversification, une corrélation faible ou négative (covariance) entre les rendements des actifs diminue simultanément avec la variance du rendement du portefeuille, d'où l'utilité de réfléchir à des méthodes d'estimation offrant une meilleure analyse du portefeuille. La matrice de variance-covariance est utilisée pour mesurer les co-mouvements entre les actifs, connus sous le nom de dépendance, en raison de la simplicité de sa mise en œuvre. Mais les preuves empiriques soulignent qu'en plus du regroupement de la volatilité (clustering volatility), les corrélations entre les actifs varient davantage pendant les périodes de turbulences que pendant les périodes de calme, par exemple, les travaux de Clare et al. (1998); Longin and Solnik (2001); Pelletier (2006). Une limite qui sera comblée par le modèle Dynamic Conditional Correlation GARCH de Engle (2002), qui étend le modèle GARCH univarié de Bollerslev (1986) à un modèle multivarié donnant ainsi la possibilité de décomposer la covariance conditionnelle en écart-type et en corrélation conditionnelle.

Malgré cette capacité à prendre en compte la dynamique dans les corrélations conditionnelles, d'autres inconvénients persistent, comme l'utilisation de la corrélation linéaire comme mesure de la dépendance. Elle ne capture que les dépendances linéaires, et est donc insuffisant pour décrire les corrélations entre les variables autres que les distributions elliptiques, et ne parvient donc pas à capturer les corrélations de queue, déformant ainsi les interprétations possibles. Jondeau and Rockinger (2006b), Harvey et al. (2010), Kim et al. (2014) font partie des auteurs qui ont considéré les moments d'ordre supérieur à 2, pour n'en citer que quelques-uns.
Ils sont souvent négligés en raison de difficultés de calcul, alors qu'elles pourraient fournir un moyen adéquat de modéliser la dépendance de la queue inférieure. Cette dépendance de la queue inférieure peut être interprétée comme la probabilité de valeurs négatives extrêmes simultanées et peut gravement nuire à la diversification du portefeuille. Une solution plus adéquate que la précédente est ensuite apparue. Elle permet de construire une distribution conjointe plus flexible sans aucune hypothèse sur les distributions marginales avec aussi la capacité de capturer la dépendance non linéaire et les dépendances symétriques et asymétriques de la queue. D'après les travaux de Sklar (1959), les copules offrent la possibilité de décomposer une distribution multivariée en une structure de dépendance et un ensemble de distributions marginales et de les modéliser indépendamment. Elles sont devenus populaires en finance et en actuariat depuis leur introduction par Frees and Valdez (1998); Embrechts et al. (2001, 2002); McNeil et al. (2005) et de nombreuses applications concluantes s'en sont suivies. Dans le contexte de la gestion de portefeuille, nous nous intéressons à l'utilisation de données à grande échelle.

Cependant, le nombre de copules multivariées est limité, contrairement au nombre de copules bivariées, c'est pourquoi les copules gaussiennes (Xue-Kun Song, 2000) et Student t (Demarta and McNeil, 2005) avec un grand succès. Au fil du temps, la copule Student s'est avéré être mieux adaptée pour des applications en finance en raison de sa capacité à capturer la dépendance contrairement à la copule gaussienne. Bien que Brechmann et al. (2014) affirment que les performances des copules gaussiennes augmentent avec le nombre de dimensions. Jusqu'à présent, les copules n'ont été construites qu'à partir de distributions elliptiques, ce qui a pour conséquence l'impossibilité de modéliser des structures de dépendance asymétriques qui correspondraient mieux aux données financières, comme cela est possible avec les copules archimédiennes.

Ce type de copule est plus facile à construire. Elles sont construites plus facilement avec des formes de dépendance explicites beaucoup moins restrictives qui peuvent capturer correctement la dépendance entre les événements extrêmes comme les copules de Gumbel et de Clayton (voir Costinot et al., 2000; Longin and Solnik, 2001; Melchiori, 2003; Koizol and Kunisch, 2005; Nguyen and Bhatti, 2012). Malgré tous les avantages que ces deux classes de copules peuvent présenter, elles sont qualifiées de rigides. Cette rigidité se traduit par une même structure de dépendance entre toutes les paires de variables et à un petit nombre de paramètres pour contrôler les dépendances de queue de toutes les paires de variables. La solution vient de cette classe relativement nouvelle de modèles de copules, qui utilisent une série de copules bivariées pour construire des copules multivariées.

Initialement introduit par Joe (1996), suivi de Bedford and Cooke (2002), Kurowicka and Cooke (2006), Aas et al. (2009), Czado (2010), cette classe de copules offre la possibilité de sélectionner différentes copules bivariées parmi de nombreuses copules bivariées en fonction de leurs propriétés. Comme pour les copules précédentes, les distributions marginales peuvent être modélisées indépendamment de la structure de dépendance, sauf que la structure de dépendance de la copule de vine est modélisée comme une cascade de différentes copules bivariées conditionnelles. On parle de construction de copules par paires (Pair-Copula Constructions). Une extension logique de tous les modèles de copule est de les coupler avec un modèle GARCH conduisant à un modèle Copula-GARCH. Il permet aux paramètres de la copule d'évoluer dans le temps en fonction de toutes les informations passées, permettant des structures de dépendance dynamiques mieux adaptées aux données financières, comme avec Patton (2006), Berg and Aas (2009) et Ausin and Lopes (2010).

Néanmoins, pour tenir compte de manière optimale des changements dynamiques entre les périodes de hausse et de baisse dans les séries temporelles financières reconnus par certains chercheurs comme Lamoureux and Lastrapes (1990), Hamilton and Susmel (1994), Hess (2003) soutenant la pertinence de recourir au modèle Markov regime-switching. Une idée reprise dans le cadre de la modélisation multivariée, où des fluctuations distinctes peuvent correspondre à des changements dans la structure dépendante. De nombreuses études suggèrent qu'il existe deux régimes avec un régime de forte dépendance caractérisée par des rendements faibles et volatils et un régime de faible dépendance avec des rendements élevés et stables. La combinaison des modèles Markov regime-switching et copule permet d'avoir suffisamment de flexibilité dans la modélisation des structures de dépendance. Sauf que cette combinaison a un défaut majeur. Il soutient que dans chaque régime, les dépendances restent constantes et que les tendances à la baisse ou à la hausse persistent dans le temps. Par conséquent, une combinaison de modèles de Markov regime-switching, de GARCH et de copule serait plus appropriée pour combler cette lacune. Il est important de noter que moins d'attention a été accordée à cette possibilité de modélisation dans la littérature, nous ne pouvons donc citer que les travaux de Okimoto (2008), Chollete et al. (2009), Stöber and Czado (2014), Fei et al. (2017), Muteba Mwamba and Mwambi (2021).

Le modèle de Markowitz d'origine considérait la variance comme une mesure du risque. Or cette mesure fausse la perception du risque par les investisseurs et conduit à des interprétations erronées parce qu'elle ne fait pas de distinction entre les variations à la hausse et à la baisse.

En termes de gestion des risques, il faut être attentif aux pertes potentielles, d'où l'émergence de nombreuses alternatives telles que la semi-variance, la Value-at-Risk (VaR) et Conditionnal Value-at-Risk (CVaR). En raison de ses propriétés mathématiques intéressantes telles que la sous-additivité et la convexité (Artzner et al., 1999; Kolm et al., 2014), la CVaR introduite par Rockafellar et al. (2000) est devenue l'une des mesures de risque les plus couramment utilisées. En plus de celles précédemment citées, nous ajoutons le lower partial moment (LPM) ou moment partiel inférieur. C'est le cas général de la semi-variance, qui a été proposé par Markowitz (1959) pour remplacer la variance comme mesure du risque.

On trouve les premières traces de ce modèle dans les travaux de Stone (1973) puis après il a été introduit dans la théorie du portefeuille par Bawa (1978), Fishburn (1977) et Nawrocki (1991, 1992, 1999). Parmi les auteurs ayant eu recours à cette mesure de risque on peut citer : Foo and Eng (2000); Sadefo Kamdem et al. (2011); Gao et al. (2017); Yao et al. (2021); Luan et al. (2022). L'avantage de ce modèle est qu'il ne nécessite aucune hypothèse distributionnelle, ce qui est très utile en gestion de portefeuille. Cependant, l'incertitude due à la nature stochastique des rendements financiers n'est pas prise en compte par ce modèle. C'est pourquoi dans ce qui suit, nous utiliserons la programmation stochastique pour intégrer cette incertitude en construisant un ensemble de scénarios représentant des distributions futures possibles.

Notre étude contribue à la littérature existante de la manière suivante.
Premièrement, mettre en évidence les avantages de l'intégration des critères ESG dans l'optimisation des portefeuilles en fournissant des preuves empiriques qu'une préférence pour les actifs ayant des notations ESG élevées pourrait réduire la performance des portefeuilles. Deuxièmement, illustrer l'importance de la prise en compte d'une structure de dépendance dynamique pour détecter les changements de régime des rendements boursiers afin de mieux comprendre leurs co-mouvements ainsi que les risques associés. Le reste du document est structuré comme suit. La section 2.2 décrit les méthodes utilisées pour modéliser la structure de dépendance et optimiser le modèle. Les sections 2.3 et 2.4 présentent respectivement les résultats empiriques et une discussion ces résultats. La section 2.5 conclut notre papier.

2.2 Cadre théorique

2.2.1 Modélisation univariée

Modèle GARCH et GJR-GARCH

Introduit par Bollerslev (1986), le modèle GARCH est une méthode de modélisation de la variance conditionnelle variable dans le temps qui, contrairement au modèle ARCH, permet une mémoire plus longue et une structure de retard plus flexible plus la variance conditionnelle précédente.
Un modèle GARCH(p,q) est un processus stochastique définit comme suit : Pour illustrer le modèle GARCH, supposons que $\{x_t\}_{t\in\mathbb{Z}}$ soit un processus stochastique décomposable en termes d'espérance de rendements et de résidus tel que :

$$x_t = \mu + \epsilon_t \tag{2.1}$$

où

$$\begin{cases} \epsilon_t & = \sigma_t z_t \\ \sigma_t^2 & = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \end{cases} \tag{2.2}$$

avec $0 < \alpha_0, \forall i = 1, \ldots, p, 0 < \alpha_1 \leqslant 1, \forall j = 1, \ldots, q, 0 < \beta_1 \leqslant 1$ et la variable z_t indépendante et identiquement distribuée selon une loi normale $\mathcal{N}(0, \sigma_t^2)$.

L'hypothèse de normalité sera revisitée dans la suite après avoir compris les spécificités des séries temporelles financières. Le modèle GARCH (1,1) est le modèle univarié le plus simple et le plus couramment utilisé et peut être défini comme

$$\begin{cases} \epsilon_t & = \sigma_t z_t \\ \sigma_t^2 & = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{cases} \tag{2.3}$$

où le processus de variance σ_t^2 dans GARCH(1,1) est stationnaire lorsque $\alpha_1 + \beta_1 < 1$.

Le taux de convergence dépend de $\alpha_1 + \beta_1$ et plus l'amplitude de la persistance est proche de 1, plus le temps pour qu'un choc soit oublié sur le marché est long. Du fait de l'incapacité du modèle GARCH à faire la différence entre les chocs négatifs et positifs et de tenir compte du phénomène de levier, Glosten, Jagannathan, and Runkle (1993) ont proposé le modèle GJR-GARCH(p,q) comme extension et représenté par l'expression :

$$\begin{cases} \epsilon_t & = \sigma_t z_t \\ \sigma_t^2 & = \alpha_0 + \sum_{i=1}^{p} (\alpha_i + \gamma_i \mathbb{1}_{\epsilon_{t-i}<0}) * \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \end{cases} \tag{2.4}$$

En particulier, le modèle GJR-GARCH(1,1) s'exprime comme suit :

$$\begin{cases} \epsilon_t & = \sigma_t z_t \\ \sigma_t^2 & = \alpha_0 + (\alpha_1 + \gamma \mathbb{1}_{\epsilon_{t-1}<0}) * \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{cases} \tag{2.5}$$

où la positivité des variances conditionnelles est assurée par $\alpha_0 > 0$, $\alpha_1 \geqslant 0$, $\beta_1 \geqslant 0$ et $\alpha_1 + \gamma \geqslant 0$, par conséquent, si $\gamma > 0$, une mauvaise nouvelle a un impact plus important sur la variance conditionnelle qu'une bonne nouvelle, en plus, la stationnarité des variances est possible à condition que $\alpha_1 + \beta_1 + 0.5\gamma < 1$.

Les études empiriques récentes, comme celles de Lamoureux and Lastrapes (1990), Hamilton and Susmel (1994), Bauwens et al. (2014) montrent généralement que la volatilité réelle est affectée par de nombreux chocs qui de toute façon ne durent pas longtemps. Alors que les méthodes qui devraient les modéliser ont une forte volatilité conditionnelle, cela affecte leur capacité à faire de bonnes prédictions. Cette incapacité à capturer ces changements structurels possibles dans le processus de variance a incité des chercheurs comme Gray (1996), Dueker (1997), Klaassen (2002) proposent d'incorporer les modèles GARCH dans un cadre de changement de régime. Cela permet de considérer de manière assez parcimonieuse l'existence de deux régimes de volatilité différents et de pouvoir basculer de l'un à l'autre. Et ainsi doter la volatilité de la possibilité de suivre une classe de modèles GARCH selon sa nature, évitant ainsi de modéliser la variance dépendante de toutes les informations mais s'adaptant rapidement aux changements dynamiques. Connu sous le nom de Markov-Switching GARCH ou Regime-Switching GARCH, ce modèle a pour objectif de permettre aux paramètres de varier dans le temps selon un processus de Markov discret régi par une variable d'état latente (observée) $s_t \in \{1, 2\}$ avec une probabilité de transition $p_{ij} = Pr(s_t = j | s_i = i, \mathcal{I}_{t-1})$ où $\mathcal{I}_{t-1}$ est un ensemble d'informations disponibles à l'instant t.

Habituellement, ces probabilités sont regroupées dans une matrice de transition

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix} \tag{2.6}$$

avec pour contraintes $0 < p_{ij} < 1$, $\forall i, j \in \{1, 2\}$ et $\sum_{j=1}^{2} p_{ij} = 1$.
Ainsi, nous définissons le modèle Markov-Switching GARCH qui sera utilisé avec des régimes hétérogènes GARCH(1,1)-GJR GARCH(1,1) comme

$$\begin{aligned} x_t &= \mu + \epsilon_t \\ \epsilon_t &= \sigma_t z_t \\ \sigma_t^2 &= \begin{cases} \alpha_{0,k} + \alpha_{1,k}\epsilon_{t-1}^2 + \beta_{1,k}\sigma_{t-1}^2 & k = 1 \\ \alpha_{0,k} + (\alpha_{1,k} + \gamma_k \mathbb{1}_{\epsilon_{t-1}<0})\epsilon_{t-1}^2 + \beta_{1,k}\sigma_{t-1}^2 & k = 2 \end{cases} \end{aligned} \tag{2.7}$$

Le point de départ d'une bonne modélisation financière est sans doute d'observer des caractéristiques très spécifiques des séries de rendements empiriques qui diffèrent de la distribution normale.
Par conséquent, il est inapproprié de supposer que la distribution conditionnelle (z_t) est une distribution normale, une distribution alternative serait :

- la distribution de Student
  C'est l'une des distributions couramment utilisées en raison de ses queues épaisses pour modéliser les rendements des actifs, et sa fonction de densité est donnée par :

$$f_\nu(z) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi s \nu}} \left(1 + \frac{z^2}{s\nu}\right)^{-\frac{\nu+1}{2}} \tag{2.8}$$

  où ν est le degré de liberté qui contrôle l'épaisseur des queues et s le paramètre d'échelle.

2.2.2 Modélisation multi-variée

Modèle DCC-GARCH

L'importance de prédire le co-mouvement entre plusieurs actifs ouvre la porte à de meilleurs outils de prise de décision en raison de l'extension des modèles GARCH univariés aux modèles GARCH multivariés. Engle (2002) propose d'assouplir l'hypothèse de corrélation conditionnelle constante dans le temps du modèle de Bollerslev (1990), qui consiste à décomposer la matrice de covariance conditionnelle en écarts-types conditionnels et à la transformer en une matrice de corrélation conditionnelle.
D'un point de vue pratique, le modèle DCC-GARCH a l'avantage d'un processus d'estimation en deux étapes, où la première étape est utilisée pour estimer un modèle GARCH univarié pour chaque série temporelle, qui dans notre cas sera le GJR-GARCH modèle conditionnel à la distribution de Student. Puis dans la deuxième étape, la matrice de covariance est estimée à partir des résidus standardisés obtenus à l'étape précédente.

Le processus DCC-GARCH est défini comme :

$$x_t = \mu + \epsilon_t \tag{2.9}$$

$$\epsilon_t = H_t^{1/2} z_t \tag{2.10}$$

$$H_t = D_t R_t D_t \tag{2.11}$$

- x_t est un vecteur $n \times 1$ des rendements logarithmiques de n actifs au temps t
- μ est un vecteur $nfois1$ de la valeur attendue du rendement conditionnel logarithmique de n actifs à l'instant t.
- ϵ_t est un vecteur $n \times 1$ des rendements logarithmiques corrigés de la moyenne
- H_t est une matrice $n \times n$ des covariances conditionnelles de ϵ_t au temps t
- $H_t^{1/2}$ représente la matrice $n \times n$ et obtenue par factorisation de $H_t^{1/2}(H_t^{1/2})^\top = H_t$ de Cholesky
- z_t est un vecteur $n \times 1$ des erreurs i.i.d

Comme mentionné précédemment, pour supprimer les contraintes associées à la matrice de corrélation en lui permettant de varier dans le temps, la variance conditionnelle peut être décomposée en écart type conditionnel (D_t) et en matrice de corrélation (R_t) sous la forme :

$$D_t = \begin{pmatrix} h_{1_t}^{1/2} & 0 & \dots & 0 \\ 0 & h_{2_t}^{1/2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_{n_t}^{1/2} \end{pmatrix} \quad R_t = \begin{pmatrix} 1 & \rho_{1,2_t} & \dots & \rho_{1,n_t} \\ \rho_{2,1_t} & 1 & \dots & \rho_{2,n_t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n,1_t} & \rho_{n,2_t} & \dots & 1 \end{pmatrix}$$

ce qui nous donne

$$H_t = D_t R_t D_t \begin{pmatrix} h_{1_t} & h_{1_t}^{1/2} h_{2_t}^{1/2} \rho_{1,2_t} & \dots & h_{1_t}^{1/2} h_{n_t}^{1/2} \rho_{1,n_t} \\ h_{2_t}^{1/2} h_{1_t}^{1/2} \rho_{2,1_t} & h_{2_t} & \dots & h_{2_t}^{1/2} h_{n_t}^{1/2} \rho_{2,n_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_t}^{1/2} h_{1_t}^{1/2} \rho_{n,1_t} & h_{n_t}^{1/2} h_{2_t}^{1/2} \rho_{n,2_t} & \dots & h_{n_t} \end{pmatrix}$$

Dans le cas actuel, les erreurs standardisées z_t sont multivariées (Eq. 2.10) et la logique voudrait qu'on utilise la généralisation multivariée des distributions gaussienne ou de student comme sa densité conjointe déterminée respectivement par

$$f(z_t) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z_t^\top z_t\right) \qquad \text{multivariate Gaussian distribution} \tag{2.12}$$

$$f(z_t; \nu) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\pi(\nu-2))^{n/2}\left[1 + \frac{z_t^\top z_t}{\nu-2}\right]^{-\frac{n+\nu}{2}}} \qquad \text{multivariate Student distribution} \tag{2.13}$$

Malgré des propriétés intéressantes, qu'en est-il des distributions multivariées non gaussiennes ou de Student, même lorsque les erreurs sont corrélées ? Pour pallier ces inconvénients, une alternative est la copule.

Modèle de copule

Analytiquement, supposons que les variables aléatoires $x_1, \dots, x_n$ aient une fonction de distribution continue $F_1, \dots, F_n$ et une fonction de distribution conjointe correspondante $F(x_1, \dots, x_n)$.

Une copule est une fonction de distribution cumulative multivariée $C : [0,1]^n \to [0,1]$, $n \in \mathbb{N}$, telle que toutes les marginales univariées sont uniformes sur l'intervalle $[0,1]$. D'après le théorème de Sklar (Sklar, 1959) pour toute distribution cumulative multivariée $F : \mathbb{R}^n \to [0,1]$ avec les marginales univariées $F_1, \ldots, F_n$ on pourrait lui associer une unique copule définie comme une fonction continue $C : [0,1]^n \to [0,1]$ qui satisfait l'équation

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \tag{2.14}$$

Si F est absolument continue et $F_1, \ldots, F_n$ sont strictement croissantes on peut passer à sa fonction de densité de probabilité et écrire

$$\begin{aligned} f(x_1, \ldots, x_n) &= \frac{\partial^n C(F_1(x_1), \ldots, F_n(x_n))}{\partial F_1(x_1) \ldots \partial F_n(x_n)} \times \frac{\partial F_1(x_1)}{\partial x_1} \times \frac{\partial F_2(x_2)}{\partial x_2} \times \ldots \times \frac{\partial F_n(x_n)}{\partial x_n} \\ &= c(F_1(x_1), \ldots, F_n(x_n)) \times f_1(x_1) \times f_2(x_2) \times \ldots f_n(x_n) \end{aligned} \tag{2.15}$$

où $f_1, \ldots, f_n$ sont les densités marginales et c est la densité de copules définie comme

$$c(u_1, \ldots, u_n) = \frac{\partial^n}{\partial u_1, \ldots, \partial u_n} C(u_1, \ldots, u_n)$$

Si nous ne nous intéressons qu'à la structure de dépendance du vecteur aléatoire x, nous pouvons appliquer le probability integral transform (PIT) à ses marginales connues pour obtenir les données dites u :

$$u_i = F_i(x_i) \rightsquigarrow \mathcal{U}[0,1] \, , j = 1, \ldots, n$$

Dans la classe des copules elliptiques, nous utiliserons la copule de student la plus courante dans les applications financières, par opposition à la copule normale, qui ne suppose aucune dépendance dans les queues de la distribution. En considérant $u = (u_1, \ldots, u_n)$, t_η^{-1} l'inverse de la fonction de distribution cumulative de la distribution standard de Student univariée avec η degrés de liberté, cette copule peut être obtenue à partir de la distribution de student multivariée avec η degrés de liberté et matrice de corrélation R et est donnée par

$$C^t_{\eta,R}(u_1, \ldots, u_n) = \int_{-\infty}^{t_\eta^{-1}(u_1)} \ldots \int_{-\infty}^{t_\eta^{-1}(u_n)} \frac{\Gamma\left(\frac{\eta+n}{2}\right)\left(1 + \frac{u^\top R^{-1} u}{\eta}\right)^{-\frac{\eta+n}{2}}}{\Gamma\left(\frac{\eta}{2}\right)\sqrt{(\pi\eta)^n |R|}} du \tag{2.16}$$

L'absence d'une forme analytique pour la modélisation de la densité conduit à des difficultés d'estimation, car les épaisseurs de queue des distributions ont des coefficients de corrélation linéaires qui ne peuvent pas être définis en raison des seconds moments infinis. Il est nécessaire d'introduire des mesures plus robustes, car elles ne dépendent pas des distributions marginales et sont invariantes aux transformations croissantes et continues telles que les transformations logarithmiques.

Pour quantifier plus précisément les dépendances et relier les corrélations de rang, les copules et les dépendances dans les queues de distribution, le τ de Kendall ou le ρ de Spearman peuvent être exprimés sous la forme d'une paire de (X, Y) dans l'équation suivante

$$\tau_{XY} = 4 \int_0^1 \int_0^1 C(u, v)\, dC(u, v) - 1 \tag{2.17}$$

$$\rho_{XY} = 12 \int_0^1 \int_0^1 C(u, v)\, dudv - 3 \tag{2.18}$$

Sur la base d'observations empiriques, l'hypothèse de la copule de Student d'une fréquence égale d'événements extrêmes associés à des rendements négatifs ou positifs simultanés semble incompatible avec la réalité. La notion de dépendance de queue permet l'analyse et la modélisation des dépendances entre événements extrêmes, avec les coefficients de la corrélation de queue supérieure λ_U et de la corrélation de queue inférieure λ_L. Dans le cas bivarié avec deux variables aléatoires X_1 et X_2, soit u le seuil

$$\lambda_U = \lim_{x \to 1^-} P(F_1(X_1) > u | F_2(X_2) > u) = \lim_{x \to 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \tag{2.19}$$

$$\lambda_L = \lim_{x \to 0^+} P(F_1(X_1) \leqslant u | F_2(X_2) \leqslant u) = \lim_{x \to 0^+} \frac{C(u, u)}{u} \tag{2.20}$$

La généralisation multivariée de ces coefficients est une alternative intéressante pour la gestion de portefeuille utilisant des copules d'archimédiennes qui permettent des dépendances non linéaires et des dépendances de queue asymétriques. Alors que les caractéristiques de cette famille correspondent à ce qu'on pourrait espérer, nous choisirons de prendre les copules des vines en raison de leur grande souplesse. Nous choisissons la copule de student comme benchmark pour modéliser la structure de dépendance entre les marginales qui seront modélisées par des GJR-GARCH(1,1).

Comme la copule donne une description complète de la distribution conjointe, son couplage au modèle DCC-GARCH permettra aux paramètres de la copule de varier dans le temps. Considérons un vecteur de n composantes, $x_t = (x_{1,t}, \ldots, x_{n,t})$, la structure de dépendance entre les séries marginales décrite par un modèle copule-GARCH variant dans le temps avec η degrés de liberté a pour densité à chaque temps t

$$c_{\eta, R_t}(F_1(x_{1,t}), \ldots, F_n(x_{n,t})) = \frac{f_{\eta, R_t}(x_{1,t}, \ldots, x_{n,t})}{\prod_{i=1}^n f_i(x_{i,t})} \tag{2.21}$$

où

- f_{η, R_t} est la densité multivariée de student avec η degrés de liberté et R_t la matrice de corrélation
- f_i est la densité de la distribution univarié de student.

Ensuite, pour la variation des paramètres de la copule, nous supposons que la matrice de corrélation R_t varie en fonction de ses valeurs passées et de ses données historiques.

Nous avons l'équation

$$R_t = (1 - a - b)R + a\Psi_{t-1} + bR_{t-1} \tag{2.22}$$

où

- a et b sont des paramètres non négatifs
- R est la matrice de corrélation inconditionnelle des erreurs standardisées $n \times n$ constante
- Ψ_{t-1} est une matrice $n \times n$ dont le (i, j)-ième élément est donné par

$$\Psi_{t-1} = \frac{\sum_{h=1}^{m} x_{i,t-h} x_{j,t-h}}{\sqrt{\sum_{h=1}^{m} x_{i,t-h}^2 \sum_{h=1}^{m} x_{j,t-h}^2}}$$

Une équation qui à quelques détails près ressemble à celle dans le processus de matrices dynamiques pour l'estimation de R_t dans le modèle DCC-GARCH, (voir Tse and Tsui, 2002; Silvennoinen and Teräsvirta, 2009).

Modèle de copule de vine

Sachant que de nombreuses copules bivariées bien connues dans la littérature n'ont pas d'extensions multivariées directes, notre idée est de tirer avantage le plus possible de la riche famille des copules bivariées. Parce que lorsque ces extensions multivariées existent, comme c'est le cas avec certaines copules archimédiennes, elles disposent d'un faible nombre de paramètres pour gérer les dépendances dans les queues de dépendance restreignant le degré de dépendance de la queue et imposent une structure de dépendance identique entre toutes les paires de variables. D'après les travaux de Joe (1996); Bedford and Cooke (2002); Aas et al. (2009), la copule de vine offre une meilleure approche, surmontant les limitations des copules précédentes en fournissant un moyen flexible de construire des copules multivariées.

Cette flexibilité s'ajoute à l'avantage inhérent à toutes les copules, à savoir la possibilité de modéliser les arêtes séparément, qui passeraient ici par GJR-GARCH (eq. 2.5) ou MS-GARCH (eq. 2.7) et une structure de dépendance avec construction de copules par paires(Pair Copula Construction). Selon l'illustration de Aas et al. (2009), le concept de copule de vine est une copule multivariée construite via une cascade de copules bivariées conditionnelles et de fonctions de distribution marginales. De façon plus explicite, considérons les variables aléatoires $x_1, \ldots, x_n$ avec sa fonction de densité factorisable en une série de densités univariées conditionnelles comme

$$f(x_1, \ldots, x_n) = f_1(x_1) \times f_{2|1}(x_2|x_1) \times f_{3|2,1}(x_3|x_1, x_2) \times \ldots \times f_{n|1,2,\ldots,n-1}(x_n|x_1, x_2, \ldots, x_{n-1}) \tag{2.23}$$

En utilisant le théorème de Sklar dans l'équation (2.15), la première densité conditionnelle dans l'équation (2.23) peut s'écrire :

$$f_{2|1}(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = c_{1,2}(F_1(x_1), F_2(x_2)) f_2(x_2) \tag{2.24}$$

avec $c_{1,2}$ la densité de copule des variables x_1 et x_2.
Pour la seconde densité conditionnelle on a

$$f_{3|2,1}(x_3|x_1, x_2) = \frac{f_{2,3|1}(x_2, x_3|x_1)}{f_{2|1}(x_2|x_1)} = c_{2,3|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) f_{3|1}(x_3|x_1) \tag{2.25}$$

où $f_{3|1}(x_3|x_1) = c_{1,3}(F_1(x_1), F_3(x_3))f_3(x_3)$ et ainsi de suite.
Cependant, face à un grand nombre de structures de vigne régulières possibles, aucune expression analytique exacte n'est proposée dans la littérature, mais des solutions existent, et pour plus de détails (voir Morales-Napoles, 2010; Dissmann et al., 2013). Dans ce nombre de décompositions possibles en paires de copules, deux cas particuliers de copules R-vine se distinguent : la canonical vine et la drawable vine avec prédominance pour de la première dans les applications liées à la gestion de portefeuille qui a l'avantage d'avoir un faible nombre de permutation avec $\frac{n!}{2}$ au lieu de $\frac{n!}{2} \times 2 \times \binom{n-2}{2}$. La densité $f(x_1, \ldots, x_n)$ correspondant à une C-vine peut s'écrire

$$f(x_1, \ldots, x_n) = \prod_{k=1}^{n} f_k(x_k) \times \prod_{j=1}^{n-1} \prod_{j=1}^{n-1} c_{i,i+j|1,\ldots,i-1} \left(F(x_i|x_1, \ldots, x_{i-1}), F(x_{i+j}|x_1, \ldots, x_{i-1}) \right) \tag{2.26}$$

Parmi la multitude des copules bivariées, nous avons sélectionné quelques-unes en fonction de leurs propriétés et sont référencées dans le tableau 2.1.

Negative tail	Centre	Positive tail
Student (t)	Gaussian (N)	Student (t)
Clayton (C)	Frank (F)	Gumbel (G)
Clayton-Gumbel (BB1)	Clayton 90 (C90)	Joe (J)
Joe-Clayton (BB7)	Clayton 270 (C270)	Clayton-Gumbel (BB1)
Gumbel 180 (SG)		Joe-Clayton (BB7)
Joe 180 (SJ)		Joe-Frank (BB8)
Joe-Frank 180 (SBB8)		Clayton 180 (SC)

TABLE 2.1 – Les copules bivariées utilisées

2.2.3 Modèle d'optimisation

Dans un article publié en 1952, Markowitz (1952b) a établi le concept de la théorie moderne du portefeuille par la diversification et la sélection de portefeuille. Il offre la possibilité de sélectionner le portefeuille avec le rendement le plus élevé pour un niveau de risque donné ou le portefeuille avec le risque le plus faible pour un rendement donné, ce que l'on appelle la frontière efficiente. Il existe différentes formulations pour l'optimisation de portefeuille, nous nous concentrerons sur le portefeuille Global Minimum Variance (GMV), qui est un portefeuille optimal spécifique avec la plus petite variance parmi tous les portefeuilles sur la frontière efficiente.
Le problème d'optimisation de portefeuille à résoudre est :

$$\begin{aligned} \min_{w} \quad & w^\top \Sigma w \\ \text{s.t.} \quad & w^\top \mu = \mu_{target} \\ & w^\top \mathbf{1} = 1 \\ & w_i \in [0,1] \, , \forall i = 1, \ldots, n \end{aligned} \tag{2.27}$$

avec

- Σ représente la matrice de covariance qui est un paramètre inconnu et doit être estimé
- $\mu = (\mu_1, \ldots, \mu_n)$ représente le vecteur des espérances des rendements des actifs
- $\mathbf{1}$ représente un vecteur de uns
- μ_{target} représente le rendement cible du portefeuille choisi
- w représente le vecteur de pondération.

Dans la littérature financière, l'utilisation de pondérations négatives dans un portefeuille est moins fréquente et son problème d'optimisation peut être formulé comme :

$$\begin{aligned} \min_{w} \quad & w^\top \Sigma w \\ \text{s.t.} \quad & w^\top \mu = \mu_{target} \\ & w^\top \mathbf{1} = 1 \\ & w_i \in [-1, 1] \, , \forall i = 1, \ldots, n \end{aligned} \tag{2.28}$$

Ce problème d'optimisation de portefeuille (eq. 2.28) est similaire à celui définit à l'équation (2.27) à un détail près, la condition qui permet aux poids de prendre des valeurs négatives. Dans l'optimisation d'un portefeuille, donner la possibilité aux pondérations de prendre des valeurs négatives permet une diversification plus importante en incluant des positions vendeuses sur des actifs dont on anticipe une baisse de rentabilité. Ce qui peut avoir comme conséquence de générer potentiellement des rendements plus élevés. Hormis les avantages qu'elles sont susceptibles d'apporter en termes de rentabilité, les pondérations négatives peuvent permettre de limiter les pertes et parer à des conditions de marché défavorables. Toutefois, considérer que les pondérations peuvent prendre des valeurs négatives a aussi son lot d'inconvénients. La mise en application de cette stratégie d'investissement peut s'avérer difficile en raison des restrictions réglementaires qui l'encadre. En plus, elle expose les investisseurs à des risques très élevés avec des pertes illimitées dans le cas où le prix de l'actif augmente au lieu de diminuer.

Il est important de noter que l'attribution des poids négatifs à des actifs dans un portefeuille peut engendrer des covariances négatives entre ces actifs et les autres, et donc impacter la matrice de covariance. Ce qui peut entraîner une matrice des covariances non définie positive si certaines covariances négatives ne peuvent pas être prises en compte. Alors que l'une des propriétés mathématiques souhaitables de la matrice des covariances est qu'elle soit définie positive. Cette propriété garantit que l'optimisation du portefeuille est bien posée mathématiquement, avec une solution unique et stable.

Mais comme nous pouvons le constater, dans certains cas, il peut arriver que cette matrice ne soit pas définie positives (voir Green and Hollifield, 1992; Jagannathan and Ma, 2003; Li et al., 2003; Bodnar et al., 2017a,b). En plus de cela, d'autres inconvénients peuvent apparaitre comme les difficultés d'estimation de la matrice des covariances liées aux effets du bruit dans les séries temporelles, à la fenêtre temporelle sélectionnée et produire des estimations biaisées. Des problématiques sur lesquelles se sont intéressées des auteurs comme Markowitz (1994), Bodnar et al. (2016, 2019, 2022), Gulliksson and Mazur (2020), Gulliksson and Mazur (2020) mais que nous n'aborderons pas dans notre papier.

Le modèle Global Minimum Variance (GMV) présente plusieurs limites notamment le choix de la corrélation linéaire comme outil de modélisation de la structure de la dépendance car elle mesure que le degré de dépendance et n'est pas invariante à des transformations strictement croissantes non linéaires. À cela, nous pouvons ajouter le choix de la variance comme une mesure du risque. La variance, par définition, est une mesure de la dispersion qui prend chaque valeur individuelle dans un ensemble de données et la compare à la moyenne. Sauf que, lorsqu'une seule valeur ou un petit nombre de valeurs sont éloignées de la moyenne (valeurs extrêmes) cela peut entraîner une variance élevée. Dans le cadre du modèle GMV, l'interprétation intuitive d'une variance élevée est qu'elle est assimilée à un risque de pertes importantes, ce qui est susceptible de fausser le processus d'optimisation du portefeuille.

Cette caractéristique de la variance s'avère problématique dans le sens où elle accorde une importance égale écarts positifs et négatifs. Alors qu'en terme de gestion de risque, l'idéal serait une mesure de risque qui pénalisera que les écarts négatifs c'est-à-dire en dessous de la moyenne contrairement à la variance.
Les rendements négatifs ont un impact psychologique et émotionnel plus important sur les agents économiques qui ont une aversion plus forte pour les pertes que pour les gains de même magnitude (voir Shefrin, 2002; Thaler and Ganser, 2015; Statman, 2017). Cette hyper-sensibilité des agents économiques aux rendements négatifs s'explique par la diminution significative de la richesse qu'ils engendrent et qui peut compromettre les objectifs financiers à long terme. Les investisseurs sont donc souvent plus sensibles aux rendements négatifs, car ils peuvent entraîner une diminution significative de la richesse et compromettre les objectifs financiers à long terme. En vue d'avoir une mesure plus appropriée pour les investisseurs qui cherchent à minimiser les pertes et à gérer le risque de manière plus spécifique, Markowitz (1959) a suggéré comme mesure de risque la semi-variance. Elle est un cas particulier du lower partial moment (LPM). L'idée de pénaliser les écarts en dessous d'un certain seuil par rapport à un taux s'inscrit dans celle de Markowitz a été concrétisée par Fishburn (1977). Dans le cas continu, le Lower Partial Moment (LPM) peut-être défini comme

$$LPM_{a,\tau}(x) = \int_{-\infty}^{\tau} (\tau - x)^a f(x) dx \tag{2.29}$$

où

- a représente le taux auquel les écarts en dessous du seuil τ sont pénalisés
- f une fonction de densité.

Dans le cas discret, on a :

$$LPM_{a,\tau}(x) = E[\max(\tau - x, 0)^a] \tag{2.30}$$

L'asymétrie des distributions de probabilité des séries temporelles financières est une raison pour laquelle le choix de la variance comme mesure de risque est inapproprié. De plus, l'optimum global d'un problème d'optimisation mean-variance est atteint au sens de Von Neumann and Morgenstern (1947) que si la fonction d'utilité est quadratique, si les rendements des actifs appartiennent à la famille des distributions elliptiques. Ces deux conditions ont été remises en cause à travers la littérature. D'abord, les fonctions d'utilité quadratiques qui ont été jugées irréalistes par Levy and Markowitz (1979), Elton et al. (2009), Kahneman and Tversky (2013).

Parce que l'utilisation d'une fonction quadratique sous-entend l'hypothèse d'une aversion au risque symétrique signifiant ainsi que par le fait que les investisseurs considèrent manière proportionnelle les gains et les pertes de magnitude égale. En plus de cela, les fonctions d'utilité quadratiques ne considèrent que les deux premiers moments d'ordre des distributions de probabilité des rendements d'où l'hypothèse de distributions elliptiques.

De nombreuses études empiriques ont montré que la distribution des rendements financiers était loin d'être elliptique parmi lesquelles nous pouvons citer sans être exhaustif celles de Constantinides and Malliaris (1995), Longin (1996) et Andersen et al. (2001). Tous ces arguments motivent l'utilisation du Lower Partial Moment (LPM) en tant que mesure de risque. Un constat qui est en accord avec les travaux de Bawa (1975); Bawa and Lindenberg (1977); Bawa (1978) qui montré que le mean-TSV où TSV est le Target SemiVariance qui correspond au cas où $a = 2$ dans l'équation (2.29) pouvait être utilisé pour trouver un ensemble admissible de dominance stochastique troisième ordre [1]. Alors que la variance et un LPM avec $a < 2$ (eq 2.29) donnent au mieux un ensemble admissible de dominance stochastique de deuxième ordre.

2.2.4 Intégration des notations ESG d'optimisation d'un portefeuille

Les notations ESG visent à fournir des informations aux acteurs du marché (investisseurs, analystes et dirigeants d'entreprise) en passant au crible des masses de données afin d'en extraire des informations sur divers éléments de performance et de risque en matière environnementale, sociale et de gouvernance. Les investisseurs s'appuient sur ces informations pour prendre des décisions d'investissement, tandis que les entreprises utilisent les notations pour obtenir un retour d'information de la part de tiers sur la qualité de leurs initiatives en matière de développement durable. Il existe deux conceptions principales de l'ESG. La première considère que l'ESG reflète l'impact d'une entreprise sur le bien-être de ses parties prenantes, telles que les employés, les fournisseurs, les clients, la communauté locale et l'environnement. Cette vision de l'ESG ("doing good") est celle qui vient à l'esprit de la plupart des investisseurs individuels lorsqu'ils pensent à la qualité de l'ESG. Selon un point de vue opposé, l'ESG mesure l'impact des facteurs sociétaux et environnementaux sur l'entreprise. Sur la base de cette définition, un cadre ESG fournit un ensemble de facteurs de risque que l'entreprise peut atténuer par des investissements ciblés ou une modification de l'activité opérationnelle. Cette vision de l'ESG (l'impact des risques environnementaux et sociaux sur la performance financière) est celle qui est principalement adoptée par les agences de notation ESG. Il existe une tension entre ces deux points de vue, car les notations ne mesurent pas l'impact d'une entreprise sur la Terre et la société. En fait, elles mesurent le contraire : l'impact potentiel du monde sur l'entreprise et ses actionnaires. Les notations ESG sont conçues pour aider les investisseurs à identifier et à comprendre les risques ESG importants d'un point de vue financier pour une entreprise. Les entreprises sont évaluées sur la base d'informations accessibles au public, telles que les sources médiatiques et les rapports annuels, et des notes sont attribuées pour chaque thème "E", "S" et "G", ainsi qu'une note globale. Les scores ESG sont utilisés par les investisseurs comme une approximation de la performance ESG dans leurs stratégies d'investissement. Un score ESG élevé indique qu'une entreprise gère bien ses risques ESG par rapport à ses pairs, tandis qu'un score ESG faible indique que l'entreprise est comparativement plus exposée aux risques ESG non gérés.

1. L'ensemble des portefeuilles qui offrent une espérance de rendement plus élevée, une volatilité plus faible et un troisième moment (asymétrie) plus favorable par rapport à un autre portefeuille de référence

Les évaluations et les notes ESG, combinées à l'analyse financière, peuvent aider les investisseurs à mieux comprendre le potentiel de croissance positive et durable à long terme d'une entreprise.
Si, par exemple, une entreprise reçoit une mauvaise note de la part d'un fournisseur de données ESG, son action peut être considérée comme un "actif non durable" par les investisseurs et être exclue de leur portefeuille d'investissement. Si plusieurs investisseurs suivent ce raisonnement, cela peut finir par avoir un impact négatif sur le cours de l'action de l'entreprise. Les notations peuvent donc constituer un outil d'étalonnage interne très précieux pour guider la prise de décision et améliorer les performances en matière de développement durable. Pour inclure les notations ESG dans un processus d'optimisation de portefeuille, nous avons suivi les étapes suivantes :

1. Sélectionner un fournisseur de notations ESG fiables sur les entreprises cotées en bourse
2. Définir des contraintes ESG
3. Intégrer les contraintes ESG dans le choix des actifs pour la composition du portefeuille
4. Collecter des données au près d'une Datastream
5. Utiliser les modèles d'optimisation pour définir les pondérations en tenant compte des objectifs d'investissement (maximisation du rendement, minimisation du risque, ...)
6. Surveiller régulièrement les notations ESG des entreprises dans votre portefeuille et effectuez des rééquilibrages si nécessaire pour maintenir l'exposition ESG souhaitée.

2.3 Résultats empiriques

Afin d'analyser l'impact de l'intégration des critères ESG dans l'optimisation des portefeuilles, nous avons fait le choix des notations de l'agence MSCI [2] parmi les agences de notations ESG les plus renommées comme Sustainalytics, FTSE Russel Bloomberg ESG et bien d'autres. Notre choix se justifie par le fait que cette agence qui fait partie des plus fiables et ses notations sont en libre accès. Il aurait été intéressant de considérer les notations ESG d'une seconde agence mais chacune d'elle a sa méthodologie spécifique pour évaluer et noter les performances ESG des entreprises. Par conséquent, nous pouvons noter des divergences dans les notations ESG ce qui peut remettre en question la cohérence et la comparabilité des résultats (voir DeFond et al., 2005; Gond et al., 2011; Kotsantonis and Serafeim, 2019; Billio et al., 2021; Berg et al., 2022), et donc nous poursuivrons qu'avec celles de l'agence MSCI. Sur la base des notations de cette MSCI (voir fig 2.1), nous avons décidé de construire trois portefeuilles. Ils seront composés respectivement des actifs de notations LEADER, AVERAGE et LAGGARD appartenant à des secteurs d'activité différents afin de garantir une diversification pour tous.

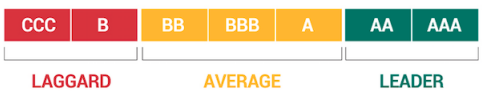


FIGURE 2.1 – Scores ESG selon MSCI

2. https://www.msci.com/our-solutions/esg-investing/esg-ratings-climate-search-tool/issuer

Les données relatives aux cotations boursières des actifs sélectionnés ont été collectées depuis la Datastream de Yahoo `https://en.finance.yahoo.com/`. Dans l'intention de maintenir l'exposition ESG souhaitée et respecter contraintes précédemment évoquées, nous sommes parvenus à construire 3 portefeuilles de 39 actifs chacun. Les 39 actifs ont été sélectionné de façon à s'assurer d'une diversification des secteurs d'activité et dont les notations ESG ne varient pas énormément durant la période d'analyse pour éviter de changer la composition des portefeuilles pour maintenir l'exposition ESG. Donc nous faisons l'hypothèse que les notations ESG des actifs sélectionnés ont très peu variées durant la période d'étude. En effet, les modèles d'optimisation utilisés pour évaluer les performances des portefeuilles sont d'une complexité élevée qui engendre des temps de calcul très longs et donc nous n'avons pas souhaité rajouter une difficulté supplémentaire en modifiant constamment la composition des portefeuilles. Le portefeuille 1 est composé uniquement de LEADER (AAA et AA), de AVERAGE (A et BBB) pour le portefeuille 2 et pour le portefeuille 3, les actifs les moins bien notés (BBB, BB, B et CCC) comme nous pouvons le voir aux tables 2.2, 2.3 et 2.4. La composition de portefeuille 3 s'explique par la difficulté de trouver des actifs avec de mauvaises notations ESG appartenant à des secteurs d'activité différents. Ces actifs sont en général concentrés dans une poignée de secteurs d'activité comme secteurs des énergies fossiles, de l'armement, du tabac, du charbon, de l'extraction minière, pétrolière et gazière. Afin de nous aider à analyser l'impact des notations ESG sur la performance des portefeuilles, nous avons construit un portefeuille benchmark qui sera composé de l'ensemble des actifs des 3 portefeuilles soit 117 actifs.

nº	Company	Ticker	Industry	ESG Rating
1	3M Company	MMM	Industrial Conglomerates	AAA
2	Adobe Systems Incorporated	ADBE	Software & Services	AAA
3	AES Corporation	AES	Utilities	AA
4	Agilent Technologies Inc.	A	Health Care Equipment & Supplies	AA
5	Alliant Energy Corp	LNT	Utilities	AA
6	Applied Materials Inc.	AMAT	Semiconductors & Semiconductor Equipment	AA
7	Autodesk Inc.	ADSK	Software & Services	AA
8	Automatic Data Processing Inc.	ADP	Software & Services	AAA
9	Ball Corporation	BLL	Containers & Packaging	AA
10	Best Buy Co. Inc.	BBY	Retail - Consumer Discretionary	AAA
11	Biogen Inc.	BIIB	Biotechnology	AA
12	BlackRock Inc.	BLK	Asset Management & Custody Banks	AA
13	CBRE Group Inc.	CBG	Real Estate Management & Services	AAA
14	Danaher Corporation	DHR	Health Care Equipment & Supplies	AA
15	Ecolab Inc.	ECL	Specialty Chemicals	AAA
16	Edwards Lifesciences Corporation	EW	Health Care Equipment & Supplies	AA
17	Fidelity National Information Services Inc.	FIS	Software & Services	AA
18	Fortune Brands Home & Security Inc.	FBHS	Building Products	AA
19	General Mills Inc.	GIS	Food Products	AAA
20	Gilead Sciences Inc.	GILD	Biotechnology	AA
21	Hasbro Inc.	HAS	Leisure Products	AAA
22	Hess Corporation	HES	Oil & Gas Exploration & Production	AAA
23	Johnson Controls International plc	JCI	Building Products	AAA
24	Kellogg Company	K	Food Products	AAA
25	Microsoft Corporation	MSFT	Software & Services	AAA
26	Molson Coors Brewing Company Class B	TAP-A	Beverages	AAA
27	Newmont Mining Corporation	NEM	Metals and Mining - Precious Metals	AA
28	NVIDIA Corporation	NVDA	Semiconductors & Semiconductor Equipment	AAA
29	Texas Instruments Incorporated	TXN	Semiconductors & Semiconductor Equipment	AAA
30	United Rentals Inc.	URI	Trading Companies & Distributors	AA
31	UnitedHealth Group Incorporated	UNH	Health Care Providers & Services	AA
32	V.F. Corporation	VFC	Textiles, Apparel & Luxury Goods	AA
33	Ventas Inc.	VTR	Real Estate Management & Services	AA
34	W.W. Grainger Inc.	GWW	Trading Companies & Distributors	AAA
35	Waters Corporation	WAT	Health Care Equipment & Supplies	AAA
36	Welltower Inc.	WELL	Real Estate Management & Services	AAA
37	Xcel Energy Inc.	XEL	Utilities	AAA
38	Xylem Inc.	XYL	Industrial Machinery	AAA
39	Zoetis Inc.	ZTS	Pharmaceuticals	AA

TABLE 2.2 – Les composants du portefeuille 1

nº	Company	Ticker	Industry	ESG Rating
1	Abbott Laboratories	ABT	Health Care Equipment & Supplies	A
2	Advanced Micro Devices Inc.	AMD	Semiconductors & Semiconductor Equipment	A
3	Akamai Technologies Inc.	AKAM	Software & Services	A
4	Allstate Corporation	ALL	Property & Casualty Insurance	BBB
5	Alphabet Inc. Class A	GOOGL	Interactive Media & Services	BBB
6	Amazon.com Inc.	AMZN	Retail - Consumer Discretionary	BBB
7	American Water Works Company Inc.	AWK	Utilities	A
8	Ameriprise Financial Inc.	AMP	Asset Management & Custody Banks	A
9	AmerisourceBergen Corporation	ABC	Health Care Providers & Services	BBB
10	Apple Inc.	AAPL	Technology Hardware, Storage & Peripherals	A
11	Arthur J. Gallagher & Co.	AJG	Multi-Line Insurance & Brokerage	BBB
12	Assurant Inc.	AIZ	Property & Casualty Insurance	A
13	AutoZone Inc.	AZO	Retail - Consumer Discretionary	BBB
14	Avery Dennison Corporation	AVY	Containers & Packaging	BBB
15	Centene Corporation	CNC	Health Care Providers & Services	BBB
16	CF Industries Holdings Inc.	CF	Specialty Chemicals	BBB
17	Chipotle Mexican Grill Inc.	CMG	Restaurants	BBB
18	ConocoPhillips	COP	Oil & Gas Exploration & Production	A
19	Devon Energy Corporation	DVN	Oil & Gas Exploration & Production	A
20	Entergy Corporation	ETR	Utilities	A
21	Exxon Mobil Corporation	XOM	Integrated Oil & Gas	BBB
22	Fiserv Inc.	FISV	Software & Services	A
23	FMC Corporation	FMC	Specialty Chemicals	BBB
24	Freeport-McMoRan Inc.	FCX	Metals and Mining - Non-Precious Metals	BBB
25	Goldman Sachs Group Inc.	GS	Investment Banking & Brokerage	A
26	Hilton Worldwide Holdings Inc	HLT	Hotels & Travel	A
27	Iron Mountain Inc.	IRM	Software & Services	A
28	Marathon Oil Corporation	MRO	Oil & Gas Exploration & Production	BBB
29	Occidental Petroleum Corporation	OXY	Integrated Oil & Gas	BBB
30	Thermo Fisher Scientific Inc.	TMO	Health Care Equipment & Supplies	BBB
31	Union Pacific Corporation	UNP	Road & Rail Transport	BBB
32	Valero Energy Corporation	VLO	Oil & Gas Refining, Marketing, Transportation	A
33	VeriSign Inc.	VRSN	Software & Services	A
34	Vertex Pharmaceuticals Incorporated	VRTX	Biotechnology	A
35	Visa Inc.	V	Software & Services	A
36	Vulcan Materials Company	VMC	Construction Materials	BBB
37	Walmart Inc.	WMT	Retail - Food & Staples	BBB
38	Waste Management Inc.	WM	Commercial Services & Supplies	A
39	Yum! Brands Inc.	YUM	Restaurants	BBB

TABLE 2.3 – Les composants du portefeuille 2

Face aux difficultés qu'on peut rencontrer en travaillant directement sur les séries de prix qui sont susceptibles de présenter une variance non constante, une tendance à long terme, effectuer une différence logarithmique comme $r_{i,t} = \ln p_{i,t} - \ln p_{i,t-1}$ est judicieux. En effet, une opération de différence logarithmique sur les séries des prix permet d'éliminer la tendance, stationnariser, stabiliser la variance et transformer les variations multiplicatives en variations additives. Dans la suite de notre analyse nous travaillerons uniquement sur les séries temporelles des rendements obtenues à l'issue de cette de différence logarithmique. Notre jeu de données couvre la période du 2 janvier 2015 au 31 août 2022 et nous l'avons subdivisé en deux sous-échantillons. Le premier sous-échantillon contient 1004 observations allant du 2 janvier 2015 au 28 décembre 2018 et constitue l'échantillon in-sample. Les 925 observations restantes constitueront l'échantillon out-sample. Nous avons effectué des prévisions itératives à travers une fenêtre mobile (rolling window) dont la taille de la fenêtre est de 1004 observations. C'est-à-dire les prévisions des structures de dépendances réalisées sur une fenêtre permettront à établir les pondérations des portefeuilles pour les 5 jours suivants. Ainsi, 185 prévisions hors échantillon ont été réalisées engendrant ainsi un rééquilibrage du portefeuille à la même fréquence à partir du 29 décembre 2018.

nº	Company	Ticker	Industry	ESG Rating
1	Ametek Inc	AME	Electrical Equipment	BB
2	Air Products and Chemicals Inc.	APD	Specialty Chemicals	BBB
3	Aon plc	AON	Multi-Line Insurance & Brokerage	BB
4	AT&T Inc.	T	Telecommunication Services	BB
5	Capital One Financial Corporation	COF	Consumer Finance	BB
6	Carnival Corporation	CCL	Hotels & Travel	B
7	Charter Communications	CHTR	Media & Entertainment	BB
8	Cincinnati Financial Corporation	CINF	Property & Casualty Insurance	BB
9	Darden Restaurants Inc.	DRI	Restaurants	BB
10	Dish Network	DISH	Media & Entertainment	CCC
11	Essex Property Trust Inc.	ESS	Real Estate Management & Services	BB
12	Extra Space Storage Inc.	EXR	Real Estate Management & Services	BB
13	Ford Motor Company	F	Automobiles	B
14	General Motors Company	GM	Automobiles	BB
15	Lennar Corporation	LEN	Construction individuelle	BB
16	M&T Bank Corporation	MTB	Banks	BB
17	Marriott International Inc.	MAR	Hotels & Travel	BB
18	Meta	META	Interactive Media & Services	B
19	Mid-America Apartment Communities Inc.	MAA	Real Estate Management & Services	BB
20	Monster Beverage Corporation	MNST	Beverages	B
21	Netflix Inc.	NFLX	Media & Entertainment	BB
22	News Corporation	NWSA	Media & Entertainment	B
23	NRG Energy Inc.	NRG	Utilities	BB
24	O'Reilly Automotive Inc.	ORLY	Retail - Consumer Discretionary	BB
25	Packaging Corporation of America	PKG	Containers & Packaging	BB
26	PG&E Corporation	PCG	Utilities	BB
27	Philip Morris International Inc.	PM	Tobacco	BBB
28	Public Storage	PSA	Real Estate Management & Services	BB
29	Royal Caribbean Cruises Ltd.	RCL	Hotels & Travel	BB
30	Simon Property Group Inc.	SPG	Real Estate Management & Services	BB
31	Southwest Airlines Co.	LUV	Airlines	BB
32	TransDigm Group Incorporated	TDG	Aerospace & Defense	BB
33	UDR Inc.	UDR	Real Estate Management & Services	BB
34	Universal Health Services Inc.	UHS	Health Care Providers & Services	CCC
35	Verisk Analytics Inc	VRSK	Professional Services	BB
36	Vornado Realty Trust	VNO	Real Estate Management & Services	B
37	Wells Fargo & Company	WFC	Banks	BB
38	Wynn Resorts Limited	WYNN	Casinos & Gaming	BB
39	Zimmer Biomet Holdings Inc.	ZBH	Health Care Equipment & Supplies	BB

TABLE 2.4 – Les composants du portefeuille 3

Les statistiques descriptives des rendements sont répertoriées aux tables 2.5, 2.6 et 2.7 fournissent certaines informations utiles pour la suite. Nous constatons que les valeurs des variances sont souvent plus élevés que les valeurs de leurs moyennes respectives qui sont quant à elles proches de 0. Ce qui est une caractéristique d'un niveau élevé d'instabilité dans le temps, et donc de volatilité élevée.

On dénombre respectivement 7, 2 et 13 actifs ayant des moyennes négatives dans les portefeuilles 1, 2 et 3. Les valeurs des coefficients d'asymétrie (skewness) sont en majorité négatives sinon une faible proportion d'entre elles sont positives. Des valeurs négatives des coefficients d'asymétrie signifient que ces actifs financiers sont susceptibles de générer des rendements négatifs. L'idéal aurait été d'avoir dans la composition de nos portefeuilles un grand nombre d'actifs avec des coefficients d'asymétrie positifs. Mais dans les portefeuilles 1, 2 et 3, les actifs ayant cette caractéristique sont respectivement au nombre de 6, 8 et 4. La totalité des valeurs des coefficients d'aplatissement (kurtosis) est différent de 3. En plus des valeurs des coefficients d'asymétrie non nulles et des coefficients d'aplatissement supérieures à 3, les statistiques du test de Jarque and Bera (1980) constituent une preuve supplémentaire militant en faveur de la non normalité des distributions des rendements.

Assets	mean (%)	variance (%)	std	skewness	kurtosis	min	max	Jarque-Bera stat	ADF stat
A	0.05968	0.02756	0.01660	-0.49460	5.09706	-0.11667	0.09385	2166.79381	-13.27099
ADBE	0.08509	0.04069	0.02017	-0.32196	7.75420	-0.15953	0.16313	4866.08543	-12.27846
ADP	0.05581	0.02513	0.01585	-0.79355	14.82632	-0.16550	0.11157	17870.46649	-12.41285
ADSK	0.06327	0.05610	0.02368	-0.23177	7.42543	-0.17269	0.15013	4448.90963	-12.25342
AES	0.03211	0.04357	0.02087	-0.86054	11.17523	-0.18996	0.13940	10275.77117	-12.34836
AMAT	0.06878	0.06505	0.02550	-0.51036	6.02169	-0.22762	0.12938	2998.20394	-12.70569
BALL	0.02546	0.02953	0.01718	-0.68028	16.60220	-0.20558	0.11414	22302.79720	-13.97826
BBY	0.03111	0.05837	0.02416	0.03427	9.94431	-0.16548	0.19458	7948.60871	-12.22359
BIIB	-0.02903	0.07399	0.02720	0.27282	59.56181	-0.34574	0.36446	285163.01136	-13.69423
BLK	0.03239	0.03134	0.01770	-0.37736	8.42781	-0.14681	0.12684	5754.65701	-12.39682
CBRE	0.04267	0.04513	0.02124	-0.40020	10.69853	-0.18881	0.15532	9251.10161	-11.79094
DHR	0.07381	0.02071	0.01439	0.04889	5.64018	-0.10231	0.09256	2557.62902	-11.96818
ECL	0.02340	0.02580	0.01606	0.00915	21.49817	-0.12491	0.20017	37147.04134	-13.12207
EW	0.07480	0.04028	0.02007	-0.14933	10.63884	-0.18744	0.16205	9104.41374	-12.36314
FBHS	0.01638	0.04128	0.02032	-0.56019	9.71336	-0.17249	0.14761	7684.22580	-11.24849
FIS	0.01976	0.02935	0.01713	-0.70734	15.18809	-0.18140	0.15109	18701.59480	-12.87326
GILD	-0.02086	0.02735	0.01654	-0.20388	4.68105	-0.09499	0.09288	1774.56272	-13.12338
GIS	0.01912	0.01869	0.01367	-0.65169	9.28235	-0.12116	0.11139	7061.81082	-13.49451
GWW	0.04027	0.03374	0.01837	0.04014	12.98751	-0.14754	0.17001	13557.81435	-12.47055
HAS	0.01868	0.04215	0.02053	-0.29049	20.17956	-0.20705	0.19614	32756.99949	-12.65614
HES	0.02530	0.08984	0.02997	-1.01181	21.00772	-0.41051	0.18495	35800.58430	-12.29743
JCI	0.00883	0.02848	0.01687	-0.64877	6.83950	-0.14295	0.11009	3895.15740	-12.07350
K	0.00545	0.01864	0.01365	-0.36355	10.47706	-0.11563	0.10116	8865.15901	-13.30684
LNT	0.03149	0.01846	0.01359	-0.32188	8.09275	-0.10110	0.09634	5297.27863	-13.62431
MMM	-0.01437	0.02185	0.01478	-0.87419	11.45321	-0.13863	0.11866	10788.96675	-13.00847
MSFT	0.08923	0.03037	0.01743	-0.20032	9.05482	-0.15945	0.13293	6602.82373	-13.27734
NEM	0.03941	0.05276	0.02297	-0.20152	4.72966	-0.14193	0.13119	1811.01612	-11.95919
NVDA	0.17631	0.08875	0.02979	0.02154	7.54711	-0.20771	0.26088	4578.21930	-12.61801
TAP.A	-0.02591	0.08245	0.02871	-0.37614	30.13867	-0.38283	0.24498	73053.26216	-13.38374
TXN	0.05847	0.03292	0.01814	-0.14046	5.62369	-0.12585	0.12689	2548.27022	-12.34299
UNH	0.08500	0.02839	0.01685	-0.43426	15.07241	-0.18967	0.12044	18320.01443	-15.28862
URI	0.05373	0.07882	0.02807	-0.47793	6.78887	-0.20302	0.20029	3777.82062	-12.18809
VFC	-0.02676	0.04298	0.02073	-0.66608	8.18550	-0.13836	0.13105	5527.95662	-11.94980
VTR	-0.02899	0.06157	0.02481	-2.93048	50.22565	-0.33677	0.19701	205516.19135	-11.12589
WAT	0.04997	0.02846	0.01687	-0.24123	9.52859	-0.13140	0.11949	7316.27538	-12.46833
WELL	-0.00020	0.04658	0.02158	-1.09604	31.87835	-0.27965	0.20847	82065.63898	-11.19977
XEL	0.03736	0.01813	0.01346	-0.86564	15.96471	-0.13573	0.10668	20726.25135	-14.20333
XYL	0.04522	0.02995	0.01731	-0.36607	7.88116	-0.12814	0.10672	5035.39580	-12.52542
ZTS	0.06661	0.02642	0.01626	-0.42796	11.42982	-0.15894	0.11314	10559.13571	-13.66713

TABLE 2.5 – Statistiques descriptives des actions du portefeuille 1

Les valeurs des statistiques du test de Jarque-Bera sont largement supérieures aux seuils soient de 5,99 et 9,21 pour les niveaux de signification de 5% et 1%, respectivement. Nous pouvons alors rejeter l'hypothèse de normalité des distributions des rendements. Les valeurs des statistiques du test augmented Dickey–Fuller (ADF stat) nous confortent dans l'idée que l'opération de différence logarithmique sur les séries temporelles des prix nous a produit des séries temporelles stationnaires. À ce stade de l'analyse, nous pouvons dire que le portefeuille 2 est le plus susceptible de produire des rendements positifs, suivie du portefeuille 1 et enfin du portefeuille 3.

Assets	mean (%)	variance (%)	std	skewness	kurtosis	min	max	Jarque-Bera stat	ADF stat
AAPL	0.09070	0.03454	0.01859	-0.28033	5.56095	-0.13771	0.11316	2510.79495	-11.59860
ABC	0.02501	0.03191	0.01786	-0.26283	7.78983	-0.13962	0.11168	4899.47395	-13.40150
ABT	0.04287	0.02397	0.01548	-0.47331	6.95553	-0.10298	0.10378	3960.52329	-12.96036
AIZ	0.04404	0.02956	0.01719	-1.71606	24.90297	-0.21985	0.13233	50791.95551	-12.43698
AJG	0.07008	0.01962	0.01401	-1.06144	19.34580	-0.16529	0.12427	30443.36559	-12.46128
AKAM	0.01845	0.04104	0.02026	-0.64952	18.58574	-0.18309	0.19230	27899.55134	-12.90295
ALL	0.02807	0.02246	0.01499	-1.15914	14.15718	-0.15186	0.10169	16541.19372	-12.20094
AMD	0.17932	0.14327	0.03785	0.58712	10.96821	-0.27746	0.42062	9780.05954	-12.49511
AMP	0.03668	0.05132	0.02265	-0.56734	22.07173	-0.27210	0.20405	39259.07710	-11.51878
AMZN	0.10919	0.04168	0.02042	0.18894	6.70422	-0.15140	0.13218	3624.05249	-12.41749
AVY	0.06561	0.03027	0.01740	0.46901	16.47453	-0.14238	0.20596	21885.30985	-12.33078
AWK	0.05248	0.02149	0.01466	-0.35914	15.91672	-0.13393	0.15334	20403.83576	-13.86481
AZO	0.06398	0.02906	0.01705	-1.56247	16.19593	-0.17361	0.11138	21867.89009	-12.14424
CF	0.03221	0.07264	0.02695	-0.31559	3.52808	-0.16713	0.15051	1032.47847	-12.75167
CMG	0.04438	0.05195	0.02279	0.43483	11.39985	-0.15762	0.21866	10506.05807	-11.58975
CNC	0.06400	0.05230	0.02287	-0.57135	8.30122	-0.21179	0.14499	5643.60800	-14.36063
COP	0.02398	0.06700	0.02589	-0.53418	14.36168	-0.28555	0.22485	16669.70774	-11.97949
DVN	0.00763	0.12576	0.03546	-1.06802	18.49292	-0.46836	0.19122	27854.02023	-11.32774
ETR	0.01418	0.02323	0.01524	-0.89071	20.39138	-0.17113	0.13845	33675.66651	-12.91146
FCX	0.01185	0.13558	0.03682	0.02069	5.19835	-0.22731	0.25994	2172.10051	-11.45991
FISV	0.05458	0.02743	0.01656	-0.81568	13.60742	-0.18236	0.11807	15096.30220	-12.91501
FMC	0.04058	0.04171	0.02042	-0.33849	12.43802	-0.21508	0.12840	12471.18903	-12.42080
GOOGL	0.07298	0.03025	0.01739	0.14101	7.00105	-0.12368	0.15065	3945.94444	-13.11457
GS	0.02785	0.03534	0.01880	-0.20070	9.64577	-0.13588	0.16195	7491.11067	-12.24873
HLT	0.04523	0.03780	0.01944	-0.14014	5.85263	-0.13435	0.13147	2759.41771	-11.97416
IRM	0.01557	0.03294	0.01815	-0.62077	8.68565	-0.16299	0.11331	6187.41948	-12.64862
MRO	-0.00576	0.13793	0.03714	-2.21606	45.10579	-0.63209	0.20900	165104.37612	-11.08697
OXY	-0.00652	0.12262	0.03502	-4.87240	122.61813	-0.76411	0.29041	1216087.24016	-11.76636
TMO	0.07509	0.02518	0.01587	-0.16451	3.82435	-0.08651	0.09957	1184.23888	-13.40305
UNP	0.03308	0.02787	0.01670	-0.50002	9.17994	-0.13966	0.12225	6853.68383	-13.12250
V	0.05694	0.02685	0.01639	-0.08049	9.41503	-0.14557	0.12965	7126.75460	-13.48792
VLO	0.04377	0.06875	0.02622	0.06201	13.08373	-0.21345	0.27157	13760.14004	-11.41540
VMC	0.04720	0.04071	0.02018	-0.03879	11.68218	-0.19376	0.16352	10969.53051	-12.89412
VRSN	0.06007	0.02831	0.01682	-0.07886	13.15614	-0.15390	0.15873	13913.63391	-12.03651
VRTX	0.04294	0.05300	0.02302	0.04756	11.78118	-0.23188	0.18922	11156.47612	-13.57709
WM	0.06136	0.01512	0.01229	-0.62101	13.75549	-0.11790	0.08664	15332.01511	-12.58123
WMT	0.02249	0.01939	0.01392	-0.20781	15.83606	-0.12076	0.11072	20170.39787	-11.99022
XOM	0.00152	0.03351	0.01831	-0.14961	6.89416	-0.13039	0.11944	3827.37640	-11.75685
YUM	0.03941	0.02576	0.01605	-0.33765	34.15992	-0.20865	0.20903	93826.22968	-11.63353

TABLE 2.6 – Statistiques descriptives des actions du portefeuille 2

En effet, dans le portefeuille 2, les actions Advanced Micro Devices Inc. (AMD), Amazon.com Inc.(AMZN), Avery Dennison Corporation (AVY), Chipotle Mexican Grill Inc. (CMG), Freeport-McMoRan Inc. (FCX) , Alphabet Inc. Class A (GOOGL), Valero Energy Corporation (VLO), Vertex Pharmaceuticals Incorporated (VRTX) ont la particularité d'avoir à la fois des moyennes positives et des coefficients d'asymétrie positifs. De plus, certaines d'entre elles opèrent dans des secteurs où n'investiraient pas les personnes concernées par les questions environnementales, comme les mines, le pétrole et le gaz. Dans les portefeuilles 1 et 3, certains actifs ont des caractéristiques similaires, à la différence que d'autres ont des moyennes négatives, par exemple Biogen Inc. (BIIB) et Ford Motor Company (F) respectivement.

Assets	mean (%)	variance (%)	std	skewness	kurtosis	min	max	Jarque-Bera stat	ADF stat
AME	0.04281	0.02524	0.01589	-1.11022	21.37040	-0.17633	0.15457	37103.04618	-12.65908
AON	0.05626	0.02166	0.01472	-0.69523	20.91157	-0.18276	0.13219	35302.87287	-12.45414
APD	0.03263	0.02414	0.01554	-0.37964	11.34769	-0.13472	0.12860	10396.23919	-12.30484
CCL	-0.08155	0.12868	0.03587	-1.38440	25.06337	-0.40319	0.33139	51105.53245	-11.18656
CHTR	0.04707	0.03341	0.01828	-0.16619	7.91474	-0.15465	0.13270	5043.81345	-12.97567
CINF	0.03269	0.03399	0.01844	-0.57475	15.38242	-0.14975	0.15556	19124.45475	-12.06513
COF	0.01291	0.05440	0.02332	-0.93540	18.01624	-0.27278	0.17189	26369.82482	-11.79434
DISH	-0.07397	0.06892	0.02625	-0.82152	10.11164	-0.21970	0.15446	8434.95075	-11.89685
DRI	0.04471	0.05878	0.02424	-0.37304	31.23979	-0.28320	0.27264	78484.68141	-10.64635
ESS	0.01198	0.02693	0.01641	-1.17566	20.36776	-0.20925	0.11345	33787.58841	-12.14304
EXR	0.06301	0.02575	0.01605	-1.00609	9.85526	-0.16805	0.07860	8131.94615	-13.86568
F	-0.00041	0.04970	0.02229	0.32201	7.90371	-0.11667	0.21060	5054.25404	-10.94069
GM	0.00479	0.05148	0.02269	-0.20042	8.77350	-0.19023	0.18185	6199.72565	-11.63098
LEN	0.02985	0.05631	0.02373	-0.22227	10.79217	-0.21834	0.19305	9377.23521	-11.68993
LUV	-0.00784	0.05125	0.02264	-0.31237	6.49899	-0.16381	0.13492	3426.16337	-11.32088
MAA	0.04030	0.02426	0.01558	-1.41567	19.37464	-0.18280	0.11004	30815.22601	-13.62740
MAR	0.03534	0.04735	0.02176	0.09125	9.57739	-0.16477	0.17800	7375.19147	-11.04417
META	0.03789	0.05131	0.02265	-1.60814	25.38249	-0.30639	0.16206	52614.70474	-12.58074
MNST	0.04675	0.03232	0.01798	0.13162	10.13674	-0.15594	0.12762	8264.37921	-12.41119
MTB	0.01925	0.04277	0.02068	0.44517	12.32729	-0.14168	0.22391	12277.65897	-12.26369
NFLX	0.07780	0.08344	0.02889	-1.73605	32.26652	-0.43258	0.17419	84649.62793	-11.91494
NRG	0.02106	0.06460	0.02542	-0.28202	11.99900	-0.19794	0.25763	11597.63965	-12.70052
NWSA	0.00424	0.03849	0.01962	-0.46900	6.99199	-0.14502	0.12714	4000.08133	-11.08704
ORLY	0.06682	0.03231	0.01797	-1.36564	21.02025	-0.20939	0.12273	36113.34555	-12.41150
PCG	-0.07593	0.18103	0.04255	-2.74349	78.54034	-0.74148	0.55729	498219.83520	-12.23818
PKG	0.02898	0.03283	0.01812	-0.26316	7.79682	-0.13875	0.14645	4908.29747	-12.67208
PM	0.00852	0.02417	0.01555	-1.49157	16.90267	-0.16932	0.09562	23678.42404	-13.48769
PSA	0.02951	0.02040	0.01428	-0.92126	6.74620	-0.12252	0.06167	3930.83095	-13.67401
RCL	-0.03672	0.12074	0.03475	-0.87628	18.62060	-0.38241	0.25298	28115.03620	-11.10046
SPG	-0.03097	0.06351	0.02520	-1.09241	33.11287	-0.31078	0.24584	88511.82365	-10.62258
T	-0.01956	0.02017	0.01420	-0.57269	8.15795	-0.09696	0.09551	5454.57697	-12.90425
TDG	0.05757	0.05173	0.02274	-0.70108	21.45305	-0.24813	0.21836	37149.28246	-11.71328
UDR	0.01855	0.02673	0.01635	-0.71873	16.47958	-0.19304	0.12938	21994.04580	-13.28266
UHS	-0.00653	0.04904	0.02214	-0.90768	21.36068	-0.25176	0.22583	36938.26820	-12.60305
VNO	-0.06237	0.04423	0.02103	-0.57045	28.12250	-0.27846	0.24213	63671.17758	-11.14377
VRSK	0.05565	0.02180	0.01477	-0.61793	12.45686	-0.11294	0.12475	12594.81650	-12.01593
WFC	-0.01163	0.04160	0.02039	-0.31452	9.70971	-0.17278	0.13571	7609.43376	-12.05157
WYNN	-0.04591	0.10980	0.03314	-0.11028	8.64449	-0.28017	0.24442	6010.11370	-11.93308
ZBH	-0.00144	0.03193	0.01787	-0.15549	11.00121	-0.15076	0.14479	9735.28115	-11.59817

TABLE 2.7 – Statistiques descriptives des actions du portefeuille 3

Sur la base des précédents résultats, nous percevons clairement qu'il y a un écart entre les distributions des rendements et les distributions gaussiennes et une présence de volatilité élevée. On peut dire que les distributions des rendements possèdent toutes les caractéristiques qui justifient de recourir à une modélisation avec les modèles GARCH. En raison des faits stylisés propres à nos différentes séries temporelles, nous avons dans un premier temps fait le choix d'un modèle dont le mécanisme de propagation de la volatilité est asymétrique de sorte que les chocs positifs et négatifs affectent différemment la volatilité future. Parce que les chocs négatifs (mauvaises nouvelles) ont un impact plus important sur la volatilité de la période suivante que les chocs positifs (bonnes nouvelles) de même ampleur. C'est ainsi que le modèle GJR-GARCH a été choisi pour la modélisation de toutes les séries.

Assets	μ	α_0	α_1	β_1	γ	ν^1
A	0.00085	0.00001	0.01213	0.88560	0.10807	5.07743
ADBE	0.00161	0.00001	0.01698	0.89755	0.15755	3.64157
ADP	0.00086	0.00002	0.08509	0.70235	0.17165	4.06813
ADSK	0.00102	0.00005	0.00000	0.79382	0.22988	3.51570
AES	0.00011	0.00001	0.04104	0.90268	0.04642	5.97621
AMAT	0.00115	0.00002	0.00000	0.89169	0.12205	4.78409
BALL	0.00031	0.00000	0.00000	0.93851	0.08899	4.50121
BBY	0.00110	0.00005	0.00000	0.86546	0.05466	3.67342
BIIB	0.00019	0.00000	0.00000	0.98839	0.01139	3.37885
BLK	0.00026	0.00001	0.04935	0.86597	0.08558	5.28004
CBRE	0.00069	0.00001	0.01197	0.89837	0.11831	4.77084
DHR	0.00036	0.00000	0.00000	0.90703	0.14151	4.78976
ECL	0.00052	0.00000	0.01883	0.91868	0.08478	4.99842
EW	0.00119	0.00007	0.04096	0.72269	0.13779	3.14186
FBHS	0.00047	0.00000	0.00367	0.95010	0.05591	5.01657
FIS	0.00087	0.00000	0.01361	0.95852	0.04551	3.66716
GILD	0.00007	0.00000	0.03075	0.98017	-0.02384	3.93493
GIS	0.00032	0.00000	0.00814	0.97864	0.01873	3.94477
GWW	0.00025	0.00000	0.01233	0.97468	0.00699	3.48899
HAS	0.00037	0.00003	0.00449	0.85983	0.12406	2.89181
HES	-0.00104	0.00001	0.00000	0.92602	0.10596	9.64599
JCI	-0.00006	0.00002	0.00467	0.86431	0.08347	4.66964
K	0.00011	0.00001	0.06647	0.83728	0.02656	3.70423
LNT	0.00074	0.00000	0.01353	0.95228	0.02900	6.19692
MMM	0.00066	0.00000	0.04774	0.91742	0.04715	3.68617
MSFT	0.00071	0.00001	0.02403	0.81645	0.27357	3.50047
NEM	0.00010	0.00000	0.01980	0.96925	0.01834	4.64542
NVDA	0.00211	0.00012	0.03404	0.67284	0.31390	3.24751
TAP.A	0.00002	0.00005	0.68114	0.22488	0.18597	18.86347
TXN	0.00134	0.00000	0.00000	0.95591	0.06072	3.89275
UNH	0.00080	0.00000	0.01738	0.94345	0.07633	5.32514
URI	0.00049	0.00002	0.00005	0.94050	0.05651	4.28518
VFC	0.00088	0.00000	0.00013	0.97011	0.02226	3.74565
VTR	0.00087	0.00003	0.00000	0.82637	0.09688	3.58693
WAT	0.00093	0.00001	0.00000	0.86407	0.11398	3.82027
WELL	0.00045	0.00001	0.03192	0.90612	0.06020	5.02201
XEL	0.00082	0.00000	0.03602	0.94475	0.01386	7.43864
XYL	0.00090	0.00000	0.00000	0.93551	0.08533	4.73732
ZTS	0.00083	0.00002	0.07183	0.75642	0.24748	4.16757

[1] conditionnal distribution parameter

TABLE 2.8 – Les estimations des paramètres GARCH du portefeuille 1

Aux tables 2.8, 2.9 et 2.10, nous avons répertorié les paramètres estimés sur la première fenêtre glissante du modèle qui sont tous statistiquement significatifs. Même si certaines incohérences sont à relever notamment au niveau du portefeuille 1 sur l'actif Gilead Sciences Inc.(GILD), sur les actifs Akamai Technologies Inc.(AKAM) et American Water Works Company Inc.(AWK) au portefeuille 2 et sur l'actif News Corporation (NWSA) au portefeuille 3. Ces incohérences se manifestent par des valeurs du paramètre γ qui sont négatives ($\gamma < 0$) ce qui sous-entend que les chocs négatifs diminuent la volatilité. Mais dans l'immense majorité, les valeurs du paramètre γ sont positives ($\gamma < 0$) ce qui montre que le modèle est cohérent avec la réalité que nous cherchons à modéliser.

Assets	μ	α_0	α_1	β_1	γ	ν^1
AAPL	0.00079	0.00001	0.00029	0.87345	0.20286	4.22282
ABC	0.00081	0.00000	0.01435	0.94118	0.06109	3.13858
ABT	0.00081	0.00001	0.00000	0.90409	0.10304	4.24538
AIZ	0.00050	0.00006	0.02089	0.58019	0.18593	3.59101
AJG	0.00057	0.00002	0.06962	0.57200	0.24593	6.22227
AKAM	0.00112	0.00001	0.01179	0.96422	-0.00273	2.84188
ALL	0.00071	0.00000	0.00000	0.97415	0.03156	3.67731
AMD	0.00158	0.00053	0.27295	0.46923	0.01121	3.48002
AMP	0.00024	0.00004	0.04059	0.70998	0.21865	5.11285
AMZN	0.00145	0.00002	0.00887	0.80976	0.29921	3.99718
AVY	0.00054	0.00002	0.00898	0.78949	0.19874	4.51580
AWK	0.00113	0.00000	0.02814	0.95508	-0.00303	5.14933
AZO	0.00044	0.00000	0.01307	0.96941	0.00658	3.65538
CF	0.00040	0.00000	0.00837	0.95937	0.05377	5.38971
CMG	-0.00010	0.00000	0.00000	0.97611	0.03953	2.76707
CNC	0.00133	0.00010	0.05305	0.57636	0.36000	3.95055
COP	-0.00021	0.00000	0.00000	0.94264	0.09637	8.05858
DVN	-0.00095	0.00001	0.00000	0.94984	0.08248	8.97239
ETR	0.00040	0.00000	0.00452	0.94970	0.04464	6.39479
FCX	-0.00094	0.00001	0.01753	0.94993	0.04635	5.60548
FISV	0.00082	0.00000	0.00843	0.88962	0.13050	5.39963
FMC	0.00028	0.00000	0.00000	0.95956	0.07031	4.44745
GOOGL	0.00090	0.00001	0.00830	0.91245	0.12108	3.94491
GS	0.00004	0.00001	0.04305	0.85284	0.10817	5.47635
HLT	0.00042	0.00001	0.02909	0.88860	0.10419	6.78160
IRM	0.00081	0.00000	0.00035	0.98124	0.02473	4.50092
MRO	-0.00056	0.00001	0.00000	0.95640	0.07186	6.29186
OXY	-0.00010	0.00000	0.00114	0.94819	0.07774	5.64233
TMO	0.00072	0.00000	0.00000	0.94483	0.08405	4.35677
UNP	0.00042	0.00000	0.01011	0.96351	0.04687	4.55937
V	0.00098	0.00001	0.01958	0.84660	0.19944	4.46706
VLO	0.00108	0.00000	0.01990	0.94671	0.03730	5.44180
VMC	0.00016	0.00002	0.01944	0.87699	0.11193	5.33852
VRSN	0.00131	0.00001	0.04258	0.87766	0.08872	4.39874
VRTX	0.00012	0.00001	0.00172	0.92790	0.11836	4.69714
WM	0.00073	0.00000	0.00000	0.95762	0.06382	3.36945
WMT	0.00041	0.00000	0.00000	0.96779	0.02260	3.47088
XOM	-0.00002	0.00000	0.03676	0.92598	0.06034	5.85677
YUM	0.00075	0.00001	0.04279	0.86784	0.11630	3.30549

[1] conditionnal distribution parameter

TABLE 2.9 – Les estimations des paramètres GARCH du portefeuille 2

De plus, les valeurs du coefficient sont toutes inférieures à 1 ($\beta_1 < 1$), ce qui assure la stationnarité des séries. Les valeurs élevées du paramètre β observées ($\beta \approx 0.9$) signifient que la volatilité future est fortement affectée par la précédente. Mais il peut arriver que lors de la modélisation de certaines séries temporelles, on obtienne des valeurs du paramètre β moyennes voire faibles ($\beta \leqslant 0.7$) comme sur les actifs Automatic Data Processing Inc. (ADP), Edwards Lifesciences Corporation (EW), Molson Coors Brewing Company Class B (TAP-A) (au portefeuille 1), Assurant Inc. (AIZ), Arthur J. Gallagher & Co. (AJG), Centene Corporation (CNC) (au portefeuille 2), et Public Storage (PSA), Universal Health Services Inc. (UHS), Verisk Analytics Inc (VRSK) (au portefeuille 3). L'une des lacunes du modèle GJR-GARCH est la persistance du choc de volatilité mesurée par $\alpha_1 + \beta_1 + \gamma$, qui est proche de 1, indiquant que les chocs mettent plus de temps à se dissiper.

Assets	μ	α_0	α_1	β_1	γ	ν^1
AME	0.00069	0.00001	0.01133	0.90516	0.09528	5.39180
AON	0.00089	0.00001	0.00000	0.87421	0.13810	4.58491
APD	0.00024	0.00000	0.00151	0.92467	0.11609	4.30252
CCL	0.00087	0.00001	0.01469	0.88242	0.09360	4.59558
CHTR	0.00004	0.00004	0.01205	0.77934	0.21079	4.12671
CINF	0.00063	0.00002	0.00000	0.75991	0.19283	4.66201
COF	0.00010	0.00003	0.03095	0.76664	0.20793	5.16520
DISH	-0.00073	0.00000	0.00000	0.96241	0.06023	3.95709
DRI	0.00073	0.00004	0.01045	0.77742	0.11217	3.74940
ESS	0.00037	0.00001	0.00000	0.89753	0.10130	7.32384
EXR	0.00074	0.00002	0.00000	0.86826	0.07692	7.74976
F	-0.00012	0.00000	0.01810	0.95164	0.02726	4.33475
GM	0.00016	0.00000	0.01379	0.95606	0.03511	4.05038
LEN	0.00015	0.00001	0.00165	0.95039	0.05776	6.31373
LUV	0.00062	0.00005	0.00569	0.81240	0.09033	3.76524
MAA	0.00060	0.00000	0.00846	0.95715	0.04800	5.32081
MAR	0.00059	0.00002	0.01833	0.80961	0.18770	4.66354
META	0.00116	0.00001	0.02074	0.91397	0.08761	3.72174
MNST	0.00070	0.00001	0.00691	0.92364	0.08648	3.23674
MTB	0.00030	0.00002	0.03331	0.76660	0.22059	4.96731
NFLX	0.00109	0.00002	0.00922	0.92895	0.10969	3.22768
NRG	0.00045	0.00001	0.00248	0.95578	0.05843	5.64946
NWSA	0.00050	0.00000	0.02181	0.97977	-0.00514	3.71220
ORLY	0.00072	0.00001	0.00703	0.89404	0.09724	3.56664
PCG	0.00073	0.00000	0.00001	0.99109	0.01580	2.91120
PKG	0.00072	0.00001	0.00000	0.92875	0.06619	3.65334
PM	0.00035	0.00000	0.00001	0.99208	0.00826	3.80680
PSA	0.00048	0.00005	0.06674	0.52965	0.12235	6.22641
RCL	0.00036	0.00001	0.00000	0.93508	0.08883	4.06688
SPG	0.00039	0.00002	0.00000	0.80387	0.09730	6.52996
T	0.00026	0.00000	0.00880	0.98019	0.01192	4.33884
TDG	0.00105	0.00002	0.00000	0.83240	0.17556	3.38670
UDR	0.00050	0.00000	0.01800	0.93642	0.05572	6.89418
UHS	0.00048	0.00008	0.01155	0.60977	0.21286	3.69586
VNO	-0.00017	0.00002	0.02269	0.82173	0.09056	7.63658
VRSK	0.00086	0.00003	0.26217	0.47395	0.32360	3.29862
WFC	-0.00022	0.00002	0.03925	0.76112	0.20033	5.62387
WYNN	-0.00024	0.00000	0.00639	0.97379	0.03765	3.27987
ZBH	0.00039	0.00001	0.00000	0.87494	0.09622	3.58628

[1] conditionnal distribution parameter

TABLE 2.10 – Les estimations des paramètres GARCH du portefeuille 3

Par conséquent, il serait logique de recourir au modèle Regime Switching-GARCH (RS GARCH) qui est beaucoup moins rigide. Dans le sens qu'Il nous donne la possibilité de modéliser l'évolution de la persistance de la volatilité tout en étant capable d'anticiper un éventuel changement de régime dans le processus de volatilité. Comme évoqué à la sous-section 2.2.1, nous avons fait le choix pour le modèle RS-GARCH avec des régimes hétérogènes GARCH(1,1)-GJR-GARCH(1,1) pour les régimes à volatilité faible et élevée respectivement. Le régime 1 correspond à la période de faible volatilité est modélisé avec un simple modèle GARCH(1,1) parce que nous estimons qu'il ne nécessite pas d'un modèle trop complexe.

Assets	$\alpha_{0,1}$	$\alpha_{1,1}$	$\beta_{1,1}$	ν_1[1]	$\alpha_{0,2}$	$\alpha_{1,2}$	γ_2	β_2	ν_2[2]	$p_{1,1}$	$p_{2,1}$	π_1	π_2	δ_1	δ_2
A	0.00000	0.02367	0.96786	4.40710	0.00029	0.00000	0.26674	0.00019	99.74813	0.98880	0.02153	0.6579	0.3421	1.0113	1.0220
ADBE	0.00003	0.12313	0.93277	4.16639	0.00044	0.00005	0.00010	0.57688	10.25922	0.98539	0.22486	0.9390	0.0610	1.0148	1.2901
ADP	0.05301	0.16557	0.73831	5.36047	0.00077	0.00005	0.00010	0.40444	95.66787	0.98421	0.79624	0.9806	0.0194	1.0160	4.9077
ADSK	0.00002	0.19408	0.87000	7.17345	0.00051	0.00290	0.00021	0.57191	2.20966	0.96885	0.06598	0.6793	0.3207	1.0321	1.0706
AES	0.00000	0.03068	0.96756	91.16627	0.00056	0.00020	0.00012	0.19568	16.39078	0.94181	0.22550	0.7949	0.2051	1.0618	1.2912
AMAT	0.00011	0.11089	0.87056	99.27470	0.00167	0.00000	0.00108	0.01708	99.89184	0.93293	0.75743	0.9187	0.0813	1.0719	4.1226
BALL	0.00003	0.01245	0.99179	3.21785	0.00001	0.00012	0.07603	0.94095	6.79357	0.99567	0.00178	0.2918	0.7082	1.0043	1.0018
BBY	0.00000	0.01293	0.92928	11.86625	0.00021	0.00005	0.00010	0.87658	2.90228	0.94312	0.21696	0.7923	0.2077	1.0603	1.2771
BIIB	0.00001	0.01054	0.98926	98.65051	0.00001	0.00001	0.01343	0.99143	2.11199	0.94486	0.04609	0.4553	0.5447	1.0584	1.0483
BLK	0.00006	0.05697	0.16252	5.59764	0.00003	0.00000	0.15156	0.84387	6.97661	0.99140	0.00901	0.5116	0.4884	1.0087	1.0091
CBRE	0.00005	0.07996	0.92360	8.71148	0.00119	0.00002	0.39899	0.00009	99.64069	0.96814	0.58333	0.9482	0.0518	1.0329	2.4000
DHR	0.00000	0.11675	0.91077	13.86033	0.00049	0.00093	0.00084	0.11724	99.61059	0.94354	0.99945	0.9465	0.0535	1.0598	1827.8194
ECL	0.02852	0.01967	0.93706	99.82761	0.00002	0.00002	0.43433	0.76519	4.55082	0.63441	0.46276	0.5587	0.4413	1.5763	1.8614
EW	0.01215	0.01055	0.67174	3.13494	0.00004	0.00044	0.00010	0.93476	85.38319	0.99041	0.04583	0.8270	0.1730	1.0097	1.0480
FBHS	0.00014	0.02934	0.97210	97.71569	0.00001	0.04846	0.00299	0.94253	4.80539	0.81603	0.20220	0.5236	0.4764	1.2254	1.2534
FIS	0.00002	0.11413	0.92210	98.00981	0.00012	0.01372	0.66233	0.34342	2.62382	0.95256	0.05247	0.5251	0.4749	1.0498	1.0554
GILD	0.01776	0.00010	0.97805	58.15092	0.00013	0.04329	0.00024	0.84491	98.64016	0.92814	0.25361	0.7792	0.2208	1.0774	1.3398
GIS	0.00055	0.00731	0.98869	99.82925	0.00016	0.00515	0.11531	0.35549	5.64588	0.90086	0.13896	0.5836	0.4164	1.1100	1.1614
GWW	0.00748	0.00416	0.97488	8.23657	0.00057	0.00259	0.00113	0.84065	2.31134	0.92819	0.80407	0.9180	0.0820	1.0774	5.1037
HAS	0.00000	0.04876	0.96313	99.92692	0.00047	0.78456	0.00110	0.16755	2.16489	0.88358	0.08632	0.4258	0.5742	1.1318	1.0945
HES	0.00001	0.03094	0.96780	89.08807	0.00004	0.00002	0.29456	0.85086	10.06921	0.94749	0.14726	0.7371	0.2629	1.0554	1.1727
JCI	0.00441	0.17367	0.87420	98.74158	0.00023	0.00005	0.06956	0.03640	3.84249	0.79134	0.13608	0.3947	0.6053	1.2637	1.1575
K	0.00104	0.00010	0.91995	98.80789	0.00003	0.00010	0.00010	0.93910	95.28183	0.82767	0.35796	0.6750	0.3250	1.2082	1.5575
LNT	0.00007	0.01386	0.97015	91.98318	0.00010	0.09373	0.00837	0.17832	6.28908	0.99225	0.00174	0.1834	0.8166	1.0078	1.0017
MMM	0.00001	0.10171	0.92779	5.35625	0.00024	0.00042	0.02872	0.00003	3.20309	0.96818	0.11024	0.7760	0.2240	1.0329	1.1239
MSFT	0.01482	0.23105	0.82140	5.42096	0.00000	0.00005	0.01275	0.99299	95.14783	0.96910	0.99865	0.9700	0.0300	1.0319	743.4391
NEM	0.01855	0.00010	0.95092	8.78091	0.00106	0.01502	0.00016	0.01494	6.63393	0.96519	0.06143	0.6383	0.3617	1.0361	1.0655
NVDA	0.00002	0.18040	0.76610	5.80064	0.00297	0.00008	0.15234	0.00005	3.43825	0.93926	0.45272	0.8817	0.1183	1.0647	1.8272
TAP-A	0.47178	0.16336	0.00015	99.86192	0.00003	0.81831	0.33405	0.01411	99.89696	0.89568	0.04925	0.3207	0.6793	1.1165	1.0518
TXN	0.00000	0.04154	0.96989	4.16961	0.00022	0.00007	0.00011	0.65506	10.59762	0.99319	0.07925	0.9209	0.0791	1.0069	1.0861
UNH	0.00003	0.15338	0.71266	28.13514	0.00000	0.00831	0.07626	0.95077	4.39596	0.98930	0.00500	0.3183	0.6817	1.0108	1.0050
URI	0.00138	0.04016	0.93480	98.10580	0.00056	0.00005	0.00011	0.82708	4.05573	0.91982	0.79710	0.9086	0.0914	1.0872	4.9286
VFC	0.03524	0.26442	0.65025	94.80371	0.00000	0.00000	0.01232	0.98408	3.17915	0.86292	0.12016	0.4671	0.5329	1.1589	1.1366
VTR	0.00001	0.08852	0.91649	5.01363	0.00004	0.00001	0.02329	0.87276	2.51803	0.99412	0.02314	0.7974	0.2026	1.0059	1.0237
WAT	0.00185	0.01822	0.95925	2.77288	0.00004	0.00008	0.32969	0.62861	16.44557	0.98189	0.01547	0.4608	0.5392	1.0184	1.0157
WELL	0.00856	0.05269	0.95802	5.07160	0.00011	0.50941	0.00010	0.03165	5.64329	0.99388	0.02120	0.7759	0.2241	1.0062	1.0217
XEL	0.14049	0.00068	0.62305	96.11305	0.00001	0.00029	0.04842	0.88039	7.91285	0.99396	0.00130	0.1766	0.8234	1.0061	1.0013
XYL	0.00001	0.04910	0.95938	14.46464	0.00025	0.00001	0.00011	0.57054	13.92896	0.95818	0.37482	0.8996	0.1004	1.0436	1.5995
ZTS	0.00122	0.09045	0.94812	6.03462	0.00016	0.29561	0.00488	0.12925	4.20348	0.95760	0.05536	0.5663	0.4337	1.0443	1.0586

[1] conditionnal distribution parameter of regime 1 (std)

[2] conditionnal distribution parameter of regime 2 (std)

TABLE 2.11 – Les estimations des paramètres GARCH avec changement de régime du portefeuille 1

Alors que le régime 2 correspondant à la période de forte volatilité quant à lui est modélisé avec un modèle GJR-GARCH(1,1) afin de pouvoir capturer les caractéristiques spécifiques des séries temporelles financières en période de fortes agitations sur les marchés boursiers. Tous les paramètres estimés sur la première fenêtre glissante sont enregistrés aux tableaux 2.11, 2.12 et 2.13. Nous y avons ajouté certaines informations utiles comme les probabilités inconditionnelles (π_1 et π_2) et les espérances des durées (δ_1 et δ_2) pour chaque régime. Elles ont été obtenues en utilisant les équations suivantes :

$$
\begin{aligned}
p_{2,2} &= 1 - p_{2,1} \\
\pi_j &= \frac{1 - p_{i,i}}{2 - p_{i,i} - p_{j,j}} \; \forall \; j = 1,2, \; \forall \; i = 1,2, \text{ and } i \neq j \\
\delta_j &= \frac{1}{1 - p_{j,j}} \; \forall \; j = 1,2.
\end{aligned}
$$

Les probabilités de rester dans les régimes à faible et forte volatilité, $p_{1,1}$ et $p_{2,2} = 1 - p_{2,1}$ parfois très élevées (≈ 0.99) ou très faibles ($\leqslant 0.01$) montrent la persistance dans chaque régime.

Assets	$\alpha_{0,1}$	$\alpha_{1,1}$	$\beta_{1,1}$	ν_1[1]	$\alpha_{0,2}$	$\alpha_{1,2}$	γ_2	β_2	ν_2[2]	$p_{1,1}$	$p_{2,1}$	π_1	π_2	δ_1	δ_2
AAPL	0.00456	0.17065	0.87948	5.79928	0.00060	0.00482	0.00011	0.00073	97.73651	0.92893	0.67118	0.9043	0.0957	1.0765	3.0412
ABC	0.00227	0.00484	0.93531	98.82616	0.00001	0.00008	0.07221	0.95040	3.98722	0.82543	0.05858	0.2512	0.7488	1.2115	1.0622
ABT	0.00001	0.15415	0.87358	5.44909	0.00028	0.10248	0.00012	0.00053	4.07461	0.98710	0.0460	0.7809	0.2191	1.0131	1.0482
AIZ	0.00202	0.93789	0.24167	74.74489	0.00024	0.00003	0.00010	0.01970	3.65447	0.85765	0.08027	0.3606	0.6394	1.1660	1.0873
AJG	0.04125	0.22520	0.66043	12.25466	0.00098	0.07355	0.03316	0.00024	99.74299	0.98921	0.92148	0.9884	0.0116	1.0109	12.7355
AKAM	0.00046	0.40650	0.65373	10.14549	0.00227	0.04237	0.00015	0.00082	2.10115	0.92800	0.09480	0.5683	0.4317	1.0776	1.1047
ALL	0.00908	0.00010	0.57112	99.53693	0.00000	0.00003	0.05444	0.96703	3.39673	0.90372	0.02130	0.1812	0.8188	1.1065	1.0218
AMD	0.00000	0.00641	0.99205	18.30965	0.00275	0.13641	0.00010	0.00036	4.90186	0.90270	0.13736	0.5854	0.4146	1.1078	1.1592
AMP	0.00003	0.41831	0.68354	97.92548	0.00018	0.14016	0.00045	0.36594	4.50058	0.83037	0.15494	0.4774	0.5226	1.2043	1.1833
AMZN	0.00548	0.25066	0.81599	6.83448	0.00490	0.00242	0.00052	0.10767	99.52549	0.98621	0.90307	0.9850	0.0150	1.0140	10.3169
AVY	0.00024	0.10620	0.90865	99.99999	0.00007	0.02333	0.69244	0.61064	2.45612	0.22486	0.92735	0.5447	0.4553	4.4472	13.7647
AWK	0.00213	0.03619	0.89814	99.45790	0.00000	0.02849	0.00010	0.96088	5.71076	0.55971	0.43868	0.4991	0.5009	1.7866	1.7815
AZO	0.00001	0.00791	0.98106	9.28159	0.00050	0.00032	0.07291	0.04458	3.44770	0.93067	0.20831	0.7503	0.2497	1.0745	1.2631
CF	0.00064	0.00010	0.05332	99.52287	0.00078	0.08516	0.00015	0.00276	7.32119	0.97868	0.00916	0.3004	0.6996	1.0218	1.0092
CMG	0.00000	0.00274	0.99335	99.71589	0.00002	0.00002	0.04388	0.96259	3.46255	0.82210	0.13467	0.4308	0.5692	1.2164	1.1556
CNC	0.00006	0.55890	0.57790	10.39101	0.00038	0.00057	0.34055	0.23084	3.55239	0.91903	0.09277	0.5340	0.4660	1.0881	1.1023
COP	0.32638	0.00026	0.00126	3.11175	0.00001	0.00000	0.09547	0.94228	98.08651	0.94415	0.02024	0.2660	0.7340	1.0592	1.0207
DVN	0.00000	0.09689	0.94525	2.84460	0.00001	0.00007	0.07268	0.95303	94.11122	0.85101	0.03910	0.2079	0.7921	1.1751	1.0407
ETR	0.00151	0.00010	0.15302	78.56778	0.00001	0.00012	0.04630	0.92592	6.74593	0.98749	0.00149	0.1067	0.8933	1.0127	1.0015
FCX	0.01380	0.01685	0.96542	5.60802	0.00023	0.02368	0.00032	0.93549	14.25649	0.99694	0.03019	0.9078	0.0922	1.0031	1.0311
FISV	0.00006	0.11977	0.89458	20.70676	0.00032	0.00021	0.04212	0.41234	99.10722	0.95822	0.72209	0.9453	0.0547	1.0436	3.5983
FMC	0.00451	0.07205	0.95403	99.99944	0.00073	0.01609	1.96724	0.00000	2.23117	0.62468	0.81995	0.6860	0.3140	1.6008	5.5541
GOOGL	0.00472	0.10891	0.92343	4.57345	0.00081	0.00001	0.94099	0.33783	2.59072	0.99219	0.21168	0.9644	0.0356	1.0079	1.2685
GS	0.01636	0.00010	0.93963	19.61097	0.00026	0.00005	0.00019	0.53918	6.11776	0.95732	0.14173	0.7686	0.2314	1.0446	1.1651
HLT	0.00001	0.32932	0.67040	10.09095	0.00001	0.00065	0.08562	0.92907	6.78164	0.98920	0.00824	0.4327	0.5673	1.0109	1.0083
IRM	0.00000	0.01701	0.96673	98.00254	0.00046	0.00015	0.16233	0.06412	99.80152	0.83750	0.50447	0.7564	0.2436	1.1940	2.0180
MRO	0.00026	0.00017	0.63528	4.01214	0.00001	0.00000	0.06520	0.95541	18.66925	0.99038	0.00635	0.3974	0.6026	1.0097	1.0064
OXY	0.08575	0.00010	0.45252	92.60255	0.00001	0.00034	0.03136	0.94862	29.88382	0.94093	0.06420	0.5208	0.4792	1.0628	1.0686
TMO	0.00028	0.00010	0.62019	18.42268	0.00028	0.00018	0.26127	0.00007	14.02634	0.92274	0.10462	0.5752	0.4248	1.0837	1.1168
UNP	0.00001	0.05027	0.95654	99.67595	0.00059	0.00000	0.43501	0.00002	99.84313	0.90536	0.55611	0.8546	0.1454	1.1045	2.2528
V	0.03981	0.00027	0.93940	4.99551	0.00003	0.00386	0.55099	0.63513	10.22353	0.82911	0.14458	0.4583	0.5417	1.2061	1.1690
VLO	0.05277	0.00010	0.83778	22.93316	0.00009	0.02569	0.01128	0.77236	7.14974	0.88867	0.08997	0.4469	0.5531	1.1253	1.0989
VMC	0.00037	0.00010	0.52897	5.51876	0.00002	0.01900	0.20452	0.84398	5.70555	0.99476	0.00188	0.2637	0.7363	1.0053	1.0019
VRSN	0.01636	0.00010	0.93963	19.61097	0.00026	0.00005	0.00019	0.53918	6.11776	0.95732	0.14173	0.7686	0.2314	1.0446	1.1651
VRTX	0.00161	0.10044	0.93280	9.99056	0.00078	0.00120	0.00011	0.80759	2.32913	0.95570	0.51872	0.9213	0.0787	1.0464	2.0778
WM	0.00000	0.10563	0.93742	99.66003	0.00000	0.00001	0.00562	0.99357	2.48893	0.87697	0.10023	0.4489	0.5511	1.1403	1.1114
WMT	0.13741	0.00040	0.06752	2.99088	0.00008	0.05157	0.00027	0.55936	99.76162	0.98566	0.03536	0.7115	0.2885	1.0145	1.0367
XOM	0.26051	0.00144	0.24305	4.94602	0.00003	0.01596	0.19418	0.76976	19.66457	0.99449	0.00543	0.4965	0.5035	1.0055	1.0055
YUM	0.02031	0.13289	0.89862	3.36420	0.00063	0.00011	0.00010	0.30477	2.48952	0.99831	0.01545	0.9016	0.0984	1.0017	1.0157

[1] conditionnal distribution parameter of regime 1 (std)

[2] conditionnal distribution parameter of regime 2 (std)

TABLE 2.12 – Les estimations des paramètres GARCH avec changement de régime du portefeuille 2

Les espérances des durées δ_1 et δ_2 mettent mieux en évidence la persistance dans chaque régime. Par exemple, dans le portefeuille 1, l'action Danaher Corporation (DHR) a une durée prévue d'environ 1 jour dans le régime à faible volatilité et de 1828 jours dans le régime à forte volatilité, avec des probabilités inconditionnelles respectives de 0.965 et 0.0535. Pendant ce temps, pour l'actif Microsoft Corporation (MSFT) reste dans le régime à faible volatilité (régime 1) avec une probabilité inconditionnelle de 0,97 que 1 jour mais peut rester environ 744 jours dans le régime à forte volatilité (régime 2) avec une probabilité inconditionnelle de 0,03. Un constat analogue peut être fait avec les actions Alphabet Inc. Class A (GOOGL) et Amazon.com Inc. (AMZN) pour le portefeuille 2 et Capital One Financial Corporation (COF), Darden Restaurants Inc. (DRI) pour le portefeuille 3. Par conséquent, considérer un modèle de changement de régime avec une structure de dépendance pour le régime à faible volatilité et une autre pour celui à haute volatilité, ne semble pas approprié pour notre étude. Ces résultats ont un impact considérable sur la suite de notre analyse parce qu'ils nous donnent des preuves empiriques qui sont incompatibles avec l'hypothèse du modèle de Fink et al. (2017) pour les structures de dépendance.

Assets	$\alpha_{0,1}$	$\alpha_{1,1}$	$\beta_{1,1}$	ν_1 [1]	$\alpha_{0,2}$	$\alpha_{1,2}$	γ_2	β_2	ν_2 [2]	$p_{1,1}$	$p_{2,1}$	π_1	π_2	δ_1	δ_2
AME	0.00013	0.20846	0.79016	12.67240	0.00026	0.00045	0.00010	0.22335	4.69822	0.95326	0.20001	0.8106	0.1894	1.0490	1.2500
AON	0.00006	0.14012	0.84295	11.33140	0.00018	0.41951	0.00732	0.08892	99.42463	0.75477	0.97370	0.7988	0.2012	1.3249	38.0257
APD	0.00567	0.05520	0.93249	29.90732	0.00000	0.00008	0.01003	0.99265	97.92875	0.92949	0.48697	0.8735	0.1265	1.0759	1.9492
CCL	0.00002	0.11951	0.88136	99.83714	0.00004	0.00000	0.12205	0.78726	4.77744	0.98921	0.00216	0.1667	0.8333	1.0109	1.0022
CHTR	0.10384	0.00061	0.58469	12.95143	0.00004	0.00007	0.13651	0.83066	4.70531	0.86127	0.01784	0.1140	0.8860	1.1611	1.0182
CINF	0.00679	0.10672	0.79557	97.38944	0.00019	0.00005	0.57722	0.52754	99.13555	0.88929	0.74445	0.8705	0.1295	1.1245	3.9131
COF	0.02433	0.18100	0.77746	11.40079	0.00346	0.01418	0.00040	0.06844	97.74473	0.98712	0.99996	0.9873	0.0127	1.0130	25125.6281
DISH	0.00086	0.04350	0.97715	2.56990	0.00018	0.04963	0.19385	0.51740	97.44751	0.85236	0.20031	0.5757	0.4243	1.1732	1.2505
DRI	0.00613	0.05221	0.84938	99.70468	0.00024	0.51019	0.93519	0.02212	5.22514	0.33256	0.87152	0.5663	0.4337	3.0069	7.7832
ESS	0.00001	0.04060	0.95371	5.18578	0.00003	0.00003	0.18416	0.74046	99.95197	0.98781	0.01466	0.5460	0.4540	1.0123	1.0149
EXR	0.00011	0.21190	0.35414	4.57332	0.00001	0.00000	0.12122	0.87351	100.00000	0.98854	0.01209	0.5135	0.4865	1.0116	1.0122
F	0.00974	0.01272	0.96477	5.32308	0.00059	0.00010	0.47109	0.00002	97.91793	0.95579	0.37099	0.8935	0.1065	1.0463	1.5898
GM	0.00334	0.00914	0.97347	99.06174	0.00045	0.00129	0.11958	0.00332	9.18327	0.73117	0.31174	0.5370	0.4630	1.3677	1.4529
LEN	0.00000	0.03640	0.97253	5.05646	0.00012	0.00033	0.23198	0.61852	93.16691	0.98904	0.05213	0.8263	0.1737	1.0111	1.0550
LUV	0.00771	0.00013	0.97429	97.90865	0.00067	0.00016	0.00010	0.52125	98.57608	0.86924	0.36687	0.7372	0.2628	1.1504	1.5795
MAA	0.00000	0.59972	0.00004	99.53784	0.00011	0.00005	0.05624	0.38001	9.66712	0.87770	0.06868	0.3596	0.6404	1.1393	1.0737
MAR	0.00001	0.20959	0.70056	4.30795	0.00019	0.00003	0.00012	0.58571	98.07654	0.98988	0.03943	0.7958	0.2042	1.0102	1.0411
META	0.00242	0.24184	0.60464	8.75828	0.00000	0.00016	0.07380	0.96236	3.25399	0.91905	0.08249	0.5047	0.4953	1.0881	1.0899
MNST	0.00000	0.10126	0.94846	8.31315	0.00032	0.05300	0.60597	0.63210	2.17142	0.95815	0.03325	0.4427	0.5573	1.0437	1.0344
MTB	0.01951	0.18044	0.78285	96.61951	0.00037	0.12720	0.00103	0.05160	99.68493	0.83531	0.64868	0.7975	0.2025	1.1972	2.8464
NFLX	0.02862	0.00026	0.93885	8.06254	0.00001	0.00000	0.03730	0.97634	2.46417	0.99108	0.00544	0.3787	0.6213	1.0090	1.0055
NRG	0.00001	0.04648	0.95925	15.41639	0.00101	0.00090	0.13204	0.79291	3.32342	0.98155	0.40562	0.9565	0.0435	1.0188	1.6824
NWSA	0.00523	0.00010	0.97603	99.24759	0.00028	0.00094	0.05452	0.41453	4.92042	0.90146	0.12648	0.5621	0.4379	1.1093	1.1448
ORLY	0.00026	0.19837	0.84395	33.17710	0.00000	0.00009	0.00522	0.99700	3.73872	0.78983	0.12111	0.3656	0.6344	1.2661	1.1378
PCG	0.01413	0.03315	0.96735	9.37597	0.00000	0.00201	0.01000	0.99270	3.10483	0.98835	0.33703	0.9666	0.0334	1.0118	1.5084
PKG	0.00006	0.00982	0.98690	86.48766	0.00003	0.00002	0.17179	0.86074	3.72145	0.62163	0.20261	0.3487	0.6513	1.6087	1.2541
PM	0.00000	0.00389	0.99512	5.10391	0.00018	0.00010	0.71749	0.15613	4.10618	0.98459	0.06146	0.7996	0.2004	1.0156	1.0655
PSA	0.00066	0.15467	0.61910	98.12459	0.00024	0.00053	0.00010	0.48519	99.68602	0.88797	0.99431	0.8987	0.1013	1.1262	175.7130
RCL	0.00001	0.14238	0.91983	5.64468	0.00044	0.02555	0.00020	0.01284	4.16093	0.43738	0.87346	0.6082	0.3918	2.2864	7.9028
SPG	0.00042	0.00022	0.26199	99.98860	0.00006	0.00018	0.00014	0.76534	99.84500	0.84914	0.21463	0.5872	0.4128	1.1777	1.2733
T	0.00862	0.01133	0.98298	3.97429	0.00000	0.00005	0.00185	0.99846	75.39618	0.97833	0.16139	0.8816	0.1184	1.0222	1.1924
TDG	0.00006	0.20411	0.79802	9.66089	0.00002	0.00053	0.00734	0.95354	3.12050	0.47927	0.99903	0.6574	0.3426	2.0865	1032.7378
UDR	0.00016	0.06968	0.94371	5.58263	0.00004	0.14884	0.02183	0.64104	15.92083	0.99832	0.00178	0.5152	0.4848	1.0017	1.0018
UHS	0.00002	0.02800	0.97755	99.24858	0.00011	0.00026	0.22766	0.53481	3.19315	0.97355	0.00559	0.1744	0.8256	1.0272	1.0056
VNO	0.00005	0.00073	0.58696	9.28327	0.00001	0.00001	0.08483	0.88615	10.08648	0.71035	0.04606	0.1372	0.8628	1.4078	1.0483
VRSK	0.20481	0.30861	0.50169	38.51200	0.00022	0.30835	0.00010	0.17291	2.45491	0.90456	0.13487	0.5856	0.4144	1.1055	1.1559
WFC	0.00019	0.22066	0.73863	5.77707	0.00025	0.20533	0.00023	0.08819	40.22130	0.99297	0.05125	0.8794	0.1206	1.0071	1.0540
WYNN	0.00001	0.01955	0.98950	2.87359	0.00038	0.00006	0.00010	0.79927	5.20686	0.99584	0.01382	0.7685	0.2315	1.0042	1.0140
ZBH	0.00000	0.16408	0.82358	99.95420	0.00002	0.00000	0.00247	0.94625	3.24187	0.79953	0.19300	0.4905	0.5095	1.2507	1.2392

[1] conditionnal distribution parameter of regime 1 (std)

[2] conditionnal distribution parameter of regime 2 (std)

TABLE 2.13 – Les estimations des paramètres GARCH avec changement de régime du portefeuille 3

La spécification du modèle Markov Switching R-vine (MS-RV) de Fink et al. (2017) est donnée par deux copules de vine (R-vine) et une matrice de transition 2×2 décrivant la chaîne de Markov sous-jacente. Ils font donc l'hypothèse qu'il y aurait une partie de la structure de dépendance pour le régime de faible volatilité et une autre pour celui de forte volatilité. Sauf que cette hypothèse ne semble pas adaptée à notre problématique dans la mesure où les actifs qui composent les portefeuilles ne restent pas le même nombre de jours dans chaque régime. C'est la raison pour laquelle nous introduirons plutôt la notion de changement de régime dans la modélisation des distributions marginales et pas pour la structure de dépendance. Même si recourir aux copules permet de modéliser indépendamment les distributions marginales et la structure de dépendance, l'incorporation d'un changement de régime dans les marginales peut indirectement influencer la dépendance conditionnelle (structure de dépendance).

Portfolio 1		
C-DCC(1,1) parameters		
a	b	η
0.0049	0.8778	7.4271

Portfolio 2		
C-DCC(1,1) parameters		
a	b	η
0.0064	0.8138	7.5816

Portfolio 3		
C-DCC(1,1) parameters		
a	b	η
0.0049	0.8933	7.3855

Portfolio Benchmark		
C-DCC(1,1) parameters		
a	b	η
1.15936e-10	0.97808	49.99987

TABLE 2.14 – Les estimations des paramètres du modèle copule-GARCH

Les différents modèles envisagés pour modéliser les dépendances sont :

- le DCC-GARCH
- le C-GARCH (copula- GARCH)
- le Pair Copula Construction (PCC)
- le Regime Switching Pair Copula Construction (RS PCC)

Nous pouvons trouver de nombreuses traces dans la littérature utilisant les modèles DCC-GARCH raison pour laquelle nous nous concentrerons sur les trois autres. Les paramètres η, a, et b de l'équation (2.23) des 4 différents portefeuilles sont enregistrés dans la table 2.14. Les valeurs élevées de b ($0.7 \leqslant b \leqslant 0.99$) et les valeurs basses de a ($a \approx 0$) indiquent respectivement que la corrélation récente a une forte influence sur la corrélation future, tandis que le choc a une faible influence sur la corrélation. Le paramètre η qui représente le degré de liberté et prend les valeurs 7.42, 7.58 et 7.35 pour les portefeuilles 1, 2 et 3 respectivement. Pour ces trois portefeuilles, les valeurs du paramètre η sont supérieures à 4, ce qui indique une probabilité relativement faible d'occurrence d'événements extrêmes. Le portefeuille Benchmark se trouve dans la même situation mais dans ce cas-ci, la valeur du paramètre η est très largement supérieure à 4. Cela signifie que cette distribution présente des queues plus fines et se rapproche de celle d'une distribution normale multivariée.

La distribution normale multivariée et la copule normale sont liées par le concept de dépendance et de corrélation. Donc d'un certain point de vue, nous pouvons dire qu'avec une valeur du paramètre η aussi élevée, la copule de Student modélisée converge en distribution vers une copule gaussienne. Parce que les copules gaussiennes présentent certaines limites raison pour laquelle nous avons fait le choix de ne pas la considérer dans notre analyse, constater que la copule de Student que nous souhaitons modéliser converge en distribution vers une copule gaussienne n'est pas problématique. Parce que des chercheurs comme Genest, Rémillard, and Beaudoin (2009) ont mis en évidence le fait que les copules gaussiennes ont de meilleures performances lorsque le nombre de dimensions est élevé comme dans le cas de notre portefeuille Benchmark avec 117 actifs. Ensuite, compte tenu de la rigidité de la copule de Student car comme nous avons pu le constater, elle ne dispose que d'un seul paramètre pour contrôler les queues de distributions entre toutes les paires de variables. Cette quête de plus de flexibilité a motivé notre choix pour l'utilisation des copules de vines (C-vine) pour la modélisation de la structure de dépendance des portefeuilles avec des distributions marginales qui devraient être générés par des processus GJR-GARCH ou SR-GJR-GARCH.

	PCC					RS PCC				
	Copula	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio Benchmark	Copula	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio Benchmark
Negative tail	2 : Student (t)	108	108	121	686	2 : Student (t)	94	110	121	662
	3 : Clayton (C)	35	47	37	520	3 : Clayton (C)	39	0	32	525
	7 : Clayton-Gumbel (BB1)	1	5	0	25	7 : Clayton-Gumbel (BB1)	5	2	1	27
	9 : Joe-Clayton (BB7)	3	0	0	15	9 : Joe-Clayton (BB7)	1	0	1	18
	14 : Gumbel 180 (SG)	29	27	22	165	14 : Gumbel 180 (SG)	18	21	19	184
	16 : Joe 180 (SJ)	7	18	17	177	16 : Joe 180 (SJ)	11	8	20	180
	20 : Joe-Frank 180 (SBB8)	35	40	45	191	20 : Joe-Frank 180 (SBB8)	18	17	7	163
Centre	1 : Gaussian (N)	81	72	74	857	1 : Gaussian (N)	80	82	75	851
	5 : Frank (F)	243	236	230	1960	5 : Frank (F)	247	231	233	1949
	23 : Clayton 90 (C90)	38	38	49	587	23 : Clayton 90 (C90)	44	37	40	612
	33 : Clayton 270 (C270)	48	41	44	593	33 : Clayton 270 (C270)	61	41	46	584
Positive tail	2 : Student (t)	108	108	121	686	2 : Student (t)	94	110	121	662
	4 : Gumbel (G)	22	32	29	223	4 : Gumbel (G)	22	34	34	226
	6 : Joe (J)	16	18	24	206	6 : Joe (J)	17	25	20	222
	7 : Clayton-Gumbel (BB1)	1	5	0	25	7 : Clayton-Gumbel (BB1)	5	2	1	27
	9 : Joe-Clayton (BB7)	3	0	0	15	9 : Joe-Clayton (BB7)	1	0	1	18
	10 : Joe-Frank (BB8)	19	16	6	85	10 : Joe-Frank (BB8)	18	17	7	82
	13 : Clayton 180 (SC)	52	46	33	496	13 : Clayton 180 (SC)	51	49	47	501

TABLE 2.15 – Resumé de la sélection de copules bivariées pour le modèle C-vine avec des marginales GJR-GARCH et RS GJR-GARCH

Les temps de calcul exagérément long précédemment évoqués sont relatifs aux nombres de paramètres à estimer pour la modélisation de la structure de dépendance avec le modèle DCC copula et surtout les modèles PCC GARCH et PCC SR GARCH. Cependant, l'utilisation de la copule de vines présentent certains inconvénients. Le premier problème est le temps de calcul exagérément long qui s'explique par le nombre de combinaisons possibles de copules bivariées ainsi que le nombre de paramètres associés à chacun d'eux. Pour un portefeuille composé de n actifs, le nombre de combinaisons possibles de copules bivariées est $\frac{n!}{2} \times 2 \times \binom{n-2}{2}$, c'est qu'une fois toutes ces combinaisons estimées que l'algorithme sélectionne la meilleure selon un critère d'information (AIC ou BIC). Bien que des techniques de calcul comme la parallélisation nous a aidé à réduire les temps de calcul, ils restent tout de même non négligeables. Le second problème de l'utilisation de la copule de vines est qu'elle génère beaucoup d'informations. Dans cette grande quantité d'informations, nous avons l'ensemble de copules bivariées utilisées et leurs paramètres, les matrices des valeurs empiriques du τ de Kendall ce qui le rend difficile sa synthèse. Pour cela nous nous référons à la *copula counting technique* de Hernandez (2015), qui nous permet de résumer à la table 2.15 la plupart des informations disponibles pour l'interprétation.

En additionnant le nombre de fois que les différentes copules bivariées qui ont été utilisées pour la modélisation de la dépendance au centre et dans les queues (inférieure et supérieure) des distributions de chaque portefeuille, nous constatons que la dépendance est majoritairement concentrée au centre de la distribution conjointe avec une dépendance de queue inférieure plus élevée que dépendance de queue supérieure. Un constat que nous pouvons vérifier au regard des résultats obtenus avec le modèle PCC (avec des marginales GJR-GARCH) pour les portefeuilles 1, 2, 3 et même pour le portefeuille Benchmark. Par exemple au portefeuille 1, les copules bivariées requises pour la modélisation de la dépendance de queue inférieure ont été utilisées 211 fois contre 201 fois pour celles de la dépendance de queue supérieure. Pour la modélisation de la dépendance au centre de la distribution conjointe, les copules bivariées ayant les propriétés de le faire ont été utilisées 420 fois. Un constat similaire peut être fait pour les 3 autres portefeuilles en utilisant les valeurs correspondantes.

Models	Short selling	Annualized Return	Annualized STD[1]	Historical VaR 5%	CVaR 5%	Skewness	Kurtosis	CF[2]CVaR 5%	MDD[3]	Annualized SR[4]	Omega Ratio
equally weighted		**0.09187**	**0.23266**	**-0.02224**	**-0.03643**	**-1.00576**	**15.59891**	**-0.05571**	**37.65920**	**0.39487**	**1.10345**
mean-var	F	0.07338	0.19684	-0.01697	-0.02996	-0.70306	18.95215	-0.02356	34.00895	0.37281	1.09868
	T	0.01794	0.20713	-0.01624	-0.03140	-0.95696	23.94140	-0.01832	38.77582	0.08659	1.04084
mean-var t Cop GARCH	F	0.00338	0.27084	-0.02567	-0.04329	0.47566	12.56480	-0.02119	39.89076	0.01248	1.03389
	T	-0.06741	0.35804	-0.03283	-0.05843	0.95713	17.85398	-0.02246	44.89314	-0.18827	0.99554
mean-var DCC-GARCH	F	-0.00861	0.28267	-0.02605	-0.04563	0.55139	12.66345	-0.02171	40.02801	-0.03044	1.02568
	T	0.03122	0.27825	-0.02191	-0.04832	-1.04080	11.64764	-0.07355	34.90672	0.11221	1.05990
mean-var PCC-GARCH	F	0.06766	0.26994	-0.02082	-0.03653	-1.86717	62.00209	-0.01419	37.15267	0.25063	1.09305
	T	**0.15002**	**0.31971**	**-0.02331**	**-0.04292**	**-0.31183**	**45.30035**	**-0.01569**	**40.98406**	**0.46925**	**1.16089**
mean-var RS PCC-GARCH	F	0.05809	0.20132	-0.01759	-0.03097	-0.66714	18.28959	-0.02509	34.03050	0.28854	1.08129
	T	0.01043	0.21198	-0.01712	-0.03261	-0.88637	21.90235	-0.02175	38.81296	0.04922	1.03292
mean-$LPM_{\tau=0,a=1}$	F	0.05880	0.20315	-0.01602	-0.03057	-0.86133	21.76486	-0.01990	35.84243	0.28943	1.08416
	T	0.01986	0.20307	-0.01583	-0.03044	-0.68488	21.45298	-0.01771	35.93036	0.09782	1.04224
mean-$LPM_{\tau=0,a=1}$ t Cop GARCH	F	0.07002	0.23191	-0.02276	-0.03588	0.15201	10.84462	-0.01981	36.40575	0.30192	1.08966
	T	-0.04656	0.36243	-0.03374	-0.05901	0.98655	16.67008	-0.02297	43.01296	-0.12846	1.01204
mean-$LPM_{\tau=0,a=1}$ DCC-GARCH	F	0.02286	0.23800	-0.01960	-0.03945	-1.72963	22.79263	-0.06790	44.23601	0.09605	1.04823
	T	0.04295	0.25343	-0.02157	-0.04296	-0.89669	11.67466	-0.06258	34.21592	0.16948	1.06665
mean-$LPM_{\tau=0,a=1}$ **PCC-GARCH**	F	0.02909	0.27628	-0.02339	-0.03897	-2.08849	54.40871	-0.01814	36.39168	0.10529	1.05830
	T	**0.08312**	**0.27752**	**-0.02425**	**-0.03975**	**-0.85630**	**33.36075**	**-0.02052**	**35.03017**	**0.29953**	**1.10225**
mean-$LPM_{\tau=0,a=1}$ RS PCC-GARCH	F	0.05664	0.20586	-0.01730	-0.03135	-0.62700	19.31416	-0.01884	34.82398	0.27512	1.08080
	T	0.02005	0.20652	-0.01714	-0.03153	-0.47718	18.30779	-0.01813	35.29160	0.09707	1.04172
mean-$LPM_{\tau=0,a=2}$	F	0.07707	0.19706	-0.01693	-0.03032	-0.90195	18.02374	-0.03677	33.93503	0.39108	1.10135
	F	0.01938	0.20822	-0.01663	-0.03168	-1.18328	21.22489	-0.04094	36.89570	0.09305	1.04126
mean-$LPM_{\tau=0,a=2}$ t Cop GARCH	F	0.03968	0.23138	-0.02234	-0.03600	0.10798	11.53724	-0.01986	36.69428	0.17151	1.06204
	T	-0.00485	0.27258	-0.02477	-0.04325	0.49341	12.42232	-0.02131	40.00172	-0.01781	1.02630
mean-$LPM_{\tau=0,a=2}$ DCC-GARCH	F	0.01255	0.24093	-0.02008	-0.04008	-1.72153	23.06046	-0.06750	44.79535	0.05211	1.03869
	T	0.06640	0.22416	-0.01919	-0.03662	-0.95906	13.65372	-0.05538	33.65439	0.29622	1.08425
mean-$LPM_{\tau=0,a=2}$ **PCC-GARCH**	F	0.05158	0.26555	-0.02253	-0.03681	-2.13880	58.27018	-0.01622	36.07080	0.19422	1.07778
	T	**0.09572**	**0.24099**	**-0.02283**	**-0.03394**	**-0.92376**	**31.57034**	**-0.01855**	**31.49148**	**0.39719**	**1.11279**
mean-$LPM_{\tau=0,a=2}$ RS PCC-GARCH	F	0.05318	0.20025	-0.01760	-0.03146	-0.91202	17.13285	-0.04068	34.76736	0.26558	1.07566
	T	0.02106	0.21523	-0.01862	-0.03319	-1.05157	19.16391	-0.04367	36.73798	0.09783	1.04232

[1] Standard Deviation
[2] Cornish-Fisher
[3] Max Drawdown
[4] Sharpe Ratio

TABLE 2.16 – Les performances des différents modèles sur le portefeuille 1

Models	Short selling	Annualized Return	Annualized STD[1]	Historical VaR 5%	CVaR 5%	Skewness	Kurtosis	CF[2]CVaR 5%	MDD[3]	Annualized SR[4]	Omega Ratio
equally weighted		0.15052	0.24769	-0.02193	-0.03943	-1.23697	16.08697	-0.06754	43.15017	0.60767	1.14699
mean-var	F	0.09956	0.20718	-0.01799	-0.03248	-0.89023	19.95410	-0.03046	38.84728	0.48055	1.12578
	T	0.04118	0.18792	-0.01524	-0.02902	-1.83141	21.74737	-0.05892	38.90198	0.21911	1.06503
mean-var t Cop GARCH	**F**	**0.14541**	**0.19334**	**-0.01655**	**-0.02988**	**-1.10780**	**18.37065**	**-0.04352**	**29.25133**	**0.75206**	**1.17387**
	T	0.11932	0.18764	-0.01505	-0.02810	-0.86892	15.20670	-0.04081	26.75110	0.63590	1.14582
mean-var DCC-GARCH	**F**	**0.14329**	**0.19384**	**-0.01655**	**-0.03007**	**-1.14114**	**18.33895**	**-0.04505**	**29.82102**	**0.73925**	**1.17135**
	T	0.12401	0.18518	-0.01565	-0.02744	-0.44543	16.00277	-0.02095	24.68714	0.66967	1.15214
mean-var PCC-GARCH	F	0.09983	0.21174	-0.01837	-0.03354	-0.81302	16.97550	-0.03875	36.12640	0.47147	1.12483
	T	0.08512	0.20781	-0.01663	-0.03244	-1.02435	20.30340	-0.03623	35.86200	0.40963	1.11121
mean-var RS PCC-GARCH	F	0.08229	0.20994	-0.01878	-0.03282	-0.87616	16.56752	-0.04263	32.69984	0.39198	1.10286
	T	0.04770	0.21385	-0.01758	-0.03428	-1.71743	19.51313	-0.06784	39.11669	0.22306	1.06920
mean-$LPM_{\tau=0,a=1}$	F	0.09789	0.20450	-0.01843	-0.03213	-0.87111	17.04841	-0.03983	37.42092	0.47867	1.12420
	T	0.05150	0.19165	-0.01557	-0.03004	-1.18193	17.77941	-0.04774	36.96312	0.26874	1.07682
mean-$LPM_{\tau=0,a=1}$ t **Cop GARCH**	**F**	**0.13836**	**0.19060**	**-0.01686**	**-0.02925**	**-1.02329**	**18.36443**	**-0.03950**	**28.09810**	**0.72591**	**1.16760**
	T	0.12442	0.18687	-0.01527	-0.02827	-0.95082	15.04346	-0.04373	26.66899	0.66579	1.15171
mean-$LPM_{\tau=0,a=1}$ DCC-GARCH	F	0.12179	0.19064	-0.01617	-0.02912	-0.76887	18.78338	-0.02657	28.78742	0.63885	1.15128
	T	0.13079	0.18536	-0.01491	-0.02754	-0.43377	15.17127	-0.02322	24.99283	0.70558	1.15873
mean-$LPM_{\tau=0,a=1}$ PCC-GARCH	F	0.10680	0.21237	-0.01812	-0.03371	-0.73471	14.99188	-0.04122	34.29037	0.50291	1.12902
	T	0.05975	0.20851	-0.01808	-0.03288	-0.79556	17.34810	-0.03622	34.93106	0.28658	1.08292
mean-$LPM_{\tau=0,a=1}$ RS PCC-GARCH	F	0.09354	0.20854	-0.01918	-0.03206	-0.68110	17.22878	-0.03088	31.66497	0.44854	1.11592
	T	0.06730	0.22265	-0.01905	-0.03570	-1.39706	17.97135	-0.06320	38.26873	0.30224	1.08658
mean-$LPM_{\tau=0,a=2}$	F	0.08299	0.20281	-0.01760	-0.03182	-1.12919	18.35575	-0.04683	38.60073	0.40921	1.10793
	T	-0.00248	0.18742	-0.01620	-0.02955	-2.03749	21.96015	-0.06332	39.86733	-0.01322	1.01664
mean-$LPM_{\tau=0,a=2}$ t **Cop GARCH**	**F**	**0.15937**	**0.19269**	**-0.01642**	**-0.02946**	**-0.94343**	**18.73840**	**-0.03517**	**27.87765**	**0.82708**	**1.18974**
	T	0.12831	0.18878	-0.01584	-0.02800	-0.69080	15.02309	-0.03476	26.04374	0.67969	1.15366
mean-$LPM_{\tau=0,a=2}$ DCC-GARCH	F	0.13581	0.19176	-0.01600	-0.02930	-0.75801	19.04837	-0.02509	28.09691	0.70826	1.16639
	T	0.12759	0.18790	-0.01581	-0.02792	-0.46307	16.03481	-0.02194	25.02658	0.67899	1.15495
mean-$LPM_{\tau=0,a=2}$ PCC-GARCH	F	0.11899	0.21451	-0.01858	-0.03382	-0.55151	17.37448	-0.02421	35.58143	0.55471	1.14412
	T	0.12009	0.20883	-0.01650	-0.03236	-0.81249	18.76940	-0.03151	35.13342	0.57508	1.14465
mean-$LPM_{\tau=0,a=2}$ RS PCC-GARCH	F	0.08182	0.20509	-0.01800	-0.03188	-0.77772	16.58418	-0.03722	32.05751	0.39897	1.10425
	T	0.05847	0.21711	-0.01906	-0.03445	-1.51992	18.52177	-0.06466	38.32385	0.26929	1.07853

[1] Standard Deviation
[2] Cornish-Fisher
[3] Max Drawdown
[4] Sharpe Ratio

TABLE 2.17 – Les performances des différents modèles sur le portefeuille 2

Maintenant, lorsque nous utilisons le modèles SR PCC (avec des marginales RS-GARCH), la dépendance reste majoritairement concentrée au centre de distribution conjointe. Mais cette fois-ci, la dépendance de queue supérieure est plus élevée que la dépendance de queue inférieure pour tous les portefeuilles. Par exemple au portefeuille 2, les copules bivariées requises pour la modélisation de la dépendance de queue inférieure et supérieure ont été utilisées 158 et 237 fois respectivement. Au même moment, les copules bivariées dont les propriétés mathématiques permettent de modéliser la dépendance au centre de la distribution conjointe ont été utilisées 391. Quel que soit le modèle utilisé PCC ou RS PCC, on peut apercevoir une hyper représentation de la copule de Frank. Elle a les mêmes caractéristiques que la copule gaussienne, c'est-à-dire la capacité de capturer la dépendance au centre de la distribution conjointe sauf qu'elle est mieux adaptée pour la dépendance non linéaire.

Models	Short selling	Annualized Return	Annualized STD[1]	Historical VaR 5%	CVaR 5%	Skewness	Kurtosis	CF[2]CVaR 5%	MDD[3]	Annualized SR[4]	Omega Ratio
equally weighted		-0.00672	0.27851	-0.02638	-0.04418	-1.28687	15.87982	-0.07896	48.64703	-0.02414	1.02431
mean-var	F	0.01843	0.22024	-0.01845	-0.03530	-1.72864	22.57858	-0.06347	41.93214	0.08366	1.04205
	T	0.02752	0.21338	-0.01625	-0.03316	-1.70466	20.18317	-0.06632	40.13764	0.12898	1.04889
mean-var t Cop GARCH	F	0.02465	0.23141	-0.01887	-0.03652	-1.30946	20.92133	-0.05346	43.90397	0.10654	1.04805
	T	0.00423	0.22053	-0.01862	-0.03419	-0.97683	16.11746	-0.05077	39.51367	0.01919	1.02672
mean-var DCC-GARCH	F	0.02410	0.23149	-0.01882	-0.03654	-1.30997	20.90332	-0.05358	43.93919	0.10411	1.04752
	T	0.02076	0.20720	-0.01731	-0.03251	-1.01929	15.84526	-0.04986	35.04837	0.10018	1.04235
mean-var PCC-GARCH	F	0.04750	0.21847	-0.01888	-0.03561	-2.06283	24.32095	-0.06927	42.76130	0.21742	1.07075
	T	-0.00292	0.21369	-0.01721	-0.03497	-1.66258	19.62443	-0.06629	44.61421	-0.01365	1.01973
mean-var RS PCC-GARCH	**F**	**0.12164**	**0.23301**	**-0.01889**	**-0.03685**	**-1.78365**	**19.09379**	**-0.07601**	**43.52882**	**0.52203**	**1.13148**
	T	**0.10637**	**0.23329**	**-0.01852**	**-0.03738**	**-1.86628**	**20.14718**	**-0.07690**	**46.94013**	**0.45595**	**1.11846**
mean-$LPM_{\tau=0,a=1}$	F	0.03977	0.22767	-0.01979	-0.03653	-1.73518	23.04828	-0.06428	41.80403	0.17468	1.06370
	T	0.02559	0.22156	-0.01723	-0.03433	-1.85489	25.15470	-0.06006	40.59910	0.11550	1.04972
mean-$LPM_{\tau=0,a=1}$ t Cop GARCH	F	0.03292	0.23111	-0.01993	-0.03623	-1.05178	20.55297	-0.04096	42.48215	0.14244	1.05527
	T	0.02228	0.21964	-0.01899	-0.03395	-1.01211	16.76328	-0.05034	39.65295	0.10145	1.04353
mean-$LPM_{\tau=0,a=1}$ DCC-GARCH	F	0.02518	0.21082	-0.01716	-0.03374	-1.89284	22.94871	-0.06527	41.29880	0.11942	1.04818
	T	0.01779	0.20979	-0.01724	-0.03312	-0.96186	15.28116	-0.04948	37.03633	0.08479	1.03959
mean-$LPM_{\tau=0,a=1}$ PCC-GARCH	F	0.05606	0.22232	-0.01927	-0.03523	-1.76352	23.89882	-0.06097	41.72396	0.25215	1.07876
	T	0.01609	0.21056	-0.01748	-0.03426	-1.29146	16.45524	-0.05903	43.17157	0.07641	1.03764
mean-$LPM_{\tau=0,a=1}$ **RS PCC-GARCH**	**F**	**0.06399**	**0.22735**	**-0.01959**	**-0.03685**	**-1.77761**	**17.87551**	**-0.07522**	**43.77396**	**0.28145**	**1.08261**
	T	**0.06601**	**0.22824**	**-0.01944**	**-0.03700**	**-1.82982**	**20.17999**	**-0.07440**	**45.28115**	**0.28920**	**1.08553**
mean-$LPM_{\tau=0,a=2}$	F	0.02065	0.21812	-0.01869	-0.03518	-1.73906	21.43558	-0.06626	41.77368	0.09466	1.04377
	T	0.00784	0.21263	-0.01664	-0.03320	-1.86691	20.85445	-0.06945	41.44703	0.03689	1.02959
mean-$LPM_{\tau=0,a=2}$ t Cop GARCH	F	0.01160	0.23225	-0.02021	-0.03687	-1.28174	19.42808	-0.05764	44.44588	0.04995	1.03574
	T	0.00326	0.21987	-0.01849	-0.03388	-0.81782	14.86156	-0.04700	39.05348	0.01481	1.02567
mean-$LPM_{\tau=0,a=2}$ DCC-GARCH	F	0.03350	0.20943	-0.01765	-0.03332	-1.83218	22.22885	-0.06466	40.66535	0.15997	1.05652
	T	0.03029	0.20967	-0.01726	-0.03291	-0.95381	16.14232	-0.04719	35.26921	0.14446	1.05225
mean-$LPM_{\tau=0,a=2}$ PCC-GARCH	F	0.06192	0.21978	-0.01886	-0.03504	-1.88094	24.52920	-0.06275	42.44811	0.28173	1.08441
	T	0.02126	0.21800	-0.01742	-0.03512	-1.75387	20.96806	-0.06773	44.10397	0.09751	1.04320
mean-$LPM_{\tau=0,a=2}$ **RS PCC-GARCH**	F	0.04995	0.22785	-0.01931	-0.03625	-1.75685	20.23175	-0.07222	41.89682	0.21921	1.07077
	T	**0.08081**	**0.23724**	**-0.02011**	**-0.03811**	**-1.66647**	**17.91588**	**-0.07581**	**45.59326**	**0.34064**	**1.09469**

[1] Standard Deviation
[2] Cornish-Fisher
[3] Max Drawdown
[4] Sharpe Ratio

TABLE 2.18 – Les performances des différents modèles sur le portefeuille 3

Models	Short selling	Annualized Return	Annualized STD[1]	Historical VaR 5%	CVaR 5%	Skewness	Kurtosis	CF[2]CVaR 5%	MDD[3]	Annualized SR[4]	Omega Ratio
equally weighted		0.07868	0.24596	-0.02321	-0.03944	-1.27520	16.60438	-0.06784	43.19419	0.31991	1.09203
mean-var	F	0.07569	0.18051	-0.01419	-0.02732	-0.97562	23.00698	-0.01784	34.11359	0.41932	1.10992
	T	-0.03437	0.18007	-0.01421	-0.02903	-2.18613	22.25618	-0.06313	45.88244	-0.19087	0.97874
mean-var t Cop GARCH	F	-0.00858	0.21682	-0.01979	-0.03416	-1.99032	22.23299	-0.07165	37.09275	-0.03959	1.01410
	T	0.07830	0.22017	-0.01946	-0.03317	-1.37142	15.50487	-0.06490	33.54988	0.35564	1.08911
mean-var DCC-GARCH	F	0.08590	0.15908	-0.01438	-0.02442	-0.55544	10.34014	-0.03226	21.65529	0.53994	1.11983
	T	0.06466	0.12990	-0.01339	-0.01896	-0.25165	1.51358	-0.01975	13.83659	0.49774	1.09678
mean-var PCC-GARCH	F	0.00369	0.23302	-0.02120	-0.03607	-2.31649	38.67979	-0.02074	40.30898	0.01584	1.02908
	T	-0.09254	0.28216	-0.02356	-0.04574	-2.56583	39.33219	-0.02610	56.62144	-0.32799	0.95813
mean-var RS PCC-GARCH	F	0.10526	0.16315	-0.01484	-0.02465	-0.52014	9.42001	-0.03281	23.69412	0.64517	1.14027
	T	0.05823	0.13606	-0.01337	-0.01958	-0.06472	2.19459	-0.02006	13.26209	0.42802	1.08696
mean-$LPM_{\tau=0,a=1}$	F	0.04448	0.18918	-0.01552	-0.02986	-1.31837	21.81503	-0.04103	37.21450	0.23513	1.07272
	T	-0.02027	0.17530	-0.01339	-0.02729	-1.71188	21.83725	-0.05185	40.55707	-0.11564	0.99419
mean-$LPM_{\tau=0,a=1}$ t Cop GARCH	F	0.02284	0.20954	-0.01934	-0.03243	-1.74736	23.66133	-0.05775	32.41687	0.10899	1.04335
	T	0.07470	0.22174	-0.02039	-0.03234	-1.24098	15.05588	-0.06211	27.98491	0.33690	1.08365
mean-$LPM_{\tau=0,a=1}$ **DCC-GARCH**	**F**	**0.10248**	**0.16758**	**-0.01418**	**-0.02555**	**-0.69513**	**12.87050**	**-0.03492**	**25.35648**	**0.61150**	**1.14046**
	T	**0.13677**	**0.14185**	**-0.01307**	**-0.02096**	**-0.19476**	**4.07101**	**-0.02319**	**15.36930**	**0.96417**	**1.19273**
mean-$LPM_{\tau=0,a=1}$ PCC-GARCH	F	0.10089	0.23927	-0.01900	-0.03526	-0.84071	28.61660	-0.01898	39.03318	0.42167	1.11449
	T	-0.01826	0.30550	-0.02668	-0.04602	-1.58615	18.94313	-0.09318	50.27628	-0.05978	1.01890
mean-$LPM_{\tau=0,a=1}$ **RS PCC-GARCH**	**F**	**0.10751**	**0.17704**	**-0.01503**	**-0.02738**	**-0.47748**	**13.16872**	**-0.02915**	**26.86869**	**0.60723**	**1.14301**
	T	**0.10918**	**0.16680**	**-0.01523**	**-0.02674**	**-0.49871**	**5.94553**	**-0.03296**	**16.01017**	**0.65454**	**1.14373**
mean-$LPM_{\tau=0,a=2}$	F	0.06130	0.17671	-0.01487	-0.02752	-1.34283	20.80381	-0.04229	34.81739	0.34688	1.09224
	T	-0.02023	0.17861	-0.01436	-0.02879	-2.19463	18.88235	-0.06391	42.76026	-0.11326	0.99525
mean-$LPM_{\tau=0,a=2}$ t Cop GARCH	F	0.02812	0.21825	-0.01851	-0.03401	-1.90682	23.17347	-0.06743	36.84549	0.12882	1.04904
	T	0.09960	0.22092	-0.02024	-0.03264	-1.38591	15.43207	-0.06548	28.64240	0.45085	1.10585
mean-$LPM_{\tau=0,a=2}$ DCC-GARCH	F	0.10044	0.15988	-0.01423	-0.02391	-0.28015	10.14004	-0.02558	22.02364	0.62823	1.13793
	T	0.09181	0.13401	-0.01260	-0.01784	0.09535	2.10164	-0.01817	21.17389	0.68509	1.12952
mean-$LPM_{\tau=0,a=2}$ PCC-GARCH	F	0.09184	0.23305	-0.01913	-0.03485	-1.06158	28.80086	-0.01927	39.41176	0.39410	1.10856
	T	-0.00014	0.27216	-0.02450	-0.04118	-1.73230	20.23508	-0.08559	45.61153	-0.00052	1.02753
mean-$LPM_{\tau=0,a=2}$ **RS PCC-GARCH**	**F**	**0.10068**	**0.17010**	**-0.01515**	**-0.02618**	**-0.29293**	**10.12987**	**-0.02761**	**24.16989**	**0.59193**	**1.13327**
	T	**0.11499**	**0.15073**	**-0.01385**	**-0.02113**	**0.02809**	**3.98787**	**-0.02176**	**16.61140**	**0.76287**	**1.14991**

[1] Standard Deviation
[2] Cornish-Fisher
[3] Max Drawdown
[4] Sharpe Ratio

TABLE 2.19 – Les performances des différents modèles sur le portefeuille Benchmark

En ce qui concerne l'optimisation de portefeuille, bien que la théorie du portefeuille de Markowitz offre la possibilité de trouver un équilibre entre la maximisation des rendements et la minimisation des risques en établissant un cadre de diversification du portefeuille, elle présente certaines limites. Parmi lesquelles, nous avons le choix de la variance comme mesure du risque ainsi que l'incertitude liée à l'estimation des paramètres d'entrées qui est susceptible de nuire à l'avantage que peut offrir la diversification et conduire à des portefeuilles sous-optimaux. C'est la motivation des travaux de DeMiguel, Garlappi, and Uppal (2009), soulignant le fait que même les versions sophistiquées du modèle de Markowitz (mean-variance) peuvent générer des portefeuilles dont les performances ne surpassent pas les performances de portefeuilles équipondérés, c'est pourquoi le modèle $1/n$ sert souvent de modèle de référence.

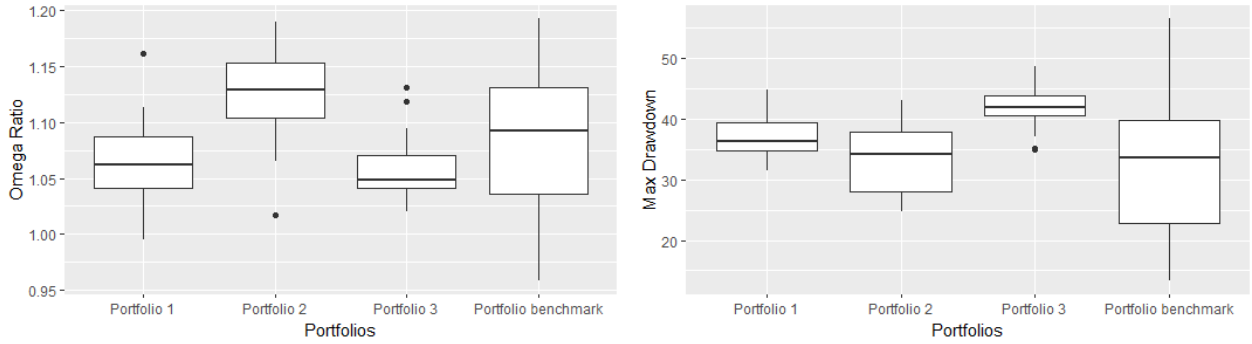
Étant donné que les rendements et les risques évoluent dans le temps, il serait inapproprié de rester dans le cadre statique qui ne tient pas compte de l'horizon temporel. Par conséquent, afin de respecter les critères de rendement et de risque préalablement définis, un rééquilibrage périodique est nécessaire, dans notre cas, tous les 5 jours. C'est dans ce sens, qu'en plus du modèle équipondéré ($1/N$), nous avons évalué les performances de nos portefeuilles avec les modèles mean-var et mean-LPM en considérant différentes structures de dépendance. Dans notre étude, pour le modèle mean-LPM, nous avons considéré $\tau = 0$ et $a = 1,2$ (voir les équations 2.29 et 2.30). Les performances des portefeuilles sont enregistrées dans les tables 2.16, 2.17, 2.18 et 2.19. En plus des tables précédemment citées, les figures A.1-A.8 nous permettent de mieux apprécier les performances de nos portefeuilles en fonction des différents modèles utilisés. Une brève analyse de tous ces résultats nous permet de constater que le portefeuille 2 composé des actifs avec les notations ESG AVERAGE (A et BBB) réalise les meilleures performances au regard des ratios Oméga sur l'ensemble des modèles. En effet, compte tenu des caractéristiques des distributions de probabilité des portefeuilles générés en l'occurrence le skewness et le kurtosis, le ratio Oméga semble être l'indicateur de performance le plus approprié.

Cet indicateur est défini comme le rapport entre l'aire sous la courbe de rendement cumulative positive et l'aire totale sous la courbe de rendement cumulative. Les valeurs du ratio Oméga des modèles obtenus après optimisation du portefeuille 2 sont toutes supérieures à 1. Cela indique que la probabilité cumulée des rendements positifs est plus élevée que la probabilité cumulée des rendements négatifs, ce qui suggère un rendement ajusté au risque favorable. Et si nous devons classer les autres portefeuilles par ordre décroissant des valeurs de leur ratio Oméga, nous auront, le portefeuille Benchmark suivi des portefeuilles 2 et 3 composés respectivement des actifs de notations ESG LEADER (AAA et AA) et LAGGARD (CCC, B, et BB).

Ensuite, en termes de risque, même si nous avons considéré la VaR historique, les CVaR (historique et Cornish-Fisher), l'indicateur de risque que nous avons souhaité mettre en avant est le Maximum Drawdown (MDD). Il donne un aperçu du risque et de la volatilité d'un portefeuille tout en mettant l'accent sur les pertes maximales encourues. En observant les valeurs du max drawdown (MDD), le portefeuille Benchmark est celui qui présente les valeurs le plus faibles et donc est le portefeuille le moins risqué. Un résultat qui semble logique dans la mesure où ce portefeuille est composé de l'ensembles des actifs des autres portefeuilles et bénéficie donc d'une diversification maximale tant au niveau des notations ESG qu'au niveau des secteurs d'activités des entreprises. Ainsi sur la base de cet indicateur, après le portefeuille Benchmark, le portefeuille 2 présente les risques les plus faibles suivi respectivement des portefeuilles 1 et 3. Nous discuterons en profondeurs de ces résultats dans la section suivante.

2.4 Discussions

Bien que les résultats obtenus semblent nous donner une certaine orientation, nous cherchons à montrer que les conclusions qui pourront en découler seront faites sur la base de preuves statistiques. C'est la raison pour laquelle, dans un premier temps, nous avons cherché à montrer qu'il y avait une différence statistiquement significative entre les médianes des valeurs des ratios Oméga des différents portefeuilles.



(a) Omega Ratio

(b) Max Drawdown

FIGURE 2.2 – Boxplot de 2 indices de performance des différents des portefeuilles

En plus de la figure 2.2a qui nous donne une tendance sur les conclusions à venir, le test de (Kruskal and Wallis, 1952) (Kruskal-Wallis Rank Sum Test) vient confirmer que le fait qu'au moins une des médianes est significativement différentes des autres. De plus, le test de Kruskal-Wallis dont les résultats sont consignés à la table 2.20 n'identifie pas spécifiquement les paires de médianes qui diffèrent les unes des autres. D'où la nécessité de recourir au test de Wilcoxon (1946) (Wilcoxon Rank Sum and Signed Rank Test). Ce test de Wilcoxon unilatéral nous permet de faire une comparaison spécifique entre les différentes paires. Il en ressort de ce test dont les résultats sont enregistrés dans la table 2.21 que sur la base du ratio Oméga comme indicateur de performances, le portefeuille 2 réalise des performances plus élevées que celles des 3 autres portefeuilles. En seconde position, nous avons le portefeuille Benchmark dont les performances surpassent celles des portefeuilles 1 et 3.

Omega Ratio test	
H_0	: $m_1 = m_2 = m_3 = m_{bench}$ (all medians are equal)
H_1	: $\exists(i,j)\ m_i \neq m_j$ (at least one median is different)
Kruskal-Wallis chi-squared = 37.977, df = 3, p-value = 2.858e-08	

Max Drawdown test	
H_0 :	$m_1 = m_2 = m_3 = m_{bench}$ (all medians are equal)
H_1 :	$\exists(i,j)\ m_i \neq m_j$ (at least one median is different)
Kruskal-Wallis chi-squared = 41.032, df = 3, p-value = 6.439e-09	

TABLE 2.20 – Tests statistique sur 2 indices de performance des différents portefeuilles

Pour ces deux derniers portefeuilles, les résultats du test Wilcoxon nous emmènent à conclure qu'il n'y a pas de différence significative entre leurs performances. C'est un résultat qui va à l'opposé d'une partie de la littérature et de la pensée collective dans la mesure où le portefeuille 1 qui est composé uniquement des best-in-class (AAA et AA) ne réalise pas des performances supérieures au portefeuille 3 dont tous ses composants sont des worst-in-class (CCC, B et BB). Cela pourrait s'expliquer par le fait qu'au fil du temps, les actifs qui composent le portefeuille 1 ont souvent réalisé des performances négatives raison pour laquelle les deux modèles d'optimisation qui ont le mieux performé étaient respectivement les modèles mean-var PCC-GARCH et mean-$LPM_{\tau=0,a=2}$ PCC-GARCH(with short selling). De plus, la modélisation de la structure de dépendance du portefeuille 1 avec la copule de vine (PCC-GARCH) nous avait permis de mettre en évidence le fait que la majorité de la dépendance soit concentrée au centre de la distribution conjointe et que la dépendance pour la queue inférieure soit légèrement plus élevée que celle de la queue supérieure. Alors nous pouvons dire que le portefeuille composé des best-in-class (portefeuille 1) a une dépendance de queue inférieure et non linéaire.

Portfolio 2 vs Portfolio 1	
H_0	: $m_2 = m_1$ (Both medians are equal)
H_1	: $m_2 > m_1$ (Median 2 is higher than median 1)
Wilcoxon test : W = 836, p-value = 4.263e-08	

Portfolio 1 vs Portfolio 3	
H_0	: $m_1 = m_3$ (Both medians are equal)
H_1	: $m_3 > m_1$ (Median 3 is higher than median 1)
Wilcoxon test : W = 548, p-value = 0.174	

Portfolio 2 vs Portfolio Benchmark	
H_0	: $m_2 = m_{bench}$ (Both medians are equal)
H_1	: $m_2 > m_{bench}$ (Median 2 is higher than median bench)
Wilcoxon test : W = 697, p-value = 0.0009953	

Portfolio Benchmark vs Portfolio 1	
H_0	: $m_1 = m_{bench}$ (Both medians are equal)
H_1	: $m_{bench} > m_1$ (Median bench is higher than median 1)
Wilcoxon test : W = 612, p-value = 0.03243	

TABLE 2.21 – Test de comparaison des valeurs des Omega ratio des différents portefeuilles

Portfolio 2 vs Portfolio 1	
H_0	: $m_2 = m_1$ (Both medians are equal)
H_1	: $m_2 < m_1$ (Median 2 is lower than median 1)
Wilcoxon test : W = 287, p-value = 0.002995	

Portfolio 1 vs Portfolio 3	
H_0	: $m_1 = m_3$ (Both medians are equal)
H_1	: $m_1 < m_3$ (Median 1 is lower than median 3)
Wilcoxon test : W = 155, p-value = 6.916e-07	

Portfolio Benchmark vs Portfolio 1	
H_0	: $m_2 = m_{bench}$ (Both medians are equal)
H_1	: $m_{bench} < m_1$ (Median bench is lower than median 1)
Wilcoxon test : W = 348, p-value = 0.03141	

Portfolio 2 vs Portfolio Benchmark	
H_0	: $m_2 = m_{bench}$ (Both medians are equal)
H_1	: $m_2 < m_{bench}$ (Median 2 is lower than median bench)
Wilcoxon test : W = 509, p-value = 0.6574	

TABLE 2.22 – Test de comparaison des valeurs des Max Drawdown des différents portefeuilles

Pendant ce temps, au portefeuille 2, les deux modèles les plus performants sont respectivement les modèles mean-$LPM_{\tau=0,a=2}$ t copula-GARCH (DCC-GARCH) et mean-var t copula-GARCH (DCC-GARCH). La capacité de la copule de Student à la dépendance dans les queues inférieure et supérieure n'est plus à démontrer. Et lorsqu'on vient en plus la coupler au modèle DCC-GARCH, on donne la possibilité à paramètres de varier dans le temps et de parvenir à capturer les changements dans la structure de dépendance au fil du temps. La structure de dépendance du portefeuille 2 qui est composé des actifs de notations ESG AVERAGE (BBB et A) est symétrique et linéaire. Quant au portefeuille 3 composé des actifs ayant les pires notations ESG, les deux modèles qui se sont les plus distingués sont respectivement les modèles mean-var RS PCC-GARCH et mean-var RS PCC-GARCH (with short selling). La prise en compte des changements de régime dans les distributions marginales a eu l'impact espéré sur la structure de dépendance modélisée avec la copule de vine. La flexibilité qu'offre la copule de vine durant la modélisation de la structure de dépendance en utilisant une cascade de copules bivariées nous emmène à dire que le portefeuille 3 a une dépendance de queue inférieure et non linéaire. Avec ses 117 actifs, le portefeuille Benchmark peut être considéré comme un portefeuille de grande dimension. La plupart du temps, les portefeuilles qui ont cette caractéristique ne nécessitent pas de structure de dépendance complexe c'est qui peut en partie expliquer le fait que le modèle mean-$LPM_{\tau=0,a=1}$ DCC-GARCH ait réalisé les meilleures performances.

Sauf qu'après ce modèle, les meilleures performances ont été réalisées avec les modèles mean-$LPM_{\tau=0,a=2}$ RS PCC-GARCH (with short selling) et mean-$LPM_{\tau=0,a=1}$ RS PCC-GARCH respectivement. Il est alors difficile de spécifier quelle pourrait être le modèle de dépendance le plus approprié à ce portefeuille et donc la nature de sa dépendance.

La même série de tests a été faite afin de déterminer le portefeuille le moins risqué. La figure 2.2b permet de visualiser la distribution et les caractéristiques principales des valeurs du maximum drawdown de chaque portefeuille. Le test de Kruskal-Wallis dont les résultats sont enregistrés dans la table 2.20 nous emmène à conclure qu'au moins une des médianes est significativement différentes des autres. Et comme nous l'avons fait précédemment, afin d'avoir plus de précision sur les différences entre chaque paire de médianes, nous avons effectué le test de Wilcoxon (1946) (Wilcoxon Rank Sum and Signed Rank Test) dont les résultats sont à la table 2.22. Les résultats de ce test nous permettent d'affirmer qu'en terme de risque, il n'y a pas de différence significative entre le portefeuille 2 et le portefeuille Benchmark. Par contre le portefeuille 2 s'avère moins risqué que le portefeuille 1 composé des actifs avec les meilleures notations ESG (AAA et AA). Parmi ces quatre portefeuilles, celui qui présente le plus de risque est le portefeuille 3 composé des actifs avec les pires notations ESG (CCC, B et BB).

2.5 Conclusion

Au terme de cette analyse empirique, nous sommes en mesure d'affirmer que l'impact de la prise en compte des critères ESG sur la performance du portefeuille est non négligeable. Et donc, par l'entremise de notre papier, nous sommes en mesure de participer au débat relatif aux informations que les scores ESG peuvent transmettre sur les niveaux de risque et de rentabilité d'un investissement qui dans notre cas est un portefeuille d'actifs financiers. Les scores ESG fournissent des informations non financières qui portent sur la durabilité et la responsabilité d'une entreprise donc susceptibles d'aiguiller les investisseurs dans l'évaluation des risques et des opportunités liés à un investissement. Par exemple, l'utilisation de ces informations non financières peut permettre d'identifier et d'évaluer les risques liés à des facteurs tels que la pollution, les litiges juridiques, les scandales de gouvernance, de potentielles réglementations qui pourraient affecter une entreprise et son secteur d'activité. Selon cette logique, les entreprises qui ont des meilleures notations ESG doivent présenter les risques les plus faibles puisse qu'elles aient mise en place une politique de gestion des risques robuste permettant de réduire les vulnérabilités et les incertitudes liées à son activité. Dans un premier temps, cela a pour conséquence d'augmenter les investissements des acteurs des marchés financiers sensibles aux considérations durables. Ensuite, cette bonne gestion des risques extra-financiers devrait engendrer une performance financière durable et élevée qui va attirer les investisseurs cherchant à anticiper de potentielles tendances haussières futures. Cela peut également encourager une augmentation des investissements, ce qui se traduit par des rendements financiers élevés. Bien que certaines recherches académiques (Achim and Borlea, 2015) aient suggéré une corrélation positive entre les performances ESG et financières, il est important de noter que la corrélation n'implique pas nécessairement la causalité. Dans le cadre de notre analyse, nous avons travaillé sur l'agrégation de ces actifs dans des portefeuilles et non sur les actifs individuellement.

Afin d'éviter que le choix de la méthode d'agrégation puisse biaiser nos résultats, nous avons considéré des modèles d'optimisation de portefeuille pour trois mesures de risques (variance, $LPM_{\tau=0,a=1}$, $LPM_{\tau=0,a=2}$) dans des configurations différentes en terme de structure de dépendance. Les différents résultats obtenus à l'issue de cette étude comparative suggèrent que le portefeuille 2 composé des actifs de notations ESG AVERAGE (BBB et A) est celui qui a généré les rendements les plus élevés et les risques les plus faibles. Suivi du portefeuille Benchmark qui quant à lui est composé de l'ensemble des actifs des trois portefeuilles et donc des actions de toutes les notations ESG (CCC, B, BB, BBB, A, AA et AAA). Composé des actifs ayant les meilleures notations ESG, le portefeuille 1 qui était censé réaliser les meilleures performances en termes de rendement et de risque parmi les quatre n'est que troisième. En dernière position, nous avons le portefeuille 3 composé de actifs de notations ESG LAGGARD (CCC, B, et BB). Bien qu'il n'y ait pas de différence statistique entre les rendements de ce portefeuille et celui du portefeuille 1, nous l'avons classé quatrième parce qu'il est celui qui a présenté les risques les plus élevés. Ainsi, pour les investisseurs concernés par les questions relatifs aux ESG, particulièrement ceux qui restent focalisés sur les concepts de risque et de rendement, nos résultats suggèrent que les deux perspectives ne sont pas contradictoires.

Une autre contribution de notre papier est la mise en évidence de l'importante de la modélisation adéquate de la dépendance dans le cadre d'optimisation de portefeuille. En effet, au regard des résultats, nous nous apercevons que les meilleures performances ont été obtenues avec lorsque que nous avons tenu compte de structure de dépendance. De plus, la procédure de modélisation de la dépendance que nous avons proposé en intégrant la notion de changement de régime dans les distributions marginales et modélisant la structure de dépendance avec une copule de vine. Cette procédure a permis d'obtenir des performances figurant parmi les meilleures des portefeuilles 3 et Benchmark.

CHAPITRE 3

Évaluation empirique de la contribution à la performance d'un portefeuille d'actifs ESG : Cas des marchés de la zone Euro

3.1 Introduction

L'émergence de l'innovation financière et de l'intelligence artificielle, couplée à l'intérêt des entités économiques pour les transitions écologique et énergétique, conduira inévitablement à l'émergence de nouvelles technologies puissantes dans la théorie financière. Ainsi, en gestion de portefeuille, plusieurs méthodes d'optimisation robustes conduisent à une meilleure allocation des actifs financiers. L'allocation optimale obtenue garantit une meilleure optimisation risque-rendement du portefeuille concerné. Bien que les deux concepts soient importants, le premier semble recevoir une attention particulière.

En fait, comprendre les différents types de risques auxquels nous pouvons être confrontés est essentiel pour mieux prévoir les rendements négatifs qui peuvent entraîner des pertes importantes. En général, il existe deux types de risque : le risque systématique et le risque non systématique. Le risque systémique est inhérent aux marchés en raison de l'interconnexion des marchés financiers internationaux, où des effets de contagion peuvent être observés d'un marché à l'autre. La crise des subprimes de 2008 et la crise sanitaire du Covid-19 de 2020 ont mis en évidence l'intensité et la rapidité avec lesquelles les chocs peuvent se propager à l'ensemble du système financier, souvent avec des externalités négatives qui peuvent se faire sentir dans l'économie réelle.

Quant au risque non systématique, également appelé risque spécifique, comme son nom l'indique, il représente le risque que la valeur d'un actif ou d'un groupe d'actifs ne soit pas présente sur le marché en raison de ses facteurs uniques. Dans notre cas, dans le cadre de la gestion de portefeuille, la dernière source de risque peut être éliminée grâce au principe de diversification. En intégrant des méthodes statistiques dans la sélection des portefeuilles financiers, on arrive à des méthodes plus raffinées d'optimisation de portefeuille, dont la plus célèbre a été développée par Harry Markowitz en 1952, qui a établi les bases fondamentales de la théorie moderne du portefeuille (MPT) dans son article "*Portfolio Selection*"(Markowitz, 1952a).

Sa méthode fournit un outil d'aide à la décision financière en formulant le problème de gestion de portefeuille comme un problème d'optimisation dans lequel le rendement et le risque sont mesurés respectivement par une moyenne et une variance empiriques, fournissant la probabilité de trouver le rendement maximal pour un niveau de risque donné ou Pour un niveau de rendement donné, le risque est minime. Cependant, ce modèle présente un certain nombre de faiblesses, à commencer par la présumée normalité de la distribution des probabilités de rendement, puis l'hypersensibilité des poids aux erreurs associées aux estimations des paramètres d'entrée, et enfin le choix de la variance comme mesure du risque.

Bien que sous l'hypothèse de normalité, la moyenne et la variance empiriques sont des estimateurs du maximum de vraisemblance, qui ont l'avantage d'être non biaisés et cohérents. Mandelbrot (1963, 1997) et Fama (1963) désapprouvent cette hypothèse car elle s'écarte trop de la réalité en termes de faits stylisés de séries temporelles financières. Comme l'ont démontré Broadie (1993), Chopra and Ziemba (1993), Michaud (1989), la qualité de l'optimisation du modèle est étroitement liée à la précision des estimations des paramètres d'entrée. Des erreurs conduisent à des allocations sous-optimales voire à une surpondération de certains actifs, compromettant les avantages que la diversification pourrait apporter. Green, Hand, and Zhang (2013) dans leur article "*The supraview of return predictive signals*" citent de nombreuses solutions pour estimer l'erreur moyenne, parmi lesquelles on peut citer la moyenne mobile pondérée exponentielle (EWMA), de Bayes-Stein(Jorion, 1986) et de Martinelli(Martellini, 2009).

Dans sa version Global Minimum Variance (GMV), nous pouvons voir le rôle central joué par la matrice de covariance, nécessitant ainsi une méthode d'estimation appropriée pour fournir une meilleure analyse de portefeuille. Le principe de diversification stipule que des corrélations faibles ou négatives (covariances) entre les rendements des actifs entraînent collectivement une réduction de la variance des rendements du portefeuille. Le lien entre les actifs est mesuré avec la covariance (corrélation) dans le cadre du modèle de Markowitz, car sous l'hypothèse d'une distribution gaussienne multivariée, le concept de lien connu sous la structure de dépendance est obtenu par la corrélation de Pearson Coefficients mesurés. L'utilisation du coefficient de corrélation de Pearson comme mesure de la dépendance peut être critiquée par beaucoup car elle ne capture que la dépendance linéaire.

En tant que tel, il ne peut pas décrire les dépendances entre les variables avec des distributions non elliptiques et n'est pas invariant aux transformations incrémentielles et continues (telles que la transformation logarithmique utilisée pour calculer les rendements). Parmi les critiques précédemment évoquées, il reste le choix de la variance comme mesure du risque. Les rendements positifs sont plus souhaitables que les négatifs mais en raison de sa nature symétrique, la variance ne les différencie pas, les gains et les pertes sont traités de façon identique. Face à cette limitation, Markowitz a suggéré qu'une meilleure mesure du risque devrait prendre en compte les rendements inférieurs aux rendements attendus connus comme une mesure du risque de baisse (downside risk), d'où l'idée d'utiliser la semi-variance au lieu de la variance.

De nombreuses autres mesures de risque de baisse (downside risk) ont été développées ces dernières années parmi lesquelles la plus populaire la Value-at-Risk(Linsmeier and Pearson, 1996; Jorion, 2000) qui même été recommandée par le Comité de Bâle.

Elle mesure la pire perte attendue pour un certain niveau de confiance et horizon temporel. Cependant, cette mesure de risque a été critiquée, d'abord parce que la sous-additivité n'existe pas dans le cas des distributions générales, Szegö (2005) et Zhu and Fukushima (2009) soulignent qu'une propriété qui porte gravement atteinte au principe de diversification. Aussi, Cornuejols and Tütüncü (2006), Krokhmal et al. (2002), Quaranta and Zaffaroni (2008) ont souligné le caractère non convexe et non lisse de la VaR, qui rend difficile la recherche d'un minimum global du fait de l'existence d'un grand nombre de minima locaux. Par conséquent, il est considéré comme une mesure incohérente au sens de Artzner et al. (1999), et des mesures alternatives de risque avec les propriétés souhaitées sont nécessaires. La Conditional Value-at-Risk (CVaR) proposée par Artzner et al. (1999) et Szegö (2005), qui peut être définie comme la moyenne de la queue de distribution dépassant la VaR, répond à ce besoin.

Avec des propriétés qui manquent à la VaR, nous pouvons ajouter de la consistance avec la dominance stochastique du second ordre (Ogryczak and Ruszczynski, 2002), la CVaR s'est avérée plus efficace et fiable, et peut même remplacer la VaR dans les accords de Bâle III. Dans le contexte de l'optimisation de portefeuille, les travaux fondamentaux de Rockafellar et al. (2000), (2002)Rockafellar and Uryasev (2002) ont développé un cadre pour minimiser la CVaR en résolvant une programmation linéaire auxiliaire qui calcule simultanément la VaR et le problème CVaR à obtenir. Leur approche évite les optimisations purement numériques plus compliquées pour les grandes dimensions qui nécessitent de connaître la VaR à l'avance.

Un grand nombre d'études empiriques, telles que Topaloglou et al. (2002), Andersson et al. (2001), Chen et al. (2012) qui privilégient l'optimisation quadratique en utilisant le cadre mean-CVaR moyen au lieu du modèle de Markowitz. Cependant, contrairement à ces chercheurs, Lim et al. (2009) ont mené une étude pour trouver la limite de la CVaR. Dans leur article, ils sont arrivés à démontrer la fragilité de la CVaR lorsque les queues de la distribution de rendement deviennent plus épaisses, de sorte que ses performances peuvent être affectées par la distribution de rendement multivariée supposée. Les erreurs dans les estimations des rendements basées sur les hypothèses de distribution des rendements ne sont pas surprenantes. A partir du moment où nous disposons d'informations imparfaites, nous savons remettre en question la fiabilité et la qualité des données du fait de la grande quantité d'informations qu'il faut inclure sur les prix. On peut parler d'incertitude de Knight (Knightian uncertainty)(Knight, 1921; Ellsberg, 1961), car de nombreux facteurs sont inconnus de la plupart des investisseurs, tels que le processus par lequel les données ont été générées et, surtout, la distribution complète des rendements.

En parcourant la littérature, nous avons réalisé que les solutions aux erreurs d'estimation peuvent être divisées en deux approches : les approches bayésiennes et les approches d'optimisation robustes. En utilisant une approche bayésienne, Garlappi et al. (2007) affirment que cette méthode est neutre vis-à-vis de l'incertitude au sens de Knight (1921), puisqu'elle ne suppose qu'une seule probabilité a priori (prior) [1] pour la distribution prédictive des rendements des actifs. Black and Litterman (1992), Barberis (2000) ont été parmi les premiers auteurs à utiliser cette approche dans le cadre de l'optimisation de portefeuille. Face au problème des données imparfaites, la communauté scientifique, consciente de l'importance de la robustesse, travaille au développement de solutions appliquées avec succès dans de nombreux domaines.

1. une probabilité basée sur des données d'observation préalable

L'optimisation robuste est de plus en plus utilisée pour résoudre des problèmes d'optimisation de portefeuille. Comme nous l'avons vu dans Ben-Tal and Nemirovski (1999), ils formulent le problème de portefeuille à travers une approche de programmation linéaire robuste. La spécificité de cette approche réside dans l'assouplissement de l'hypothèse forte que la distribution est connue en faisant l'hypothèse plus faible qu'un seul ensemble de distributions qui devrait inclure la vraie distribution des paramètres est connu. Contrairement aux méthodes bayésiennes qui attribuent une seule probabilité a priori à différents modèles (scénarios) et dérivent des prédictions en faisant la moyenne des prédictions de probabilité, les méthodes d'optimisation robustes, d'autre part, considèrent plusieurs probabilités a priori pour déterminer une distribution empirique. L'application de cette approche à un problème de sélection de portefeuille dans le cadre mean-CVaR ou Robust CVaR (RCVaR) revient à minimiser la CVaR à condition que les distributions appartiennent à un ensemble d'incertitude ellipsoïdal [2].

Dans les articles de Benati (2003), Zhu and Fukushima (2009), Polak et al. (2010), Fertis et al. (2012), Kakouris and Rustem (2014) et bien d'autres, on suppose que les rendements appartiennent à un ensemble d'incertitude ellipsoïdal, et l'optimisation est basé sur la distribution du pire cas. Le portefeuille optimal est obtenu après optimisation sur la distribution la plus défavorable appartenant à l'ensemble d'incertitude qui donne la plus mauvaise valeur de la fonction objectif, on obtient la robustesse face à d'éventuelles erreurs de spécification de la distribution du portefeuille. Comme référence, nous pouvons citer les travaux de Zhu and Fukushima (2009) qui ont fait le choix de maximiser les rendements espérés dans les pires cas sous contrainte de la worst-case CVaR dans le cas de l'incertitude de distribution mixte de combinaison convexe. Pour protéger les portefeuilles contre les pires scénarios, Kakouris and Rustem (2014) proposent d'étendre le cadre worst-case CVaR(WCVaR) en utilisant des copules mixtes multidimensionnelles (multidimensional mixture Archimedean copulas).

La prédominance des distributions normales et elliptiques a généralement limité jusqu'à présent notre connaissance des dépendances de corrélation linéaire, de sorte que les copules nous permettent d'étendre cela. Suite aux travaux de Sklar (1959), les copules offrent une certaine flexibilité pour la modélisation afin de séparer les structures de dépendance et les marginales individuelles. Les copules peuvent être scindées en deux grandes familles : les copules elliptiques et les copules archimédiennes. Le concept de dépendance de queue (tail dependence) permet de faire ressortir leurs limites. En effet, pour la copule gaussienne, les dépendances de queue assimilables à des probabilités de de baisses ou de hausses simultanément sont supposées nulles, alors que pour la copule student elles sont supposées être les mêmes dans les deux cas.

Cette caractéristique soutient l'idée que les copules archimédiennes sont les mieux adaptées aux marchés financiers car elles permettent des structures de dépendance asymétriques. Cependant, des preuves empiriques nous ont montré que l'ampleur des dépendances de la queue inférieure est bien supérieure à celle des dépendances de la queue supérieure, de sorte qu'aucune des familles de copules précédemment citées ne semble bien adaptée à la nature complexe de la structure de dépendance.

2. L'ensemble d'incertitude ellipsoïdal est l'intersection d'un nombre fini d'ensembles d'ellipsoïdes résultant en un nombre fini d'inégalités quadratiques convexes

Sachant qu'il existe de nombreuses copules avec chaque caractéristique particulière, l'ensemble des copules candidates est non négligeable, donc la stratégie vise à les essayer toutes et à sélectionner la meilleure sur la base de certains critères du critère d'information bayésien (BIC)(Schwarz, 1978) et le test de Vuong(Vuong, 1989) ne sont pas réalisables. Connaissant les propriétés du modèle de dépendance que nous espérons, Chollete et al. (2005) et Hu (2006) ont proposé la solution de la copule mixte (mixture copula). Il s'agit d'une combinaison linéaire ou d'une moyenne pondérée de plusieurs copules individuelles avec des structures de dépendance suffisamment différentes pour construire une structure de dépendance qui n'appartient à aucune copule individuelle existante. C'est pourquoi la structure de dépendance est divisée en trois parties, généralement les copules utilisées pour construire l'ensemble candidat sont des copules de Clayton ou de rotated Gumbel pour les dépendances dans la queue inférieure, des copules de Student ou de Frank pour les dépendances au centre de la distribution et enfin les dépendances dans la queue supérieure de Gumbel.

Maintenant que les bases d'un cadre théorique solide ont été définies, nous pouvons aborder l'objectif principal de cette étude, qui est de juger de la pertinence des critères ESG, en l'occurrence la stratégie de composition du portefeuille Best-in-Class. Le comportement des investisseurs a changé, tout comme les marchés financiers, dépassant les critères purement financiers pour s'intéresser au positionnement des entreprises dans lesquelles ils investissent sur les questions environnementales, sociales et de gouvernance (ESG). Comme révélé par un article de Bloomberg [3] d'ici quelques, plus précisément en 2025, en raison de la croissance rapide des investissements ESG, les actifs sous gestion des fonds ESG seraient en mesure de dépasser 50 000 milliards de dollars. Cette tendance oblige les gérants de portefeuille à considérer les critères ESG comme de nouvelles contraintes en plus de la performance financière.

Dès lors, la mise en œuvre de telles stratégies nécessite un dispositif d'analyse financière complémentaire de la performance des entreprises au niveau environnemental, social et de bonne gouvernance, et il est donc logique qu'un grand nombre d'auteurs dont Amel-Zadeh and Serafeim (2018) soutiennent l'idée de choisir un score ESG. Parmi les stratégies d'intégration ESG les plus couramment utilisées, on peut citer Screening, Thematic et surtout Best-in-class avec pour spécificité d'exclure, d'inclure, de sélectionner ou de se concentrer sur des actions basées sur des scores ESG. Nous nous concentrerons sur la stratégie Best-in-class, car les entreprises avec des scores ESG élevés sont perçues comme plus performantes en termes de rendements et de risques. Après avoir analysé un grand nombre d'articles de la littérature sur l'investissement ESG et la performance financière associée, citefriede2015esg suggère que la plupart concluraient que l'ESG est positivement corrélé avec la performance financière d'une entreprise. De même, Lins et al. (2017), Garel and Petit-Romec (2021) soutiennent qu'en temps de crise, les entreprises les mieux notées ESG surperforment les autres.
En termes de risque, à travers leurs études empiriques, Fombrun, Gardberg, and Barnett (2000), Boutin-Dufresne and Savaria (2004), Lee and Faff (2009), Godfrey et al. (2009) concluent que les entreprises avec les meilleures notes ESG courent moins de risques de catastrophes environnementales, de boycotts de clients et d'autres types de scandales en raison de leur forte responsabilité dans les facteurs ESG.

3. https://www.bloomberg.com/company/press/esg-assets-rising-to-50-trillion-will-reshape-140-5-trillion-of-global-aum-by-20

Á l'opposé du courant de pensée précédent, d'autres chercheurs ont eu un discours plus élogieux à l'égard des entreprises ayant des faibles scores ESG souvent appelées "sin stocks" à cause des secteurs d'activités dans lesquels ils exercent comme l'alcool, les armes, les jeux de hasard, l'extraction minière, le nucléaire. Fabozzi et al. (2008) et Hong and Kacperczyk (2009) défendaient la thèse que ces entreprises avaient des performances supérieures. Un constat qui selon Demers et al. (2021), Pástor et al. (2022) était toujours valable même en période de crise. Mais de toutes les critiques, celles de Rudd (1981) résonnent le plus. Selon lui, intégrer des critères ESG dans la composition des portefeuilles compromettrait le principe de diversification.

Dans la littérature sur l'investissement responsable, on retrouve d'autres chercheurs comme Barnett and Salomon (2006), Renneboog et al. (2008) qui soutiennent que, contrairement aux stratégies d'investissement traditionnelles, intégrer des filtres ESG dans les stratégies d'investissement nuit aux avantages de la diversification. En parcourant la littérature, nous nous sommes rendu compte que les études menées concentrer autour d'un coté du modèle de Markowitz ainsi que ces variantes (voir Gasser et al., 2017; Chen et al., 2021; Schmidt, 2022) car lorsqu'ils évoquaient la diversification et le risque associé, très souvent la corrélation linéaire contribuait à la modélisation, d'un autre coté l'approche multi-objectifs ou multi-critères (voir Utz et al., 2014; Garc4́a et al., 2019)). C'est pourquoi l'apport de ce chapitre est double : enrichir la littérature sur les portefeuilles sains et l'investissement responsable de résultats empiriques.

La suite de ce chapitre est organisée comme suit. La section 3.2 présente les différents modèles mathématiques, y compris le cadre d'optimisation robuste CVaR et la spécification de deux modèles de copules hybrides utilisés avec le modèle GARCH. Dans la section 3.3, les applications numériques et les analyses des résultats empiriques obtenus seront abordées. La section 3.4 conclut le chapitre.

3.2 Cadre théorique

Les principales références de cette section sont Rockafellar, Uryasev, et al. (2000); Rockafellar and Uryasev (2002), Pflug (2000), Cornuejols and Tütüncü (2006), Würtz et al. (2009), Salahi et al. (2013), Kakouris and Rustem (2014), B Sabino da Silva and Ziegelman (2017).

3.2.1 Minimisation de la CVaR

Nous considérons un portefeuille d'actifs $A_1, \ldots, A_n$, $n \geqslant 2$ avec des rendements aléatoires. On suppose w_i et μ_i respectivement le poids et le rendement espéré pour le i-ème actif. L'espérance des rendements du portefeuille peut-être représenté comme

$$E[r_P] = \sum_{i=1}^{n} w_i \mu_i = \left(w_1, \ldots, w_n\right) \times \left(\mu_1, \ldots, \mu_n\right)^\top = w\mu^\top \tag{3.1}$$

Tout comme dans l'article de Cornuejols and Tütüncü (2006), nous supposerons que l'ensemble des portefeuilles réalisables est un ensemble polyédrique non vide et le représenterons comme $\Omega = \{w | Mw = b, Cw \leqslant d\}$ où M est une matrice $m \times n$, b est un vecteur de dimension m, C est une matrice de $p \times n$ et d est un vecteur de dimension p avec comme contrainte particulière $\sum_{i=1}^{n} w_i = 1$.

Soit $f(w,r)$ la fonction de perte lorsque l'on choisit un portefeuille w parmi un ensemble de portefeuilles réalisables Ω et soit r la réalisation des événements aléatoires (le vecteur des rendements des n actifs). Alors que l'espérance des rendement du portefeuille est définie comme dans l'équation (3.1), la fonction de perte est son négatif et est aussi une fonction convexe (linéaire) des variables du portefeuille w, c'est-à-dire :

$$f(w,r) = -w^\top r = -\big(w_1 r_1 + \ldots + w_n r_n\big) \tag{3.2}$$

Pour un vecteur de décision w fixé, la distribution de probabilité $p(r)$ connue, la probabilité que la perte ne dépasse pas un seuil γ est donnée par

$$\psi(w,\gamma) = \int_{\{f(w,r)\leqslant\gamma\}} p(r)dr \tag{3.3}$$

La fonction $\psi(w,\gamma)$ est une fonction de distribution cumulative de la perte et possède des propriétés mathématiques souhaitables. Elle détermine complètement le niveau de la VaR ainsi que de la CVaR. Alors, pour un niveau de confiance α donné, la VaR_α et la $CVaR_\alpha$ associées au portefeuille sont données par les équations :

$$VaR_\alpha(w) = min\{\gamma \in \mathbb{R} | \psi(w,\gamma) \geqslant \alpha\} \tag{3.4}$$

$$CVaR_\alpha(w) = \frac{1}{1-\alpha}\int_{f(w,r)\geqslant VaR_\alpha(w)} f(w,r)p(r)dr \tag{3.5}$$

Nous pouvons ainsi définir le problème de sélection de portefeuille mean-CVaR comme suit :

$$\begin{aligned} &\min_w \ CVaR_\alpha(w) \\ &\text{s.t } \ w^\top \mu \geqslant \bar{r} \\ &\qquad w \in \Omega \end{aligned} \tag{3.6}$$

Sur la base de la définition de CVaR à l'équation (3.5), on constate qu'elle implique explicitement la VaR, et donc son optimisation est lié à celle de la VaR qui est ni convexe ni linéaire ce qui peut compliquer la tâche. Afin de contourner cette difficulté, Rockafellar, Uryasev, et al. (2000); Rockafellar and Uryasev (2002) ont démontré que le calcul de la CVaR peut être réalisé sans nécessairement calculer la VaR au préalable en minimisant la fonction auxiliaire suivante par rapport à la variable $\gamma \in \mathbb{R}$:

$$F_\alpha(w,\gamma) = \gamma + \frac{1}{1-\alpha}\int_{f(w,r)\geqslant\gamma} \big(f(w,r)-\gamma\big)\, p(r)dr \tag{3.7}$$

Alternativement, nous pouvons écrire $F_\alpha(w,\gamma)$ comme suit :

$$F_\alpha(w,\gamma) = \gamma + \frac{1}{1-\alpha}\int \big(f(w,r)-\gamma\big)^+ p(r)dr \tag{3.8}$$

où $a^+ = max\{a,0\}$.

Proposition 3.2.1 *La fonction $F_\alpha(w,\gamma)$ est convexe par rapport à γ. De plus, minimiser $F_\alpha(w,\gamma)$ par rapport à γ donne la CVaR et la VaR est un point minimum de cette fonction par rapport à γ, c'est-à-dire*

$$\min_{\gamma\in\mathbb{R}} F_\alpha(w,\gamma) = F_\alpha(w, VaR_\alpha(w)) = CVaR_\alpha(w) \tag{3.9}$$

De plus, nous pouvons utiliser $F_\alpha(w,\gamma)$ pour trouver les poids optimaux conjointement de la CVaR et de la VaR sur un ensemble possible admissible, c'est-à-dire

$$\min_{w\in\Omega} CVaR_\alpha(w) = \min_{(w,\gamma)\in\Omega\times\mathbb{R}} F_\alpha(w,\gamma) \tag{3.10}$$

Pour finir, Pflug (2000) montre que sous des conditions assez générales, $F_\alpha(w,\gamma)$ est une fonction lisse.
De plus, si $f(w,r)$ est convexe par rapport à w, alors $F_\alpha(w,\gamma)$ est également convexe par rapport à w. Par conséquent, si l'ensemble admissible Ω est également convexe, nous devons alors résoudre un problème d'optimisation convexe lisse (smooth convex optimization problem).

Au lieu de supposer que la distribution est normale ou que nous avons une connaissance précise de la distribution du vecteur aléatoire r alors qu'en réalité elle n'est pas connue. Rockafellar et al. (2000) montrent qu'un échantillonnage par discrétisation via la simulation de Monte Carlo serait une meilleure alternative. Cela signifie que nous pouvons supposer J scénarios différents $r = (r_1,\dots,r_J)$ qui sont échantillonnés à partir de simulations numériques avec $P(r_j) = \pi_j$ les probabilités de scénarios r_j telles que $\sum_j^J \pi_j = 1$, $\pi_j \geqslant 0$, $j = 1,\dots,J$, et donc

$$\bar{F}_\alpha(w,\gamma) = \gamma + \frac{1}{1-\alpha}\sum_{j=1}^{J} \pi_j \big(f(w,r_j)-\gamma\big)^+ \tag{3.11}$$

Pour $\pi_j = 1/J$, $\forall j = 1,\dots,J$, l'équation (3.11) peut-être réduit à

$$\bar{F}_\alpha(w,\gamma) = \gamma + \frac{1}{(1-\alpha)J}\sum_{j=1}^{J} \big(f(w,r_j)-\gamma\big)^+ \tag{3.12}$$

Et donc, dans le problème $\min_w\ CVaR_\alpha(w)$, on remplace $F_\alpha(w,\gamma)$ par $\bar{F}_\alpha(w,\gamma)$

$$\min_{w,\gamma}\ \gamma + \frac{1}{(1-\alpha)J}\sum_{j=1}^{J} (f(w,r_i)-\gamma)^+ \tag{3.13}$$

Pour résoudre ce problème d'optimisation, nous introduisons des variables artificielles z_j pour remplacer $\big(f(w,r_j)-\gamma\big)^+$. Ceci est obtenu en imposant les contraintes $z_j \geqslant 0$ et $z_j \geqslant f(w,r_j)-\gamma$ et alors le problème d'optimisation (3.13) se réduit au problème de programmation linéaire (LP) suivant :

$$\begin{aligned}
\min_{w,z,\gamma}\ & \gamma + \frac{1}{(1-\alpha)J}\sum_{j=1}^{J} z_j \\
\text{s.t}\ \ & w^\top \mu \geqslant \bar{r} \\
& z_j \geqslant 0\,, \qquad\qquad j = 1,\dots,J \\
& z_j \geqslant f(w,r_j)-\gamma\,,\ j = 1,\dots,J \\
& w \in \Omega
\end{aligned} \tag{3.14}$$

Sous cette forme, ce problème peut être résolu en utilisant la méthode du simplexe ou d'autres algorithmes de programmation linéaire à même de résoudre efficacement des problèmes à très grande échelle dans une durée raisonnable. Cependant, cette minimisation de la CVaR basée sur des scénarios n'est pas un estimateur robuste car l'un des paramètres incertains du modèle, en l'occurrence μ.

Les moindres petites différences dans les estimations de ce paramètre engendre d'importants changements dans la composition d'un portefeuille optimal. Un moyen de réduire la sensibilité du modèle à l'estimation des paramètres consiste à utiliser une optimisation robuste pour déterminer le portefeuille optimal. À défaut supposer la connaissance précise de la distribution du vecteur aléatoire r, l'idée était de plutôt de supposer la fonction de densité $p(.)$ (Eqs 3.7 & 3.8) appartient à un certain ensemble $\mathcal{P}$ de distributions.

3.2.2 Minimisation de la Worst-Case CVaR

Soient n actifs risqués $(A_1, \ldots, A_n)$, $r = (r_1, \ldots, r_n)^\top$ le vecteur du log des rendements aléatoires, $\mu = (\mu_1, \ldots, \mu_n)^\top$ le vecteur des rendements attendus, $\Sigma_{n\times n}$ la matrice de covariance et $w = (w_1, \ldots, w_n)^\top$ le vecteur de pondérations du portefeuille. Aussi, nous supposons que $w \in \Omega$, où Ω est un ensemble réalisable et que la fonction de perte de rendement du portefeuille $f(w, r)$ est une fonction convexe (linéaire) donnée par $f(w, r) = -w^\top \mu$.

Définition 3.2.1 (Ambiguïté dans la distribution) *La variable aléatoire ξ suppose une distribution de*

$$\mathbb{D} = \left\{\pi \mid E_\pi[X] = \bar{\mu},\ Cov_\pi = \bar{\Gamma} > 0\right\} \tag{3.15}$$

où $\bar{\mu}$ et $\bar{\Gamma}$ sont donnés.

En supposant maintenant que nous ne disposons pas d'informations précises sur la distribution du vecteur aléatoire r, toutefois nous savons que sa densité appartient à une famille de distributions $\mathcal{P}$ définie par

$$\mathcal{P} = \left\{r \mid E[r] = \mu, Cov(r) = \Sigma\right\} \tag{3.16}$$

où Σ est une matrice définie positive.
Avec différentes définitions de l'optimisation robuste dans la littérature est scindable en deux catégories : l'optimisation robuste absolue et l'optimisation robuste relative. Également appelée optimisation du pire cas (worst-case optimization), l'optimisation robuste absolue suppose que la décision finale n'est influencée que par le pire cas d'un ensemble de paramètres d'entrée incertain. L'optimisation robuste relative, quant à elle, considère le problème d'optimisation en tenant compte à la fois de la décision la plus défavorable et aussi de la meilleure décision possible pour chaque réalisation des paramètres d'entrée.

Les travaux de Zhu and Fukushima (2009) rentrent dans le cadre de l'optimisation robuste absolue. Ils considèrent que la distribution de probabilité $p(.)$ du vecteur aléatoire des rendements r est seulement connue pour appartenir à un certain ensemble $\mathcal{P}$ qui définit le worst-case CVaR (WCVaR) comme la CVaR lorsque la distribution de probabilité la plus défavorable dans l'ensemble $\mathcal{P}$ survient.

Définition 3.2.2 *Étant donné un niveau de confiance α, $\alpha \in [0, 1]$, la CVaR dans le pire des cas (worst-case CVaR) pour $w \in \Omega$ fixé par rapport à l'ensemble d'incertitude $\mathcal{P}$ est définie comme*

$$\begin{aligned} WCVaR_\alpha(w) &= \sup_{p(.)\in\mathcal{P}} CVaR_\alpha(w) \\ &= \sup_{p(.)\in\mathcal{P}} \min_{\gamma\in\mathbb{R}} F_\alpha(w, \gamma) \end{aligned} \tag{3.17}$$

Dans leur volonté d'éviter les hypothèses sur les distributions des rendements, Zhu and Fukushima (2009) étaient arrivé à la conclusion que la distribution de r appartient à un ensemble d'ambiguïté de distribution mixte (mixture distribution) de combinaison convexe de distributions prédéterminées. Supposer que la fonction de densité de r est seulement connue pour appartenir à un ensemble de distributions qui se compose de toutes les distributions mixtes de certains scénarios de distribution possibles peut s'écrire :

$$p(.) \in \mathcal{P}_M = \Big\{ \sum_{i=1}^{K} \lambda_i p^i(.) \ : \ \sum_{i=1}^{K} \lambda_i = 1, \ \lambda_i \geqslant 0, \ i = 1, \ldots, K \Big\} \tag{3.18}$$

où $p^i(.)$ désigne le i-ième scénario de distribution, et K désigne le nombre de scénarios possibles.
Nous pouvons alors ré-écrire l'équation (3.8) comme

$$\begin{aligned} F_\alpha(w, \gamma, \lambda) &= \gamma + \frac{1}{1-\alpha} \int_{r \in \mathbb{R}^n} \big(f(w,r) - \gamma\big)^+ \sum_{i=1}^{K} \lambda_i p^i(r) dr \\ &= \sum_{i=1}^{K} \lambda_i F_\alpha^i(w, \gamma) \end{aligned} \tag{3.19}$$

où

$$F_\alpha^i(w, \gamma) = \gamma + \frac{1}{1-\alpha} \int_{r \in \mathbb{R}^n} \big(f(w,r) - \gamma\big)^+ p^i(r) dr \ , \ i = 1, \ldots, K$$

Théorème 3.2.2 (Zhu and Fukushima (2009)) *Pour chaque w, $WCVaR_\alpha(w)$ par rapport à $\mathcal{P}_M$ est donné par*

$$WCVaR_\alpha(w) = \min_{\gamma \in \mathbb{R}} \max_{i \in \mathcal{L}} F_\alpha^i(w, \gamma) \tag{3.20}$$

où $\mathcal{L} = \{1, \ldots, K\}$.

D'après le théorème 3.2.2, on obtient immédiatement le corollaire suivant.

Corollaire 3.2.3 *La minimisation de $WCVaR_\alpha(w)$ sur Ω peut être obtenue en minimisant $F_\alpha^{\mathcal{L}}(w, \gamma)$ sur $\Omega \times \mathbb{R}$, c'est-à-dire*

$$\min_{w \in \Omega} WCVaR_\alpha(w) = \min_{(w,\gamma) \in \Omega \times \mathbb{R}} F_\alpha^{\mathcal{L}}(w, \gamma) \tag{3.21}$$

où $F_\alpha^{\mathcal{L}}(w, \gamma) = \max_{i \in \mathcal{L}} F_\alpha^i(w, \gamma)$ et plus précisément, si (w^*, γ^*) atteint le minimum de droite, alors w^* atteint le minimum de gauche et γ^* atteint le minimum de $F_\alpha^{\mathcal{L}}(w^*, \gamma^*)$ et vice versa.

L'utilisation des distributions mixtes en finance n'est pas nouvelle et plus particulièrement de distributions mixtes de copules. Quelques années auparavant, Hu (2006) avait déjà exploré plusieurs modèles de structures de dépendance dont particulièrement les modèles mixtes de copules normale, gumbel et gumbel de survie. Cette approche a permis aux auteurs Kakouris and Rustem (2014) de faire une extension du cadre Worst Case CVaR (WCVaR) en introduisant les distributions mixtes de copules archimédiennes. Notre étude rentre dans ce nouveau et pour ce faire nous définirons certains concepts.

3.2.3 Minimisation de la Worst Case Copula CVaR

Copules

Définie comme une fonction permettant de construire des fonctions de distribution multivariées en reliant leurs fonctions de distribution marginale.La copule est un outil statistique utilisée pour modéliser la structure de dépendance d'une distribution conjointe, indépendamment des distributions marginales. Elle a été introduite pour la première fois dans le théorème de Sklar (1959) qui énonce ce qui suit :

Théorème 3.2.4 (Sklar (1959)) *Soit F une fonction de distribution de dimension n avec des marginales $F_1, \ldots, F_n$. Alors il existe une unique copule C telle que pour tout $x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$,*

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \tag{3.22}$$

Si $F_1, \ldots, F_n$ sont continues, alors C est unique, sinon, C est uniquement déterminé sur $Ran\, F_1 \times \ldots \times F_n$. Inversement, si C est une copule et $F_1, \ldots, F_n$ sont des fonctions de distribution, alors la fonction F est une fonction de distribution jointe avec les marginales $F_1, \ldots, F_n$.

Corollaire 3.2.5 *Soient F une fonction de répartition à n dimensions à marginales continues $F_1, \ldots, F_n$, et une copule C. Donc, pour tout $u = (u_1, \ldots, u_n)$ dans $[0,1]^n$*

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)) \tag{3.23}$$

où $u_i = F_i^{-1}$, $i = 1, \ldots, n$ sont les quasi-inverses des marginales.

Parce que la fonction de distribution F de l'équation (3.22) est n-fois différentiable, en la différenciant et en utilisant l'équation (3.23), nous pouvons déduire une relation entre les fonctions de densité de probabilité et les copules. On peut définir la densité de copule d'une copule de dimension n comme

Définition 3.2.3 *Soient f la fonction de densité de probabilité multivariée de la distribution de probabilité F et $f_1, \ldots, f_n$ les fonctions de densité de probabilité univariées des marginales $F_1, \ldots, F_n$. La fonction de densité d'une copule C de dimension n est la fonction $c : U[0,1]^n \to [0, \infty)$ telle que*

$$c(u_1, \ldots, u_n) = \frac{\partial^n C(u_1, \ldots, u_n)}{\partial u_1 \ldots \partial u_n} = \frac{f(x_1, \ldots, x_n)}{\prod_{i=1}^n f_i(x_i)} \tag{3.24}$$

On constate que la copule est un outil statistique flexible et puissant pour la construction des distributions multivariées. Elle nous permet de séparer le choix de la dépendance du choix des marginales sur lesquelles aucune restriction n'est imposée. L'estimation d'une distribution multivariée se fait en deux parties :

(i) trouver la distribution marginale de pour chaque x_i

(ii) trouver la dépendance entre les données filtrées de l'étape précédente.

Avec des caractéristiques qui s'éloignent d'une distribution normale, les séries temporelles financières présentent de l'asymétrie, un excès d'aplatissement et un clustering de volatilité largement mis en évidence à travers de nombreuses études empiriques.

Afin de décrire ces caractéristiques, les modèles GJR-GARCH(1,1) seront utilisés avec des innovations de distribution de student ou distribution d'erreur généralisée (generalized error distributions) pour filtrer les données initiales. Le choix du modèle GJR-GARCH de Glosten, Jagannathan, and Runkle (1993) se justifie par la volonté de prendre en compte l'effet asymétrique dans la dynamique de la variance et capturer les queues asymétriques et épaisses. Le modèle GJR-GARCH(1,1) peut-être défini comme :

$$\begin{aligned} x_{i,t} &= \mu_{i,t} + \epsilon_{i,t} \\ \epsilon_{i,t} &= h_{i,t}^{\frac{1}{2}} z_{i,t} \\ \sigma_{i,t}^2 &= h_{i,t} = \alpha_0 + \left(\alpha_1 + \gamma_1 \mathbb{1}_{\epsilon_{i,t-1}}\right) \epsilon_{i,t-1}^2 + \beta_1 \sigma_{i,t-1}^2 \end{aligned} \tag{3.25}$$

où $x_{i,t}$, $\mu_{i,t}$ et $z_{i,t}$ désignent respectivement la série de rendements i, son espérance et $z_{i,t}$ les innovations distribuées selon les distributions de student ou ged.

Ensuite, les distributions marginales des résidus filtrés à partir du processus GJR-GARCH(1,1) sont ajustées avec une fonction de distribution cumulative utilisant les fonctions des copules. Il existe deux grandes familles de copules distinctes pour modéliser : les copules archimédiennes et les copules elliptiques. Parmi les copules elliptiques on a :

- la copule gaussienne est définie par :

$$C_\Sigma^N(u_1, \ldots, u_n) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))$$

  pour $u_i \in [0,1]$, $\forall i = 1, \ldots, n$ et avec Σ la matrice de corrélation linéaire.

- la copule de student est définie par :

$$C_{\nu,\Sigma}^T(u_1, \ldots, u_n) = t_{\nu,\Sigma}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_n))$$

  pour $u_i \in [0,1]$, $\forall i = 1, \ldots, n$, avec ν et Σ respectivement le degré de liberté et la matrice de corrélation linéaire.

Quant aux copules archimédiennes, il en existe plusieurs mais nous nous limiterons à ceux que nous utiliserons. On a :

- la copule de Clayton est définie par

$$C_\alpha^C(u_1, \ldots, u_n) = \Big(\sum_{i=1}^n u_i^{-\alpha} - (n-1)\Big)^{-\frac{1}{\alpha}}$$

  pour $u_i \in [0,1]$, $i = 1, \ldots, n$, $\alpha > 0$ sachant que le générateur est donné par $\varphi(u) = u^{-\alpha} - 1$ donc $\varphi^{-1}(u) = (u+1)^{-\frac{1}{\alpha}}$, il est complètement monotone si $\alpha > 0$.

- la copule de Frank est définie par

$$C_\alpha^F(u_1,\ldots,u_n) = -\frac{1}{\alpha}\ln\left(1+\frac{\prod_{i=1}^n(e^{-\alpha u_i}-1)}{(e^{-\alpha}-1)^{n-1}}\right)$$

pour $u_i \in [0,1]$, $i=1,\ldots,n$, $\alpha>0$ et le générateur est donné par $\varphi(u)=\ln\left(\frac{e^{-\alpha u}-1}{e^{-\alpha}-1}\right)$, donc $\varphi^{-1}(u)=-\frac{1}{\alpha}\ln\left(e^u(e^{-\alpha}-1)\right)$ il est complètement monotone si $\alpha>0$.

- la copule de Gumbel est définie par :

$$C_\alpha^G(u_1,\ldots,u_n) = \exp\left(-\left(\sum_{i=1}^n -\ln(u_i)^\alpha\right)^{\frac{1}{\alpha}}\right)$$

pour $u_i \in [0,1]$, $i=1,\ldots,n$, $\alpha \geqslant 1$ et le générateur est donné par $\varphi(u)=(-\ln(u)^\alpha)^{-\alpha}$, donc $\varphi^{-1}(u)=\exp\left(-t^{\frac{1}{\alpha}}\right)$, il est complètement monotone si $\alpha>1$.

Cependant, l'utilisation d'une seule copule pour la modélisation financière peut également être problématique à cause des limites qu'elles présentent. La copule elliptique ne rend correctement compte de la structure de dépendance que si la distribution marginale est elliptique. Par conséquent, l'utilisation de la corrélation linéaire comme mesure de dépendance est réductrice car elle ne parvient pas à saisir les dépendances non linéaires.

De plus la copule gaussienne néglige la dépendance de queue inférieure et supérieure, un aspect très important qui fait référence à la probabilité de baisse ou de hausse simultanée. Alors que la copule de student a une dépendance de queue symétrique (inférieure et supérieure) et donc pas nulle, les résultats empiriques suggèrent plutôt une structure de dépendance asymétrique. C'est ce qui explique les applications réussites avec les copules archimédiennes comme les copules de Clayton pour la dépendance de queue inférieure et de Gumbel la dépendance de queue supérieure. Bien que la copule de Frank, comme les autres copules archimédiennes, ait la spécificité de prendre en compte les dépendances non linéaires, elle présente des caractéristiques similaires à la copule gaussienne.

La difficulté est maintenant de choisir une copule appropriée pour un problème donné selon certains critères (tels que les critères d'information AIC ou BIC), mais le meilleur ajustement n'est pas garanti.
La copule mixte (mixture copula) est une combinaison linéaire de plusieurs copules individuelles qui peut être utilisée pour générer des structures de dépendance qui n'appartiennent pas aux familles de copules existantes. Elle nous permet de combiner les copules en fonction des caractéristiques que nous voulons, dans notre cas, les copules sont utilisées pour les dépendances de la queue inférieure, de la queue supérieure et du centre de distribution. Mathématiquement, une fonction de copule mixte est formulée comme :

$$C(u_1,\ldots,u_n,\theta) = \sum_{l=1}^L \lambda_l C_l(u_1,\ldots,u_n,\theta_l) \tag{3.26}$$

où $\{C_1(.),\ldots,C_L(.)\}$ est un ensemble de copules de base, qui est une séquence de copules connues avec des paramètres inconnus $\{\theta_l\}_{l=1}^L$, $\{\lambda_l\}_{l=1}^L$ sont les poids satisfaisant $0 \leqslant \lambda_l \leqslant 1$, $\sum_{l=1}^L \lambda_l = 1$ et L est le nombre de copules candidates.

Cai and Wang (2014) ont proposé un algorithme EM pour trouver l'estimateur de vraisemblance pénalisé (penalized likelihood estimator) afin d'estimer simultanément les poids et les paramètres de copules.

Worst Case Copula CVaR

Maintenant que nous avons énoncer les concepts nécessaires au cadre théorique de l'extension du worst case CVaR avec les copules mixtes de Kakouris and Rustem (2014), on a :

Proposition 3.2.6 *Soient $u = (u_1, \dots, u_n) \in [0,1]^n$ un vecteur aléatoire, $\tilde{h}(w,u)$ la fonction de perte et $F(x) = (F_1(x), \dots, F_l(x))$ un ensemble de distributions marginales où $u = F(x)$. Supposons également que u suit une distribution continue avec une fonction de densité de copules $c(.)$. Alors VaR_α et $CVaR_\alpha$ pour un niveau de confiance α sont définis comme*

$$VaR_\alpha = arg\ \min\{\gamma \in \mathbb{R}\ :\ C(u|\tilde{h}(w,u) \leqslant \gamma) \geqslant \alpha\} \tag{3.27}$$

$$CVaR_\alpha = \frac{1}{1-\alpha} \int_{\tilde{h}(w,u) \geqslant VaR_\alpha(w)} \tilde{h}(w,u)c(u)du \tag{3.28}$$

En utilisant l'équation (3.8), on peut écrire

$$F_\alpha(w,\gamma) = \gamma + \frac{1}{1-\alpha} \int_{u\in[0,1]^n} \left(\tilde{h}(w,u) - \gamma\right)^+ c(u)du \tag{3.29}$$

De façon similaire à l'équation (3.19)

$$\begin{aligned} F_\alpha(w,\gamma,\lambda) &= \gamma + \frac{1}{1-\alpha} \int_{u\in[0,1]^n} \left(\tilde{h}(w,u) - \gamma\right)^+ \sum_{i=1}^{L} \lambda_i c_i(u)du \\ &= \sum_{i=1}^{L} \lambda_i F_\alpha^i(w,\gamma) \end{aligned} \tag{3.30}$$

Puisque la copule de la distribution de x ne peut pas être exacte, on suppose qu'elle appartient à un ensemble de copules $c(.) \in \mathcal{C}_M$

$$c(.) \in \mathcal{C}_M = \{\sum_{i=1}^{L} \lambda c_i(.)\ :\ \sum_{i=1}^{L} \lambda_i = 1\ ,\ i = 1, \dots, L\}$$

Définition 3.2.4 *La WCVaR pour $w \in \Omega$ fixé par rapport à $\mathcal{C}_M$ est définie comme :*

$$WCVaR_\alpha(w) = \sup_{c(.)\in\mathcal{C}_M} CVaR_\alpha(w) \tag{3.31}$$

Suivant Zhu and Fukushima (2009), on peut obtenir la WCVaR pour les copules comme le problème de minimisation linéaire suivant :

$$\begin{aligned} \min_{w\in\mathbb{R}^n, v\in\mathbb{R}^K, \lambda\in\mathbb{R}}\ & \gamma + \frac{1}{(1-\alpha)S_i} \sum_{i=1}^{S_i} v_i \\ \text{s.t}\ & w \in \Omega \\ & v_i^{[i]} \geqslant \tilde{h}(w, u^{[j]}) - \gamma \\ & v_i^{[i]} \geqslant 0\ ,\ j = 1, \dots, J,\ i = 1, \dots, n \end{aligned} \tag{3.32}$$

où v_i, $i = 1, \ldots, J$, sont des variables indicatrices auxiliaires et $m = \sum_{i=1}^{L} S_i$.

3.3 Applications et résultats

Dans l'application empirique, nous utiliserons deux portefeuilles composés chacun de 30 valeurs cotées dans la zone euro. Les actions ont été sélectionnées sur la base de leur notation ESG par l'agence MSCI ESG [4] comme le montre la figure 3.1. Constitué d'actions d'entreprises de différents domaines d'activité, le portefeuille 1 est composé de Best-in-class, c'est-à-dire les actions avec les meilleures notations ESG (AAA et AA), les actions du portefeuille 2 avec des notations bien inférieures (A, BBB, BB et B) .

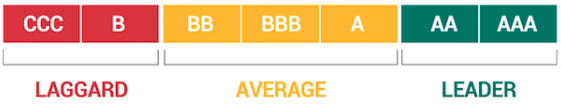


FIGURE 3.1 – Les scores ESG selon MSCI

Les différentes compositions des portefeuilles sont répertoriées aux tables 3.1 et 3.2. Les rendements logarithmiques quotidiens sont calculés comme $r_t = log(P_t/P_{t-1})$, où P_t fait référence au cours de clôture du jour t. Les données quotidiennes sont extraites de l'API Yahoo Finance `finance.yahoo.com`, couvrant un total de 1842 entrées du 2015-05-08 au 2022-08-12, dont les 1202 premières observations sont utilisées comme échantillon d'apprentissage et 640 pour les échantillons d'essai Dernière observation. La période de l'échantillon test va du 2020-02-07 au 2022-08-12, elle contient donc les chocs sur les marchés financiers liés à la crise sanitaire de la COVID-19 et au début de la guerre en Ukraine. Nous implémentons une fenêtre glissante (rolling window) de 1202 observations.

Par conséquent, la prévision hors échantillon est effectuée 640 fois comme si nous effectuons un rééquilibrage du portefeuille chaque jour. Les statistiques descriptives pour chaque rendement sont présentées dans les tableaux 3.3 et 3.4. Les moyennes sont presque nulles, mais 16 actions sur 30 du premier portefeuille ont une valeur moyenne négative, tandis que 7 actions sur 30 du second portefeuille ont une valeur moyenne négative. Cela remet en question la notion selon laquelle les actions avec des scores ESG plus élevés ont des rendements plus élevés. Les valeurs élevées des variances des rendements peuvent s'expliquer comme une opportunité d'obtenir des rendements élevés en bénéficiant d'importants changements à la hausse. Par conséquent, la performance du portefeuille 1, qui se compose davantage d'actions à rendement négatif, n'augure rien de bon pour l'avenir. De plus, les distributions de ces rendements logarithmiques quotidiens sont pour la plupart asymétriques négativement, à l'exception de ITX.MC, KPN.AS, OR.PA, et URW.AS pour le portefeuille 1 et ARGX.BR, CLNX.MC, ERF.PA, HO.PA pour le portefeuille 2. Les valeurs du skewness négatives suggèrent que les rendements négatifs sont plus probables. À l'exception de l'action MC.PA du portefeuille 2, toutes les séries présentent un excès de kurtosis et peuvent être plus risquées en raison de leur probabilité plus élevée de rendements extrêmes. Les valeurs du skewness et de kurtosis indiquent que les distributions de probabilité pour tous les rendements sont asymétriques et leptokurtiques, de sorte que l'hypothèse de normalité et les modèles basés sur celle-ci sont inappropriés.

4. `https://www.msci.com/our-solutions/esg-investing/esg-ratings-climate-search-tool/issuer`

nº	Company	Ticker	Industry	ESG Rating
1	Aegon	AGN.AS	Life & Health Insurance	AA
2	Alstom	ALO.PA	Construction & Farm Machinery & Heavy Trucks	AA
3	Amadeus	AMS.MC	Software & Services	AA
4	Bouygues	EN.PA	Construction & Engineering	AA
5	Caixabank	CABK.MC	Banks	AA
6	EssilorLuxottica	EL.PA	Textiles, Apparel & Luxury Goods	AA
7	Generali	G.MI	Multi-Line Insurance & Brokerage	AA
8	Heineken	HEIA.AS	Beverages	AA
9	Henkel	HEN3.DE	Household & Personal Products	AA
10	Orange	ORA.PA	Telecommunication Services	AA
11	Randstad	RAND.AS	Professional Services	AA
12	UCB	UCB.BR	Pharmaceuticals	AA
13	Akzo Nobel	AKZA.AS	Specialty Chemicals	AAA
14	Allianz	ALV.DE	Multi-Line Insurance & Brokerage	AAA
15	ASML Holding	ASML.AS	Semiconductors & Semiconductor Equipment	AAA
16	Banco Bilbao Vizcaya Argentaria	BBVA.MC	Banks	AAA
17	Danone	BN.PA	Food Products	AAA
18	Endesa	ELE.MC	Utilities	AAA
19	Industria de Diseno Textil	ITX.MC	Retail-Consumer Discretionary	AAA
20	Intesa Sanpaolo	ISP.MI	Banks	AAA
21	KBC Groep	KBC.BR	Banks	AAA
22	Kering	KER.PA	Textiles, Apparel & Luxury Goods	AAA
23	Koninklijke DSM	DSM.AS	Specialty Chemicals	AAA
24	Koninklijke KPN	KPN.AS	Telecommunication Services	AAA
25	L'oréal	OR.PA	Household & Personal Products	AAA
26	Michelin	ML.PA	Auto Components	AAA
27	SAP	SAP.DE	Software & Services	AAA
28	Schneider Electric	SU.PA	Electrical Equipment	AAA
29	Umicore	UMI.BR	Specialty Chemicals	AAA
30	Unibail-Rodamco-Westfield	URW.AS	Real Estate Management & Services	AAA

TABLE 3.1 – La composition du portefeuille 1

Le choix du modèle GJR-GARCH pour la modélisation des séries des rendements est raisonnable car elles présentent des phénomènes asymétriques et un comportement extrême. Pour évaluer la performance de ces deux portefeuilles, nous leur appliquons plusieurs modèles, du plus basique au plus complexe. D'abord le modèle $1/N$ plus connu sous le nom de portefeuille équi-pondéré, le modèle mean-CVaR, le modèle mean-CVaR avec copule gaussienne et les deux versions du modèle worst case avec copules mixtes pour $\alpha = 0.95$. En effet, parmi les deux versions du modèle worst case avec copules mixtes on a :

$$c_{mix\,1}(u) = \lambda_1 C^C(u) + \lambda_2 C^T(u) + \lambda_3 C^G(u)\ , \text{ avec } \sum_{i=1}^{3} \lambda_i = 1$$

$$c_{mix\,2}(u) = \lambda_1 C^C(u) + \lambda_2 C^F(u) + \lambda_3 C^G(u)\ , \text{ avec } \sum_{i=1}^{3} \lambda_i = 1$$

Ce sont deux structures de dépendance, la différence étant la modélisation de la dépendance au centre de la distribution.

nº	Company	Ticker	Industry	ESG Rating
1	ACS	ACS.MC	Construction & Engineering	A
2	Airbus	AIR.PA	Aerospace & Defense	BBB
3	ArcelorMittal	MT.AS	Steel	BB
4	Argenx	ARGX.BR	Biotechnology	A
5	Bayer	BAYN.DE	Pharmaceuticals	BB
6	Bayerische Motoren Werke	BMW.DE	Automobiles	A
7	Carrefour	CA.PA	Retail - Food & Staples	A
8	Cellnex Telecom	CLNX.MC	Telecommunication Services	A
9	Daimler Truck Holding	DTG.DE	Construction & Farm Machinery & Heavy Trucks	BBB
10	Dassault Aviation	AM.PA	Aerospace & Defense	B
11	Deutsche Post	DPW.DE	Air Freight & Logistics	A
12	ENI	ENI.MI	Integrated Oil & Gas	A
13	Eurofins Scientific	ERF.PA	Health Care Equipment & Supplies	A
14	Exor	EXO.MI	Diversified Financials	BBB
15	Ferrovial	FER.MC	Construction & Engineering	A
16	Groep Brussel Lambert	GBLB.BR	Diversified Financials	A
17	Hermes International	RMS.PA	Textiles, Apparel & Luxury Goods	A
18	LVMH	MC.PA	Textiles, Apparel & Luxury Goods	A
19	Recordati Industria Chimica	REC.MI	Pharmaceuticals	B
20	Safran	SAF.PA	Aerospace & Defense	BBB
21	Saint Gobain	SGO.PA	Building Products	A
22	Sanofi	SAN.PA	Pharmaceuticals	A
23	Siemens Gamesa Renewable Energy	SGRE.MC	Electrical Equipment	A
24	Sofina	SOF.BR	Diversified Financials	BB
25	Thales	HO.PA	Aerospace & Defense	A
26	TotalEnergies	TTE.PA	Integrated Oil & Gas	A
27	Veolia Environnement	VIE.PA	Utilities	A
28	Volkswagen	VOW.DE	Automobiles	B
29	Vonovia	VNA.DE	Real Estate Management & Services	A
30	Warehouses De Pauw	WDP.BR	Real Estate Management & Services	A

TABLE 3.2 – La composition du portefeuille 2

Et pour les queues inférieure et supérieure et le centre de la distribution, nous choisissons une copule qui prend en compte les dépendances non linéaires dans la version 2 du modèle. Dans la version 1, nous avons choisi la copule student qui, contrairement à la copule de Frank, ne capture que les dépendances linéaires au centre de la distribution. Il a également la capacité de modéliser les dépendances de queue de distribution. Les tableaux 3.5 et 3.6 rapportent les performances du portefeuille pour chaque modèle. Il n'est donc pas surprenant que les modèles worst case copula-CVaR 1 et worst case copula-CVaR 2 soient plus performants que les autres, tant en termes de rendement que de risque, comme le montre la figure 3.2. Au cours de la période de turbulences intenses de février-mars 2020, les plus faibles baisses ont pu être observées à l'aide de modèles robustes, bien que la plupart des données liées à leurs estimations aient été tirées pendant une période de calme considérable sur les marchés financiers. Parce que dans la littérature existante, la dépendance tant dans les queues qu'au centre des distributions est de nature non linéaire. Les résultats montrent que le modèle worst case CVaR est moins risqué, ce qui peut s'expliquer par la surpondération de certaines actions par le modèle mean-CVaR dans les figures B.1 et B.2.

assets	mean (%)	variance (%)	std	skewness	kurtosis	min	max
AGN.AS	-0.0200	0.0598	0.0245	-0.8979	12.0819	-0.2431	0.1638
AKZA.AS	-0.0037	0.0239	0.0155	-0.0830	7.6863	-0.1181	0.1222
ALO.PA	-0.0071	0.0360	0.0190	-0.2084	6.2484	-0.1214	0.1109
ALV.DE	0.0111	0.0254	0.0160	-0.6153	16.2633	-0.1664	0.1467
AMS.MC	0.0190	0.0398	0.0200	-0.1465	9.4116	-0.1756	0.1398
ASML.AS	0.0963	0.0403	0.0201	-0.2397	3.0084	-0.1311	0.1170
BBVA.MC	-0.0344	0.0521	0.0228	-0.3889	7.8268	-0.1765	0.1540
BN.PA	-0.0089	0.0177	0.0133	-0.2492	4.7563	-0.0889	0.0743
CABK.MC	-0.0197	0.0522	0.0228	-0.4013	6.6777	-0.2000	0.1396
DSM.AS	0.0593	0.0222	0.0149	-0.3567	4.5740	-0.1102	0.0848
EL.PA	0.0229	0.0256	0.0160	-0.0327	5.2690	-0.1097	0.1120
ELE.MC	0.0023	0.0204	0.0143	-1.6723	16.1031	-0.1735	0.0650
EN.PA	-0.0090	0.0335	0.0183	-0.8453	16.0112	-0.1901	0.1367
G.MI	-0.0067	0.0273	0.0165	-1.0634	14.6076	-0.1835	0.1049
HEIA.AS	0.0161	0.0185	0.0136	-0.1923	5.7979	-0.0972	0.1055
HEN3.DE	-0.0258	0.0195	0.0140	-0.9998	8.1082	-0.1203	0.0684
ISP.MI	-0.0275	0.0495	0.0222	-1.2934	16.5922	-0.2606	0.1349
ITX.MC	-0.0054	0.0298	0.0173	0.1978	5.2347	-0.1113	0.1313
KBC.BR	-0.0045	0.0458	0.0214	-1.0434	12.8547	-0.2125	0.1316
KER.PA	0.0716	0.0392	0.0198	-0.1504	4.5994	-0.1314	0.1007
KPN.AS	-0.0036	0.0236	0.0154	0.3096	9.4249	-0.1128	0.1467
ML.PA	0.0036	0.0294	0.0171	-0.3742	8.0151	-0.1567	0.1226
OR.PA	0.0412	0.0206	0.0143	0.1084	3.5782	-0.0755	0.0810
ORA.PA	-0.0194	0.0188	0.0137	-0.0932	7.0802	-0.1093	0.1066
RAND.AS	-0.0025	0.0373	0.0193	-0.6871	7.2206	-0.1512	0.1003
SAP.DE	0.0177	0.0259	0.0161	-2.0211	34.5041	-0.2477	0.1182
SU.PA	0.0382	0.0296	0.0172	-0.4526	7.1468	-0.1510	0.1134
UCB.BR	0.0094	0.0296	0.0172	-0.9626	16.9753	-0.1990	0.1395
UMI.BR	0.0272	0.0455	0.0213	-0.5886	8.3015	-0.1903	0.1396
URW.AS	-0.0732	0.0728	0.0270	0.2363	17.0583	-0.2435	0.2175

TABLE 3.3 – Statistiques descriptives des actions du portefeuille 1

assets	mean (%)	variance (%)	std	skewness	kurtosis	min	max
ACS.MC	-0.0128	0.0464	0.0215	-0.0818	18.2009	-0.2091	0.2299
AIR.PA	0.0318	0.0591	0.0243	-0.3916	15.2750	-0.2506	0.1862
AM.PA	0.0095	0.0333	0.0182	-0.1331	4.7482	-0.1213	0.0874
ARGX.BR	0.2003	0.0767	0.0277	3.2898	43.7804	-0.1255	0.4469
BAYN.DE	-0.0456	0.0355	0.0188	-0.6867	6.9563	-0.1511	0.0838
BMW.DE	-0.0155	0.0343	0.0185	-0.3036	6.6526	-0.1389	0.1352
CA.PA	-0.0315	0.0300	0.0173	-0.1215	9.0860	-0.1407	0.1259
CLNX.MC	0.0617	0.0320	0.0179	0.0084	4.0091	-0.1084	0.1154
DPW.DE	0.0170	0.0273	0.0165	-0.2148	6.9070	-0.1281	0.1173
DTE.DE	0.0070	0.0191	0.0138	-0.3552	5.1526	-0.1127	0.0631
ENI.MI	-0.0199	0.0351	0.0187	-1.5724	23.6664	-0.2339	0.1392
ERF.PA	0.0600	0.0426	0.0206	0.2360	5.6201	-0.1047	0.1604
EXO.MI	0.0248	0.0471	0.0217	-0.1666	6.8945	-0.1368	0.1956
FER.MC	0.0169	0.0264	0.0162	-0.4844	8.8307	-0.1333	0.1277
GBLB.BR	0.0047	0.0177	0.0133	-2.8478	45.8518	-0.2238	0.0651
HO.PA	0.0443	0.0283	0.0168	0.2531	7.9032	-0.1195	0.1285
MC.PA	0.0811	0.0309	0.0176	-0.0607	2.8200	-0.0908	0.0912
MT.AS	0.0136	0.0986	0.0314	-0.1891	3.9068	-0.2046	0.1703
REC.MI	0.0474	0.0313	0.0177	-0.4186	13.8624	-0.1799	0.1716
RMS.PA	0.0773	0.0230	0.0152	-0.2005	3.6398	-0.0921	0.0800
SAF.PA	0.0284	0.0534	0.0231	-0.6563	21.3805	-0.2597	0.1901
SAN.PA	-0.0027	0.0191	0.0138	-0.1862	3.1093	-0.0853	0.0612
SGO.PA	0.0072	0.0363	0.0190	-0.9600	12.6816	-0.1876	0.1126
SGRE.MC	0.0161	0.0708	0.0266	-0.3731	8.0477	-0.1942	0.1724
SOF.BR	0.0448	0.0196	0.0140	-0.7587	10.2213	-0.1151	0.1067
TTE.PA	0.0057	0.0340	0.0184	-0.6341	15.2506	-0.1816	0.1404
VIE.PA	0.0184	0.0298	0.0173	-0.6749	9.4663	-0.1753	0.1157
VNA.DE	0.0152	0.0251	0.0159	-0.1885	4.7375	-0.0888	0.1063
VOW.DE	-0.0060	0.0545	0.0233	-0.4127	9.9779	-0.1880	0.1469
WDP.BR	0.0696	0.0257	0.0160	-0.1160	9.8568	-0.1293	0.1119

TABLE 3.4 – Statistiques descriptives des actions du portefeuille 2

Indicators	Equally weighted	mean-CVaR	Gaussian Copula-CVaR	Worst Case Copula-CVaR 1	Worst Case Copula-CVaR 2
Annualized Return	-0.0687	-0.0846	-0.0522	-0.0421	-0.0424
Annualized Standard Deviation	0.2431	0.1797	0.1830	0.1827	0.1850
Historical VaR 5%	-0.0239	-0.0169	-0.0182	-0.0186	-0.0175
Cornish-Fisher VaR 5%	-0.0261	-0.0202	-0.0206	-0.0203	-0.0204
CVaR 5%	-0.0397	-0.0294	-0.0304	-0.0297	-0.0305
Cornish-Fisher CVaR 5%	-0.0688	-0.0582	-0.0572	-0.0540	-0.0582
Max Drawdown (%)	37.9360	32.1363	35.1418	33.8510	33.6669
Conditional Drawdown 5%	32.2995	27.3399	33.3847	32.1584	28.6256
Annualized Sharpe Ratio	-0.2826	-0.4709	-0.2851	-0.2303	-0.2290
Omega Ratio	0.9675	0.9264	0.9626	0.9733	0.9736
Burke ratio	-0.1762	-0.2482	-0.1403	-0.1144	-0.1149
Kappa	-0.0325	-0.0736	-0.0374	-0.0267	-0.0264

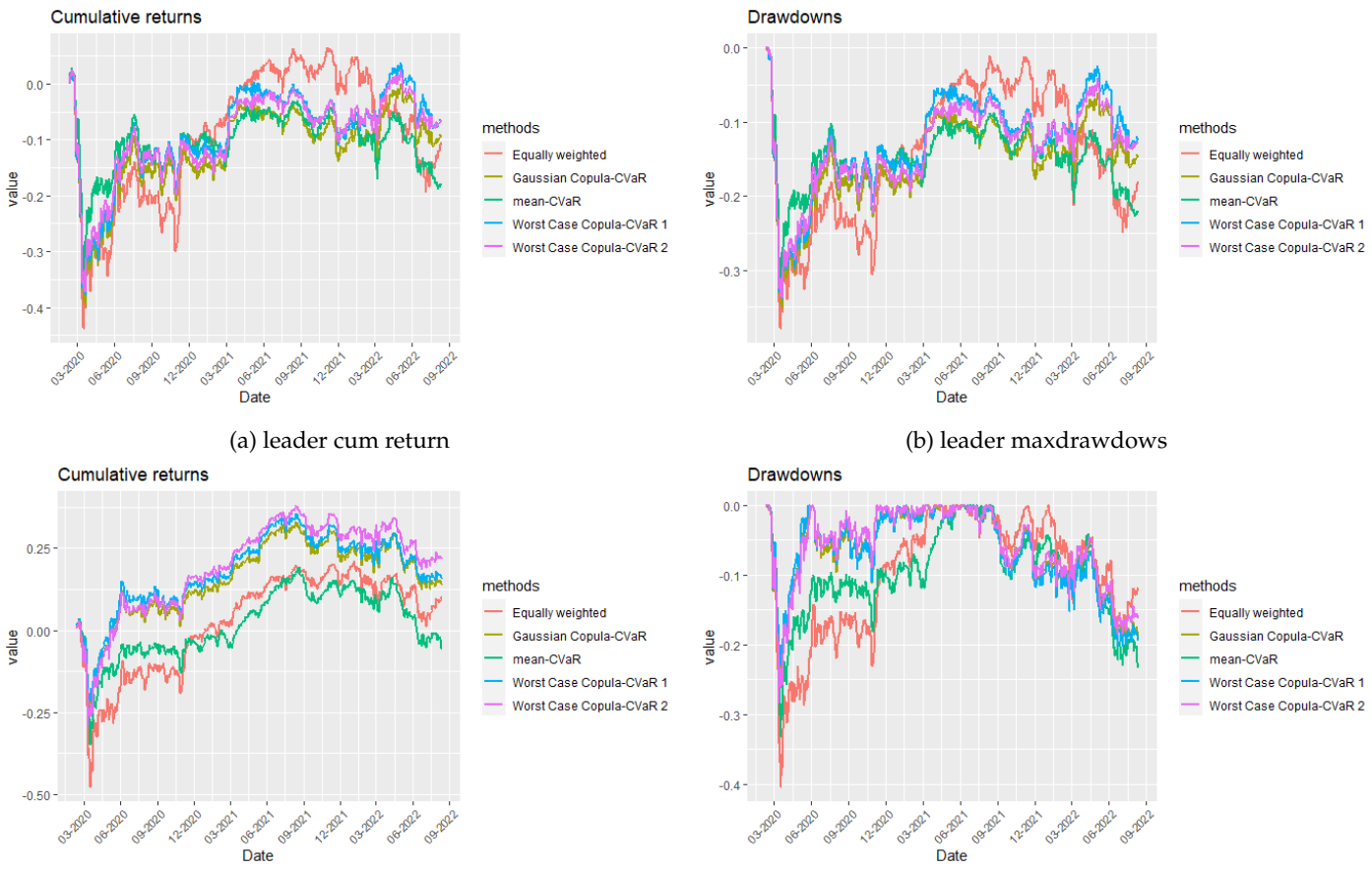
TABLE 3.5 – Les performances du portefeuille 1

Indicators	Equally weighted	mean-CVaR	Gaussian Copula-CVaR	Worst Case Copula-CVaR 1	Worst Case Copula-CVaR 2
Annualized Return	0.0086	-0.0417	0.0348	0.0418	0.0674
Annualized Standard Deviation	0.2460	0.2010	0.2001	0.1994	0.1996
Historical VaR 5%	-0.0218	-0.0187	-0.0188	-0.0184	-0.0193
Cornish-Fisher VaR 5%	-0.0265	-0.0215	-0.0214	-0.0212	-0.0212
CVaR 5%	-0.0402	-0.0305	-0.0312	-0.0306	-0.0308
Cornish-Fisher CVaR 5%	-0.0748	-0.0595	-0.0510	-0.0495	-0.0549
Max Drawdown (%)	40.4983	33.2299	26.5140	25.3916	28.2688
Conditional Drawdown 5%	13.8931	17.8249	8.7332	10.0239	4.9511
Annualized Sharpe Ratio	0.0349	-0.2076	0.1741	0.2098	0.3377
Omega Ratio	1.0317	0.9794	1.0520	1.0580	1.0836
Burke ratio	0.0194	-0.1126	0.0934	0.1150	0.1827
Kappa	0.0317	-0.0206	0.0520	0.0580	0.0836

TABLE 3.6 – Les performances du portefeuille 2

En l'absence de méthode adéquate pour estimer les paramètres d'entrée du modèle, nous observons que les poids sont concentrés sur les valeurs ayant les rendements attendus les plus élevés. La robustesse corrige l'incertitude sur la distribution et engendre des rendements élevés avec un faible risque de baisse (downside risk). Concernant la comparaison du portefeuille 1, composé uniquement d'actifs avec les meilleures notations ESG, et du portefeuille 2, qui a une composition plus diversifiée, les résultats ne sont pas surprenants.

L'analyse statistique descriptive nous donne une première indication de ces résultats, puisqu'un portefeuille est simplement une agrégation d'actions individuelles. Ainsi, résumer les actions financières avec des rendements moyens négatifs et une variance élevée ne peut que conduire à ce résultat. Dans la figure 3.3, nous pouvons voir que le portefeuille 1 est moins performant que le portefeuille 2 sur tous les modèles. L'analyse de la figure 3.2 montre clairement que les allocations selon le modèle Worst Case Copula CVaR (WCVaR) surperforment systématiquement les autres stratégies en termes de rentabilité, quel que soit le portefeuille.
De plus, une valeur plus élevée du portefeuille 2 n'implique aucunement un risque plus élevé, bien au contraire, surtout pour le modèle worst-case CVaR. Par ailleurs, par rapport à la stratégie de diversification naïve ($1/n$) au portefeuille 2, le modèle mean-CVaR a des rendements annuels inférieurs. C'est à ce moment que nous nous rendons compte de l'impact de l'incertitude associée à la distribution des rendements et de l'importance de considérer les structures de dépendance.



(a) leader cum return

(b) leader maxdrawdows

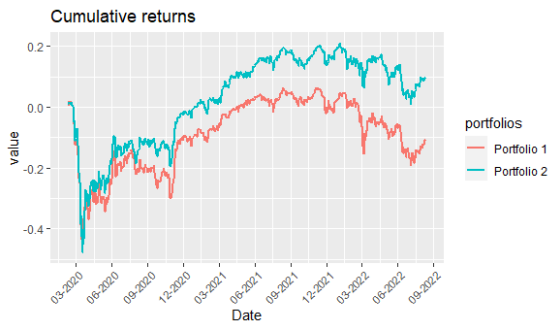
(c) mix cum return

(d) mix maxdrawdows

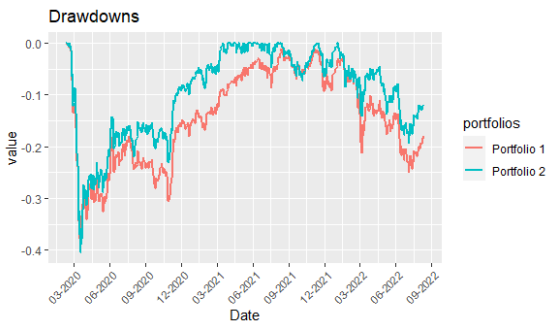
FIGURE 3.2 – Les différents modèles

L'évaluation de la performance et du risque de deux portefeuilles en utilisant uniquement le ratio de Sharpe peut fausser l'interprétation en raison de la non-normalité. Car le ratio de Sharpe peut être sensible aux valeurs extrêmes, puisqu'il utilise l'écart type. D'autre part, le ratio Oméga utilise une approche plus robuste pour mesurer le risque en considérant la distribution des rendements positifs et négatifs. Les modèles worst case mean-CVaR soulignent cette constatation car ces modèles peuvent générer des niveaux de rendement qui dépassent les niveaux de risque. Les résultats obtenus suggèrent que les actions de la zone euro ont une structure de dépendance asymétrique, car la baisse simultanée de presque toutes les actions est caractéristique des dépendances à queue basse.

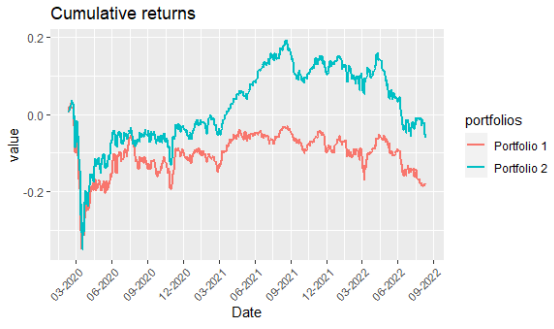
D'un point de vue pratique, cela signifie que les acteurs financiers apparaissent plus sensibles aux mauvaises nouvelles du marché qu'aux bonnes nouvelles. D'autre part, nous pouvons également remarquer une dépendance non linéaire au centre et à la queue supérieure de la distribution. C'est pourquoi la flexibilité offerte par la copule mixte est largement conforme à nos attentes, car nous avons le potentiel d'effectuer des ajustements en fonction de la dynamique des marchés financiers. Par conséquent, l'adoption d'un modèle des copules mixtes Clayton, Frank et Gumbel dans notre étude nous donne entière satisfaction conformément à d'autres résultats empiriques. Par conséquent, sur la base d'autres résultats empiriques, nous sommes entièrement satisfaits de l'utilisation du modèle de copule mixte Clayton, Frank et Gumbel dans notre étude.



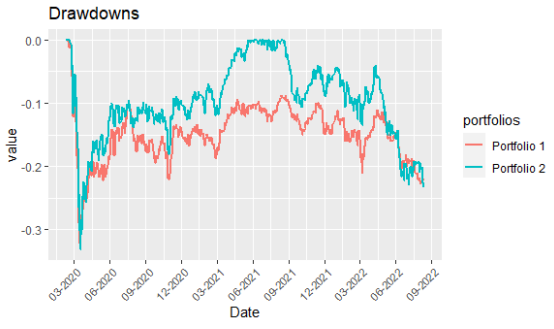
(a) $1/n$



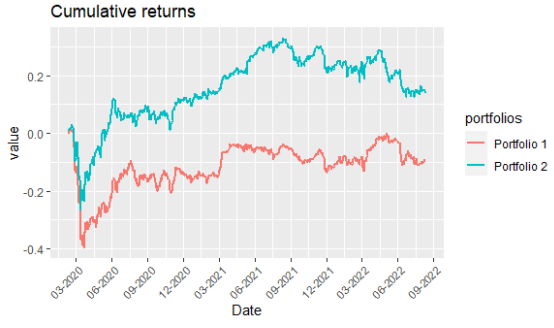
(b) $1/n$



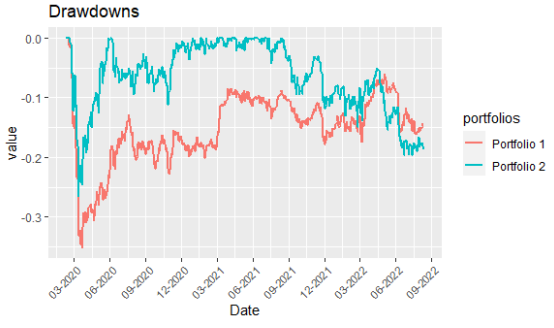
(c) mean-cvar



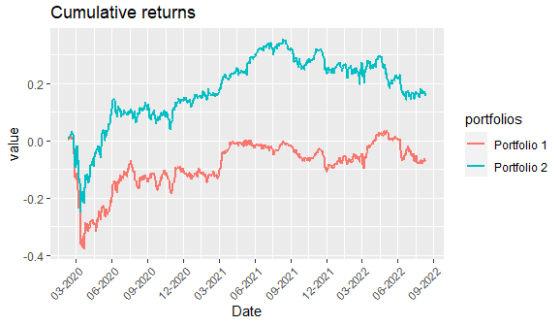
(d) mean-cvar



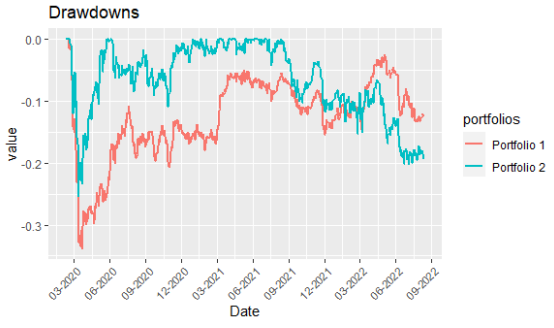
(e) gaussian copula mean-cvar



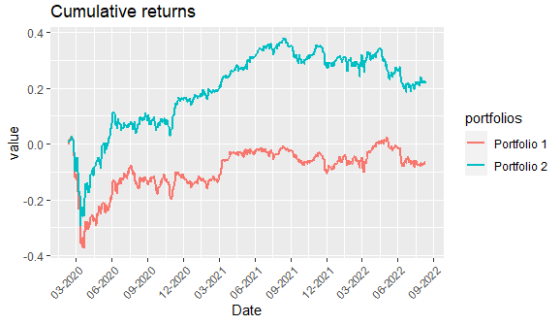
(f) gaussian copula mean-cvar



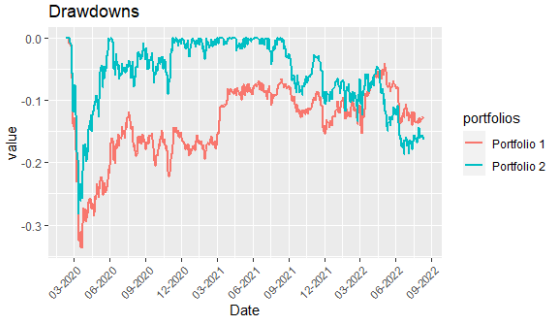
(g) Worst Case Copula mean-cvar 1



(h) Worst Case Copula mean-cvar 1



(i) Worst Case Copula mean-cvar 2



(j) Worst Case Copula mean-cvar 2

FIGURE 3.3 – Les performances des portefeuilles selon les approch

3.4 Conclusion

On assiste à une profonde évolution des préférences des investisseurs, avec des priorités qui ne se limitent plus à la performance financière, mais aussi aux activités des entreprises allouant des fonds, notamment avec la menace du réchauffement climatique et des scandales potentiels. Les critères ESG commencent à devenir des contraintes souvent imposées aux fonds et gérants de portefeuille face à l'intérêt croissant des investisseurs pour les actions visant à améliorer leur comportement sur les questions environnementales, sociales et de gouvernance (ESG). La difficulté réside dans le respect de ces deux objectifs potentiellement contradictoires. En ce sens, parmi les différentes méthodes de sélection d'actifs pour la construction de portefeuilles, nous mettons en avant les méthodes best-in-class. Cette méthode de sélection des valeurs avec les meilleures notations ESG suppose implicitement que ces valeurs performent mieux, donc leur agrégation dans un portefeuille devrait se traduire par une surperformance. Ainsi, les actions des sociétés ayant des activités dans les secteurs de l'automobile, de l'armement, de la pharmacie et du pétrole sont naturellement exclues en raison de leur mauvaise notation. Comme le montre Hong and Kacperczyk (2009), l'exclusion (screening out) de ces actions affecte la performance du portefeuille, et ces résultats sont cohérents avec ceux de Pedersen et al. (2021) qui soutiennent l'idée que les notations ESG ne sont pas nécessairement corrélées aux performances financières. C'est ce que nous voulons démontrer à travers notre modèle de gestion de portefeuille robuste. Afin de faire face à l'incertitude de la distribution des probabilités de rendement, le problème de minimisation de worst-case CVaR permet de calculer les poids optimaux du portefeuille. Cette recherche de robustesse offre plus de flexibilité dans l'analyse des décisions de portefeuille.

En fait, l'utilisation de copules, en particulier les copules mixtes, sont des outils statistiques très puissants qui peuvent capturer efficacement la complexité des dépendances, en particulier la non-linéarité et l'asymétrie. Grâce aux copules mixtes de Clayton, Frank et Gumbel, nous avons montré l'importance de prendre en compte les structures de dépendance de modélisation, qui permettent de prédire les éventuelles pertes associées aux périodes de forte volatilité des marchés financiers. Le portefeuille qui en a résulté était le plus rentable, le moins affecté par de fortes baisses et non concentré sur quelques titres. Ce cadre de modélisation met l'accent sur la performance supérieure des portefeuilles composés d'actions avec des notations différentes car elles surperforment individuellement les autres actions. Idéalement, ce critère devrait donc être assoupli en fonction de la tolérance des investisseurs à des facteurs de risque ESG spécifiques, permettant d'inclure certaines valeurs dans leurs portefeuilles pour améliorer la performance financière. Sinon, les investisseurs qui privilégient l'intégration stricte des critères ESG doivent être prêts à accepter des pertes à plus ou moins long terme afin de respecter leurs convictions. Certaines limites de notre étude sont l'absence de coûts de transaction et une stratégie de rééquilibrage quotidien, qui dégrade les performances observées. En terme de perceptive d'extension, les copules mixtes peuvent être remplacées par des copules de vines et des copules mixtes de vines.

CHAPITRE 4

Hybridation d'ARIMA avec des modèles d'apprentissage pour une meilleure prévision des séries temporelles financières

4.1 Introduction

La prévision est un élément important de la prise de décision en finance susceptible de permettre la création de la valeur et donc la richesse de certains agents économiques. Ainsi, l'un des objectifs de la finance quantitative est de prédire le couple rendement/risque des actifs financiers dans le temps. Nous voyons que la plupart des recherches académiques sur les prévisions financières sont principalement concentrées autour de deux approches : l'analyse fondamentale et l'analyse technique. La première approche repose sur la méthode d'analyse de la performance financière de l'entreprise, grâce aux indicateurs obtenus à la suite de l'analyse financière.

En revanche, la deuxième approche consiste à utiliser des données historiques pour construire des modèles capables de prédire l'évolution des actifs financiers ou les fluctuations de prix. Nous aborderons l'analyse technique, en particulier l'analyse des séries chronologiques, pour comprendre les caractéristiques des actifs afin de mettre en œuvre diverses méthodes de prévision qui peuvent permettre aux gestionnaires de portefeuille de prendre des décisions d'investissement optimales. L'adoption de méthodes de prévision performantes peut conférer aux gestionnaires de portefeuilles d'actifs financiers un avantage concurrentiel (notamment en termes "d'arbitrage statistique"). L'idée de pouvoir prédire le cours d'un indice boursier ou d'actions individuelles a divisé la communauté scientifique. En effet, deux courants de pensée s'opposent. Alors que des chercheurs comme Shiller (1980),Lo and MacKinlay (1988) , Fama and French (1992) soutiennent l'hypothèse que le marché boursier est prévisible, d'autres chercheurs comme Sharpe (1964), Fama (1995), Malkiel (1999) soutiennent l'hypothèse inverse. Ceux qui soutiennent l'idée que les marchés financiers sont imprévisibles se basent sur l'hypothèse de l'efficience des marchés.

L'hypothèse de l'efficience de marché "*Efficient Market Hypothesis*" (EMH) stipule que sur les marchés de capitaux, les prix des instruments financiers reflètent parfaitement les informations disponibles et que toute nouvelle information est immédiatement intégrée dans les prix. Par conséquent, il n'y aurait pas d'opportunités d'arbitrage et donc pas de modèles statistiques pour aider à prédire les mouvements de prix futurs ou les relations entre les actifs financiers pour en tirer un quelconque avantage. Lo et al. (2000) rejettent cette hypothèse avec une étude empirique montrant que les prix des actifs ne sont pas des marches aléatoires mais prévisibles à "court terme". Suivant cette logique, de nombreuses méthodes ont été développées pour prévoir les séries chronologiques financières.

Les premiers modèles de prévision n'étaient pas nécessairement conçus pour des problématiques liées à la finance. Comme on peut le voir dans les travaux de Yule (1927) et Walker (1931) qui ont jeté les bases conceptuelles des modèles AR (AutoRegressive) et MA (Moving Average). Sur cette base, de nombreuses méthodes ont été développées, dont l'une des premières et l'une des plus connues, le modèle ARIMA. L'idée de modéliser les séries temporelles à l'aide d'une combinaison AR (AutoRegressive) et MA (Moving Average) a été formalisée par Box and Jenkins (1976). Ils ont formulé les propriétés statistiques et les méthodologies d'estimation d'un des modèles les plus connus, le modèle ARIMA. Le modèle d'origine repose sur l'hypothèse que les séries ou les termes d'erreur sont normalement distribués. Cette hypothèse s'explique principalement par les difficultés rencontrées lors de l'estimation des paramètres du modèle dans une distribution autre que la distribution normale.

Une hypothèse jugée trop restrictive par Tiku, Wong, Vaughan, and Bian (2000) dans de nombreuses applications. En effet la modélisation de certaines séries temporelles doit prendre en compte ses caractéristiques spécifiques. Comme par exemple, la modélisation de données présentant des caractéristiques similaires à celles d'une distribution Gamma ou Béta requiert un outil capable de les prendre en considération. C'est en ce sens que des chercheurs comme McKenzie and Hipel (1994), Hipel and McLeod (1994) ont proposé le modèle Gamma ARIMA et également Rocha and Cribari-Neto (2009) avec leur modèle béta ARMA. En outre, les séries temporelles financières présentent plusieurs caractéristiques spécifiques connues sous le terme de "faits stylisés" qui sont susceptibles de remettre en question l'applicabilité directe du modèle ARIMA. Notamment, l'hétéroscédasticité de la volatilité qui va à l'encontre de l'hypothèse de volatilité constante au fil du temps du modèle ARIMA.

Nous pouvons trouver des preuves empiriques de l'hétéroscédasticité de la volatilité à travers les travaux de Granger (1969, 1992) et cette hétéroscédasticité peut être marquée par une tendance des périodes de volatilité élevée à être suivies par des périodes de volatilité faible. On parle de clustering de volatilité. Des évènements extrêmes peuvent survenir durant les périodes de haute volatilité ce qui est une caractéristique des distributions à queues épaisses. Des événements extrêmes qui se produisent moins fréquemment ou jamais pour des distributions gaussiennes. La notion d'autocorrélation qui fait référence à la corrélation linéaire entre les valeurs présentes et futures ne peut que capturer la dépendance temporelle linéaire.

Etant dans cadre non gaussien, nous pouvons d'affirmer que la dépendance est très probablement non linéaire. Afin de tenir compte de ces faits stylisés et surmonter les limites du modèle ARIMA, (Bollerslev, 1986) a développé le modèle GARCH (Generalized Autoregressive Conditional Heteroskedasticity) qui n'est autre qu'une extension du modèle ARCH(Autoregressive Conditional Heteroskedasticity) de Engle (1982) conçu pour capturer les effets d'hétéroscédasticité conditionnelle dans les séries temporelles. Le modèle GARCH peut également être critiqué pour sa structure quadratique qui ne lui permet pas de différencier les chocs positifs et négatifs sur la volatilité alors que les chocs négatifs sont censés augmenter la volatilité. A cela s'ajoutent d'autres inconvénients, comme le manque de flexibilité, la non prise en compte de l'effet de levier, qui rendent ce modèle très irréaliste.

C'est pourquoi les modèles EGARCH(Nelson, 1991) et GJR-GARCH(Glosten, Jagannathan, and Runkle, 1993) semblent plus appropriés. De nombreux travaux de recherche ont été consacrés à la mise en œuvre de modèles non linéaires paramétriques pour améliorer l'exploration de la non-linéarité dans les données grâce à des formulations spécifiques. Ces modèles peuvent être mieux adaptés pour prédire certaines séries chronologiques financières, mais sont moins efficaces pour d'autres sur lesquelles certains faits stylisés ne sont pas pris en compte. À la recherche de méthodes efficaces de prévision de séries temporelles capables de prendre en compte la dynamique complexe ou non linéaire du processus de génération de revues considéré, nous utilisons le machine learning, qui a été appliqué avec succès dans d'autres domaines scientifiques.

En effet, ils sont particulièrement adaptés pour trouver des solutions précises dans des environnements caractérisés par des informations complexes, non pertinentes voire partielles. Simon (2013) définit le machine learning comme la capacité des ordinateurs à découvrir des modèles, à les généraliser et à apprendre sans être explicitement programmés. Cette innovation scientifique a inspiré les premières applications du machine learning à la prévision des prix des actifs, notamment avec l'avènement de l'automatisation et du big data dans le secteur financier. Parmi les différentes branches du machine learning, on peut distinguer plusieurs branches l'apprentissage par renforcement, non supervisé, semi-supervisé et supervisé. Notre attention sera portée sur cette dernière branche.

En raison des taux d'erreur élevés observés lorsque la normalité et la linéarité des données ne sont observés, l'apprentissage supervisé est assez plébiscité pour la prédiction des séries temporelles financières. Cela est dû à sa capacité à utiliser des données étiquetées, à modéliser des relations complexes, à s'adapter aux changements de marché, à exploiter des variables explicatives multiples et à capturer les tendances et les motifs dans les données. Dans la littérature, nous avons des auteurs comme Cao and Tay (2001); Tay and Cao (2001); Huang et al. (2005) qui ont utilisé des Support Vector Machine (SVM) pour prédire la direction des mouvements de prix de certains actifs financiers.

Le SVM a la particularité de contrôler la fonction de décision, à l'aide d'une fonction noyau.
De plus, il utilise la théorie du principe de minimisation du risque structurel pour estimer la fonction (en minimisant une borne supérieure sur l'erreur de généralisation), ce qui la rend très résistante au problème de sur-apprentissage.

Partageant certaines similarités avec le SVM dont entre autres la capacité à gérer les données bruitées et l'utilisation des fonctions noyau le "Gaussian Process Regression" (GPR) est particulièrement adapté à la modélisation des séries temporelles. Cet algorithme a la possibilité de capturer les dépendances séquentielles, les tendances, les saisons et les fluctuations et plus largement la structure temporelle intrinsèque des données essentielles pour obtenir des prévisions précises. (voir MacKay et al., 1998; Lawrence and Hyvärinen, 2005; Williams and Rasmussen, 2006; Wilson and Adams, 2013; Hernández-Lobato and Adams, 2015)

Notons que le GPR est un modèle probabiliste non paramétrique qui estime la distribution des valeurs cibles conditionnelles contrairement au SVM. Récemment, avec l'amélioration de la puissance de calcul des ordinateurs, des algorithmes tels que les réseaux de neurones artificiels (ANN) ont gagné en popularité, notamment dans la modélisation de systèmes non linéaires. Parmi les nombreux types de réseaux de neurones, les réseaux de neurones artificiels feed-forward et les réseaux de neurones récurrents sont les plus fréquemment utilisés dans le domaine financier.

Le perceptron multicouche MLP (Multi Layer Perceptron) est à l'origine issu des travaux de Rosenblatt (1961), et est un modèle de type réseau de neurones artificiels feedforward. Il a la particularité que toutes les connexions pointent dans une seule direction, et ses couches sont généralement entraînées à l'aide de l'algorithme de rétropropagation(backpropagation) statique, qui généralise la méthode de descente de gradient ou une variante de celle-ci. Des applications financières ont rapidement émergé, par exemple Trippi and DeSieno (1992) qui ont mis en évidence la performance des contrats "*futures*" S&P 500 utilisant des réseaux de neurones dans les systèmes de trading intra-journaliers.

Plusieurs travaux de recherche ont montré que les méthodes d'apprentissage sont performantes dans l'analyse et la prévision des taux de change (voir Anastasakis and Mort, 2000; Gourinchas and Tornell, 2004; Huang et al., 2004; Kim and Sheen, 2006). Contrairement aux ANN de type feed-forward artificial neural network, les ANN de type recurrent neural network permettent aux signaux d'être connectés via des boucles, constituant ainsi une mémoire efficace qui peut être utilisée pour améliorer la prévision des séries temporelles financières. Une autre amélioration pour mieux comprendre la dynamique séquentielle des séries temporelles est l'utilisation d'une technique de backpropagation dynamique appelée Backpropagation Through Time (BPTT). Comme soutenu par Sutskever et al. (2014); Chung et al. (2014), les RNN fonctionnent aussi bien que les systèmes existants. Il suffit d'illustrer son application réussie dans divers domaines du traitement séquentiel des données, tels que la reconnaissance vocale (Speech Recognition), la modélisation du langage et la génération de texte (Language Modeling and Generating Text), la prédiction d'instruments financiers (Stock Market Prediction).

Cependant, des résultats intéressants obtenus avec les réseaux de neurones récurrents ont été obtenus avec des versions complexes telles que les modèles Long Short-Term Memory (LSTM) et Gated Recurrent Unit (GRU) proposés respectivement par Hochreiter and Schmidhuber (1997) et Cho et al. (2014a) pour leur capacité à représenter des dépendances à court ou à long terme dans la chronique considérée. Ces versions sophistiquées se sont avérées extrêmement performantes sur de nombreux problèmes, y compris pour les défis de la modélisation prédictive.

Excepté les réseaux de neurones feedforward et récurrents, un autre type de réseaux de neurones commencent a se faire une place dans la modélisation financière.

Récemment, LeCun, Bottou, Bengio, and Haffner (1998) ont proposé le Convolutional Neural Network (CNN) qui est une sorte de réseau de neurones feedforward très populaire dans la vision par ordinateur et le traitement d'image pour son efficacité. Il possède des propriétés particulières qui lorsqu'elles sont transposées sur des problématiques de prévision de séries temporelles financières qui lui confèrent la capacité d'identifier des motifs récurrents et des structures locales tels que les tendances, les cycles de volatilité. Comme tous les autres modèles, celui-ci a également des inconvénients dont un singulièrement qui peut s'avérer problématique. Ils ne sont pas dotés de dépendance à long terme et donc incapable de modéliser des dépendances à plus long terme dans les séries temporelles.

Bai, Kolter, and Koltun (2018) ont introduit le Temporal Convolutional Network (TCN) pour contourner les limites du CNN pour la modélisation de séquences dans des tâches de prédiction séquentielle. Ce modèle est très flexible en raison de son aptitude d'adaptation à des séries temporelles de longueurs variables car il n'a pas besoin d'une fenêtre de temps fixe pour effectuer les prédictions. (voir Wan et al., 2019; Sen et al., 2019; Hewage et al., 2020; Cao et al., 2021).

Bien que les modèles de machine learning continuent de s'améliorer au-delà du point de saturation, ces nouveaux modèles appartiennent à l'apprentissage en profondeur (Deep Learning) et peuvent traiter de grandes quantités de données comme celles obtenues dans le cadre du trading à haute fréquence. Parmi les applications de ces modèles à des problèmes de prédiction des mouvements de prix d'instruments financiers utilisant le deep learning, nous pouvons citer Bernal et al. (2012), Xiong et al. (2015), Nelson et al. (2017), Roondiwala et al. (2017), Fischer and Krauss (2018), Kamdem et al. (2020).

Considérées comme des approximateurs de fonctions non linéaires, les architectures de réseaux de neurones peuvent produire des résultats incohérents pour les chroniques avec des structures linéairement dépendantes. Cependant, il est important de noter que les réseaux de neurones sont le plus souvent utilisés pour résoudre des problèmes non linéaires complexes mais lorsqu'il s'agit des problèmes linéaires certaines limites apparaissent. Cybenko (1989), Zhang and Benveniste (1992) ont souligné les difficultés des réseaux de neurones à approximer des fonctions linéaires qui sont liées à des problèmes de convergence de l'algorithme d'apprentissage. Au regard de certaines caractéristiques complexes des séries temporelles financières, des auteurs tels que Engle (1982), Bollerslev (1986, 1987), Granger (1991), Glosten, Jagannathan, and Runkle (1993) ont proposé d'associer les modèles ARIMA aux modèles GARCH pour capturer les dépendances temporelles non linéaires en plus des linéaires. C'est dans cet esprit que Zhang (2003) a jugé opportun de combiner les modèles ARIMA et les modèles de réseaux de neurones artificiels. Connu sous le terme modèle hybride, cette combinaison a la capaciter de bénéficier de leurs avantages respectifs.

Ce constat justifie l'émergence de modèles dits hybrides pour obtenir des modèles capables d'intégrer la linéarité (dans les tendances) et la non-linéarité (dans la volatilité). Il existe un nombre important de combinaisons possibles pour générer des modèles hybrides comme nous pouvons le constater avec les modèles ARIMA-ANN (RNN, LSTM, GRU, TDNN, BNN)(Pan et al., 2009; Ji and Chee, 2011; Kia et al., 2012; Guo et al., 2016; Choi, 2018; Fathi, 2019; Xu et al., 2022; Temur et al., 2019; Rajalakshmi et al., 2022), ARIMA-SVM(Kumar and Thenmozhi, 2014; Pan et al., 2009; Zhu et al., 2017; Ordóñez et al., 2019; Xu et al., 2020; Nawi et al., 2021).

Par ailleurs, il existe d'autres versions de modèles hybrides combinant également des modèles non linéaires comme GRU-SVMAgarap (2018) ou CNN-LSTM(Livieris et al., 2020; Li et al., 2020; Vidal and Kristjanpoller, 2020; Lu et al., 2020). Mais ces autres versions hybrides ne rentrent pas dans notre analyse. Dans la grande majorité des applications du modèle ARIMA, la distribution normale ou généralement des distributions appartenant à la famille exponentielle (gamma) ou encore la distribution béta ont été utilisé dans une volonté de réduire la difficulté de l'estimation des paramètres. L'estimation par la méthode de vraisemblance peut se révéler être une tache laborieuse si on désire passer par une démonstration entièrement analytique et requiert inévitablement une approximation numérique.

À la lumière de ces observations, nous estimons que les lois de distributions normales n'ont pas les caractéristiques des séries temporelles financières. Et nous proposons plutôt d'utiliser la distribution de Student, car il existe de nombreuses preuves dans la littérature qu'elle se rapproche des propriétés notamment la leptokurticité des queues de distributions des rendements des actifs financiers. Afin de vulgariser ces modèles d'apprentissage automatique sous-estimés en économétrie, notre article vise à analyser les performances de différents modèles, en particulier des modèles hybrides où les paramètres du modèle ARIMA sont estimés sous l'hypothèse de distribution de Student. L'ensemble de ces modèles seront confrontés à la complexité de séries temporelles de divers instruments financiers comme un indice, un actif, une paire de devise et un ETF.

La suite de l'article est organisée comme suit : les modèles utilisés sont présentés dans la section 4.2. Les résultats empiriques et une discussion sur ces résultats seront faits dans les sections 4.3 et 4.4 respectivement, et enfin, la conclusion dans la section 4.5.

4.2 Théorie

4.2.1 Les modèles de prévision traditionnels

Modèle ARIMA

Le modèle AutoRegressive Moving Average (ARMA) est issu des travaux de Box and Jenkins (1976), dont l'approche combinait les modèles de moyenne mobile et autorégressive. Le modèle suppose la stationnarité de la série temporelle et doit donc différencier la série temporelle une ou plusieurs fois jusqu'à ce que la stationnarité soit atteinte, ce qui conduit Pesaran (2015) à suggérer le modèle AutoRegressive Integrated Moving Average(ARIMA) comme généralisation du modèle ARMA. Le terme "Integrated" fait référence l'ordre d'intégration qui est le nombre d'opérations de différenciation nécessaire pour rendre la série stationnaire.

Dans ce modèle, la valeur future d'une série chronologique de données X_t est une combinaison linéaire de valeurs passées et d'erreurs passées, il est défini mathématiquement comme

$$\left(1 - \sum_{i=1}^{p} \varphi_i B^i\right)(1 - B)^d X_t = \left(1 + \sum_{j=1}^{q} \theta_i B^i\right)\epsilon_t \tag{4.1}$$

où B l'opérateur de retard, φ_i et θ_i sont les paramètres respectifs de la partie autorégressive et moyenne mobile avec d l'ordre d'intégration.
Il n'est pas rare qu'on fasse l'hypothèse que le terme d'erreur soit normalement distribué. Cependant, nous estimons que la distribution de Student fait partie des distributions les plus en adéquation avec les faits stylisés des séries temporelles financières. Notamment à cause de sa robustesse aux valeurs aberrantes parce qu'elle a des queues de distribution épaisses. De plus, elle permet de prendre en compte l'incertitude liée aux variations de la volatilité qui peut affecter l'estimation des paramètres.

Modèles de la famille GARCH

Afin d'améliorer le modèle ARIMA, qui présente certaines limites, l'idée était de le combiner avec le modèle GARCH. Elle a pour but d'assouplir cette hypothèse trop stricte d'homoscédasticité conditionnelle du terme d'erreur. L'introduction de termes autorégressifs dans les termes d'erreur conditionnels qui dépendent de la volatilité passée permet de modéliser la volatilité qui en réalité varie, connue sous le terme hétéroscédasticité conditionnelle. Elle a été développée par Bollerslev (1986), c'est un outil de modélisation de la volatilité plus flexible que celui de Engle (1982), dont il s'est inspiré. Il est défini comme :

$$\epsilon_t = \sigma_t z_t \tag{4.2}$$

$$h_t = \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{4.3}$$

avec $\omega > 0$, $\alpha_i \geqslant 0$, $\beta_j \geqslant 0$ et $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$ garantit la stationnarité.
La principale critique du modèle est que la variance conditionnelle répond exactement de la même manière à un choc négatif qu'à un choc positif de même ampleur, alors que des études empiriques montrent que les chocs négatifs ont un impact plus important sur la volatilité que les chocs positifs. D'autres modèles ont été développés pour pallier cette lacune grâce aux travaux de Nelson (1991) pour le modèle exponentiel GARCH et le modèle GJR-GARCH proposé par Glosten, Jagannathan, and Runkle (1993). Une solution plus permanente que de choisir la mauvaise distribution asymétrique est de modifier l'équation de volatilité (Eq 4.3) en incorporant un paramètre pour gérer l'asymétrie, tout comme les modèles EGARCH et GJR-GARCH et leurs équations respectives :

$$\ln(\sigma_t^2) = h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \frac{|\epsilon_{t-i}| + \gamma \epsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{4.4}$$

l'idéal serait que $\sum_{j=1}^{p} \beta_j < 1$ pour des raisons pratiques

$$\sigma_t^2 = h_t = \omega + \sum_{i=1}^{q} \left(\alpha_i + \gamma_i \mathbb{1}_{t-i}\right)\epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-i}^2 \tag{4.5}$$

où ω, $\forall i,\ \alpha_i, \gamma_i, \beta_i > 0$ garantissent la positivité et $\sum_{i=1}^{p} \alpha_i + c \sum_{i=1}^{p} \gamma_i + \sum_{i=1}^{q} \gamma_i < 1$ est nécessaire pour la stationnarité (où c est une constante dépendant de la distribution des erreurs et vaut 0.5 si la distribution est symétrique), et $\mathbb{1}_{t-i}$ est une variable indicatrice telle que

$$\mathbb{I}_{t-i} = \begin{cases} 1 & \text{si } \epsilon_{t-i} < 0 \\ 0 & \text{si } \epsilon_{t-i} \geqslant 0 \end{cases}$$

Le point de départ d'une bonne modélisation financière est sans doute d'observer des caractéristiques très particulières de la série qui s'écartent de la distribution normale.

4.2.2 Les modèles d'apprentissage avancés

Des modèles d'apprentissage automatique ont été proposés pour surmonter les limites des modèles traditionnels basés sur des hypothèses restrictives en raison de leur capacité à apprendre des modèles cachés (relations potentielles dans les données).

Modèle Support Vector Machine

Le terme "Support Vector Machine" a tendance à désigner uniquement la classification alors que la régression fait également partie de ces domaines d'application, connus sous le nom de Support Vector Regression (SVR). Développé par Vapnik (1995), le SVM appartient à la famille des réseaux de neurones caractérisés par la minimisation structurelle du risque, c'est-à-dire la minimisation d'une borne supérieure sur l'erreur de généralisation plutôt que la minimisation empirique du risque.

C'est similaire à la résolution d'un problème de programmation quadratique avec des contraintes linéaires, il n'y a donc aucun risque de minima locaux, la condition qui garantit une solution optimale unique. Cette aptitude à résoudre les problèmes de régression non linéaire n'a été possible qu'avec l'introduction de la fonction de perte ϵ-insensible [1] de Vapnik, une aptitude mise en évidence par ses performances plus élevées que les autres techniques de régression conventionnelles comme nous pouvons le voir à travers les articles de Mukherjee et al. (1997) et Müller et al. (1997).

Le problème d'optimisation pour le cas non linéaire de SVR est formulé comme décrit comme dans les articles de Smola and Schölkopf (2004),Welling (2004) :
Soit un ensemble de points de données $\{(x_i, y_i)\}_{i=1}^{n}$ (x_i est le vecteur d'entrée, y_i est la valeur espérée et n est le nombre total d'observations), la fonction de régression SVM est formulée comme suit

$$f(x) = \sum_{i=1}^{n} w_i x_i + b \tag{4.6}$$

où chaque poids w_i est défini sur l'espace d'entrée à valeur réelle $\mathbb{R}^n$, c'est-à-dire $w_i \in \mathbb{R}^n$. La taille de la marge maximale est définie par la norme euclidienne des poids et donc la planéité (flatness) dans le cas de l'équation (4.6) requiert la minimisation de la norme des poids.

1. loss ϵ-insensible function
$$L_\epsilon(d, y) = \begin{cases} |y - f(x)| - \epsilon & \text{si } |y - f(x)| \geqslant \epsilon \\ 0 & \text{sinon.} \end{cases}$$

Pour chaque paire (x_i, y_i), l'erreur peut être exprimée sous la forme

$$E_i(x_i) = y_i - (w_i x_i + b) \tag{4.7}$$

Dans la mesure qu'on autorise chaque écart E_i à être inférieur à ϵ alors l'équation (4.6) peut s'écrire

$$\begin{cases} y_i - (w_i x_i + b) & \leqslant \epsilon \\ (w_i x_i + b) - y_i & \leqslant \epsilon \end{cases} \tag{4.8}$$

Le problème de minimisation pour w peut être

$$\min \frac{1}{2}||w||^2$$

sujet à

$$\begin{cases} y_i - (w_i x_i + b) & \leqslant \epsilon \\ (w_i x_i + b) - y_i & \leqslant \epsilon \end{cases} \tag{4.9}$$

Lorsque les contraintes sont irréalisables, le modèle peut-être reformuler avec l'introduction des variables d'écart ξ_i, ξ_i^* et décrit par le problème :

$$\min \ \frac{1}{2}||w||^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

sujet à

$$\begin{cases} y_i - (w_i x_i + b) & \leqslant \epsilon + \xi & \forall \text{ i} \\ (w_i x_i + b) - y_i & \leqslant \epsilon + \xi^* & \forall \text{ i} \\ \xi_i & \geqslant 0 & \forall \text{ i} \\ \xi_i^* & \geqslant 0 & \forall \text{ i} \end{cases} \tag{4.10}$$

Modèle Gaussian Process Regression (GPR)

Le Gaussian Process Regression est une méthode de machine qui s'appuie sur les théories bayésiennes et de machine learning respectivement. Raison pour laquelle, on le décrit comme une technique de modélisation probabiliste qui donne la possibilité d'approximer une fonction inconnue à partir d'un ensemble de données observées. Il est repose sur l'hypothèse que les valeurs des fonctions sont distribuées selon un processus gaussien, une propriété qui lui permet d'estimer les valeurs non observées. L'idée sous-jacente du Gaussian Process Regression est de considérer que chaque point de données est associé à une variable aléatoire gaussienne, et les relations entre les différentes observations sont déterminées par une fonction de covariance.

Cette approche a des similarités avec le Support Vector Regression dont leur aptitude à capturer la non linéarité dans les données et surtout l'utilisation de fonctions noyau. D'autre part, elle offre certains avantages non négligeables surtout pour la modélisation des séries temporelles financières. D'abord sa capacité à capturer les structures sous-jacentes malgré la présence de perturbations aléatoires comme avec des données bruitées.

Et ensuite, sa capacité à gérer des ensembles de données de tailles variables pour prédire des séries temporelles à différentes échelles temporelles. Pour plus de détails, nous pouvons ne référer aux travaux de Bernardo, Berger, Dawid, Smith, et al. (1998), MacKay et al. (1998), Williams (1998) et Rasmussen, Williams, et al. (2006) qui principales références bibliographiques sur cette approche.

Dans le cadre du Gaussian Process, la modélisation d'une série temporelle se fait à l'aide d'un a priori gaussien paramétré par la fonction moyenne $m(\mathbf{x})$ et une fonction de covariance (CovF) $k(\mathbf{x}, \mathbf{x}')$ comme décrit ci-dessous : Soit un ensemble de données de n observations $\{(\mathbf{x}_i, y_i)|i = 1, 2, \ldots, n\}$ représentant le jeu de données d'apprentissage du modèle gaussien où

- $\mathbf{x} = [\mathbf{x}_i]_{i=1}^n$ représente la matrice input du jeu de données d'apprentissage et $x_i = [y_{t-i}, y_{t-(i+1)}, \ldots,]^\top$
- $y = \{y_i\}_{i=1}^n$ représente la séquence d'output qui est composé des scalaires $y_i = y_{t-i+1}$.

Un Gaussian process $f(\mathbf{x})$ peut être entièrement spécifié par sa fonction moyenne $m(\mathbf{x})$ et sa fonction de covariance $k(\mathbf{x}, \mathbf{x}')$ et écrites sous la forme $f(x) \rightsquigarrow \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$. La relation entre l'input et l'output est exprimée par

$$y = f(\mathbf{x}) + \epsilon \tag{4.11}$$

où ϵ suit une distribution gaussienne indépendante et identiquement distribuée avec une moyenne nulle et une variance σ^2, c'est-à-dire $\epsilon \rightsquigarrow \mathcal{N}(0, \sigma^2)$. Selon l'équation (4.11), la vraisemblance est donnée par

$$p(\mathbf{y}|\mathbf{f}) = \mathcal{N}(\mathbf{y}|\mathbf{f}, \sigma^2 I) \tag{4.12}$$

où $y = [y_1, y_2, \ldots, y_n]^\top$, $\mathbf{f} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n)]^\top$ et I désigne la matrice unitaire $n \times n$. Selon de la définition d'un processus gaussien, la distribution marginale $p(\mathbf{f})$ est donnée par une gaussienne dont la moyenne est nulle et dont la covariance est définie par une matrice de Gram K telle que

$$p(\mathbf{f}) = \mathcal{N}(\mathbf{f}|\mathbf{0}, K) \tag{4.13}$$

où $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. Étant donné que la vraisemblance (4.12) et l'a priori (4.13) suivent la distribution gaussienne, la distribution marginale de y est donnée par

$$p(\mathbf{y}) = \int p(\mathbf{y}|\mathbf{f}) p(\mathbf{f}) d\mathbf{f} = \mathcal{N}(\mathbf{f}|\mathbf{0}, K_y) \tag{4.14}$$

où $K_y = K + \sigma_n^2 I$.
Afin de faire une prédiction de la variable cible y_* pour un nouvel input $\mathbf{x}_*$, la distribution conjointe sur $y_1, y_2, \ldots, y_n, y_*$ est donné par

$$\begin{bmatrix} y \\ y_* \end{bmatrix} = \left(\begin{bmatrix} \mathbf{f} \\ y_* \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon} \\ \epsilon_* \end{bmatrix} \right) \rightsquigarrow \mathcal{N} \left(\mathbf{0}, \begin{bmatrix} K_y & \mathbf{k}_* \\ \mathbf{k}_*^\top & k_{**} + \sigma_n^2 \epsilon_* \end{bmatrix} \right) \tag{4.15}$$

où $f_* = f(\mathbf{x})$ est la valeur de la fonction latente pour le vecteur d'entrée $\mathbf{x}$ et ϵ est le bruit correspondant; $k_* = [k(\mathbf{x}_*, \mathbf{x}_1), \ldots, k(\mathbf{x}_*, \mathbf{x}_N)]^\top$ et $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$.

En utilisant les règles de conditionnement des gaussiennes, la distribution prédictive $p(y_*|y)$ est une distribution gaussienne dont la moyenne et la covariance sont données par

$$m(\mathbf{x}_*) = \mathbf{k}_*^\top K_y^{-1}\mathbf{y} \tag{4.16}$$

$$\sigma^2(\mathbf{x}_*) = k_{**} - \mathbf{k}_*^\top K_y^{-1}\mathbf{k}_* + \sigma_n^2 \tag{4.17}$$

Pour des raisons de stabilité numérique, l'inverse de la matrice de covariance K_y peut être calculé en utilisant la décomposition de Cholesky.

La fonction de covariance $k(.,.)$ est également appelée fonction noyau dans le machine learning. Parmi les fonctions de noyau couramment utilisées dans la littérature, nous avons le noyau exponentiel au carré (kernel function squared exponential (SE))(Rasmussen et al., 2006) défini comme :

$$k_{SE}(d) = \theta_1^2 \exp\left(-\frac{d^2}{2\theta_2^2}\right) \tag{4.18}$$

où d dans l'équation (4.18) est la distance euclidienne entre deux points d'inputs $d = |\mathbf{x} - \mathbf{x}'|$.
Les paramètres θ_1, θ_2 dans l'équation (4.18) sont les hyperparamètres qui doivent être optimisés. Au cours de l'apprentissage du modèle, le negative log marginalized likelihood (NLML) dans l'équation (4.19) est minimisé, de sorte que les hyper-paramètres de la matrice noyau K puissent être estimés.

$$\begin{aligned} NLML &= -\log(p(\mathbf{y}|\mathbf{x},\theta)) \\ &= -\frac{1}{2}\log|K + \sigma_n^2 I| - \frac{1}{2}y^\top (K + \sigma_n^2 I)^{-1} y - \frac{n}{2}\log(2\pi) \end{aligned} \tag{4.19}$$

L'hyper-paramètre inconnu θ dans l'équation (4.19) est déterminé en minimisant le NLML. Le problème d'optimisation pour l'estimation des paramètres s'écrit :

$$\hat{\theta} = arg\min_{\theta} -log(p(\mathbf{y}|\mathbf{x},\theta))$$

Étant donné que le NLML est une fonction convexe, il peut être optimisé par des algorithmes d'optimisation classiques, tels que la descente de gradient.

Réseaux de neurones artificiels

Afin de trouver une explication au traitement de l'information dans le fonctionnement du cerveau, les chercheurs le considèrent comme un système très complexe capable d'effectuer des tâches très complexes. En raison de sa composition, nous appelons un ensemble de cellules interconnectées un neurone. C'est de ce système biologique, notamment de la topologie de base des neurones, que s'inspirent les réseaux de neurones artificiels et permettent d'envisager des tentatives de représentation mathématique du traitement de l'information, d'où émergeront des modèles dits d'intelligence artificielle. L'un des modèles précurseurs remonte aux travaux de Rosenblatt (1957) dans les années 1950, le Perceptron est un modèle qui peut apprendre à reconnaître les formes, et de nombreuses extensions et améliorations ont été apportées depuis, telles comme le Perceptron multicouche (Multi-Layer Perceptron) et bien d'autres.

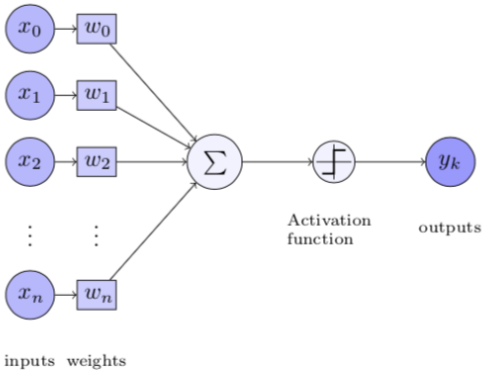
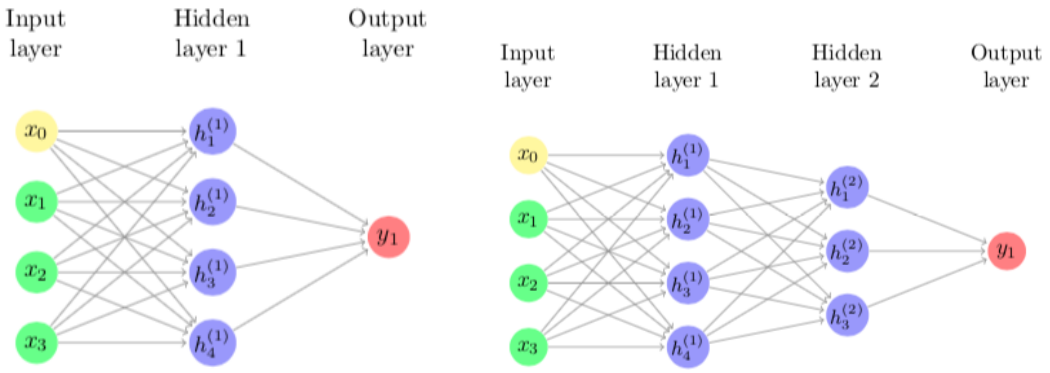


FIGURE 4.1 – Réseau neuronal avec un biais en entrée avec $x_0 = +1$ et poids $\omega_{k0} = b_k$

La figure 4.1 montre la structure d'un neurone artificiel en tant qu'une unité de traitement d'information avec les signaux d'entrée ($x_i, i = 1, \ldots, n$) qui sont connectés à la couche cachée par des poids ($\omega_i, i = 1, \ldots, n$) et cet ensemble est résumé par une fonction appelée fonction d'activation avec comme paramètre la somme pondérées des signaux d'entrée constituant ainsi le signal de sortie comme on peut le voir avec l'équation (4.20)

$$y_t = f(\sum_{i=0}^{n} \omega_i x_i) = f(\sum_{i=1}^{n} \omega_i x_i + b) \tag{4.20}$$

avec $x_0 = 1$ un biais sur les signaux d'entrée et $\omega_0 = b$.



(a) Perceptron mono couche (b) Perceptron multicouche

FIGURE 4.2 – Réseau de neurone Feedforward

Dans un cadre plus général comme nous pouvons le voir sur la figure 4.2 avec une couche d'entrée composée de n unités, p couches cachées avec un nombre d de neurones pouvant être différent d'une couche à l'autre et une couche de sortie dont sa représentation mathématique est la suivante :

$$y_t = f\Big(\sum_{j=0}^{p} \omega_{jd} f\big(\sum_{i=0}^{n} \omega_{ji} x_{ij}\big)\Big)^O \tag{4.21}$$

avec

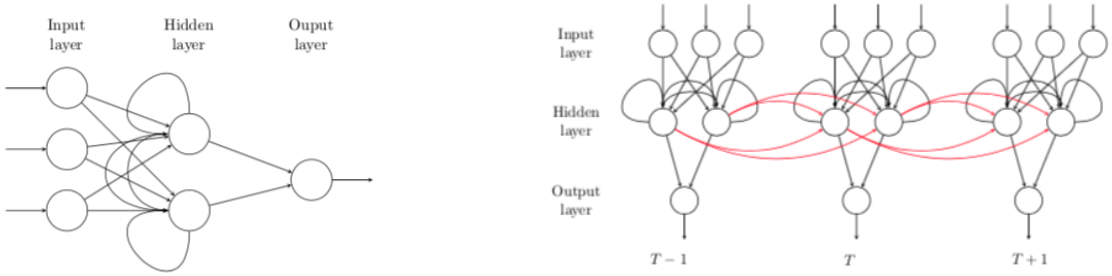
- ω_{jd} et ω_{ji} les poids respectifs de connexion entre couche d'entrée et les couches cachées puis entre les couches cachées et la couche de sortie.
- $f(.)$ [2] est une fonction d'activation pour l'information issue des couches cachées qui peut changer d'une couche à une autre et $f(.)^O$ une fonction d'activation pour l'information en provenance des couches cachées et à destination de la couche de sortie.

Les perceptrons multicouches utilisent l'algorithme de rétropropagation statique pour optimiser les poids des neurones en rétropropageant les erreurs à travers le réseau afin d'améliorer l'apprentissage sur les problèmes statiques tels que la reconnaissance d'images. Pour plus de détails sur le concept de rétropropagation(voir Azoff, 1994). Estimez l'erreur pour chaque neurone de sortie à l'aide d'une fonction d'erreur (également appelée fonction de coût) et additionnez-les pour obtenir l'erreur totale, l'erreur quadratique moyenne (mean squared error)) est l'une des fonctions de coût classiques en rétropropagation, qui est défini comme suit :

Considérons la paire $(\vec{x}_i, y_i)$ ou $\vec{x}_i$ est les vecteurs des variables d'entrée et y_i la variable de sortie puis $X = \{(\vec{x}_1, y_1), \ldots, (\vec{x}_{n*}, y_{n*})\}$ l'ensemble des paires entrée-sortie de taille n^*, erreur quadratique moyenne est :

$$E_{MSE} = \frac{1}{n^*} \sum_{i=1}^{n^*} (y_i - \hat{y}_i)^2 \tag{4.22}$$

avec $\hat{y}_i$ est la valeur de sortie calculée par le réseau à partir des valeurs d'entrée $\vec{x}_i$.
Ce processus de rétropropagation utilise principalement la méthode de descente de gradient (Gradient Descent) (Rumelhart et al., 1986) est limité dans la mesure où il est difficile d'avoir une configuration connue qui lui permette toujours de trouver une solution optimale, et nécessite donc d'utiliser des heuristiques pour s'assurer que la minimisation atteint bien un plus bas global.



(a) Architecture d'un réseau neuronal récurrent (b) Traitement d'une séquence temporelle dans un RNN

FIGURE 4.3 – Réseau de neurones récurrents

Face aux limites de la perception multicouche, les réseaux de neurones récurrents sont apparus comme une meilleure alternative. Mais contrairement au premier modèle ci-dessus, ils utilisent les informations actuelles et passées comme entrée. Cette utilisation des sources d'information actuelles et passées permet de comprendre les réponses à de nouvelles séquences de données.

2. les fonctions d'activation les plus couramment utilisées sont :
   - sigmoide $\sigma(x) = \frac{1}{1+e^{-x}}$
   - tangente hyperbolique $tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

Et cela à l'aide de boucles de rétroaction dans lesquelles le signal de sortie à chaque instant est également le signal d'entrée à l'instant suivant. Cette propriété justifie l'affirmation selon laquelle les réseaux de neurones récurrents ont de la mémoire, puisque les séquences d'entrée contiennent une grande quantité d'informations. Elles sont stockées dans des couches cachées et utilisées de manière récursive à travers le réseau à la vitesse à laquelle de nouvelles séquences d'informations sont traitées. Le concept de base d'un réseau de neurones récurrent (RNN) est de prédire des valeurs futures en se basant sur un entraînement sur des valeurs précédentes. En fait, la couche cachée doit mémoriser la séquence d'observations des étapes précédentes afin de prédire la séquence future de l'étape suivante, comme le montre la figure 4.3. La relation entre la couche de sortie (y_t), les couches cachées représentées par une une séquence vectorielle cachée $h = (h_1, \ldots, h_t)$ et la couche d'entrée $X_t = (x_{1,t}, \ldots, x_{n,t})$ des couches d'entrées précédentes $(X_1, \ldots, X_{t-1})$ peut-être décomposée en deux parties :

- de la couche d'entrée vers les couches cachées

$$h_t = f_h(W_{xh}X_t + W_{hh}h_{t-1} + b_h) \tag{4.23}$$

  avec

  * h_t la couche cachée à l'instant t
  * W_{xh} et W_{hh} respectivement les matrices des poids des neurones entre la couche d'entrée et les couches cachées puis entre les couches cachées elles-mêmes
  * X_t la couche d'entrée à l'instant t
  * h_{t-1} la couche cachée à l'instant $t-1$
  * f_h une fonction d'activation
  * b_h le biais

- des couches cachées vers la couche de sortie

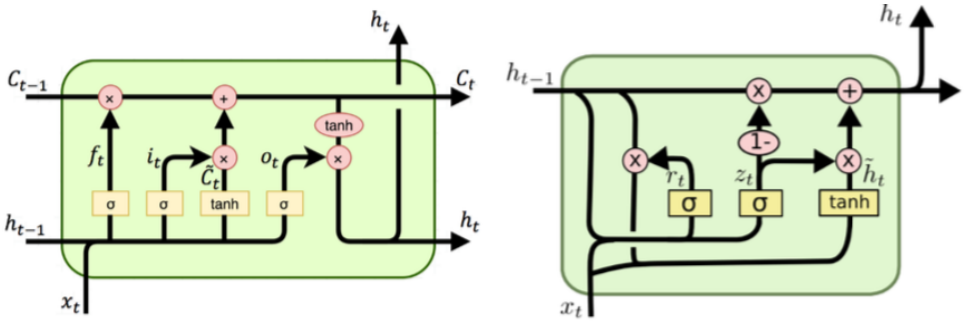
$$y_t = W_{yh}h_t + b_y \tag{4.24}$$

  où

  * y_t est la couche de sortie
  * W_{yh} est la matrice des poids des neurones entre les couches cachées et la couche de sortie
  * b_y est le biais.

Néanmoins, les RNN vanille (dans leur forme standard) ont deux limitations majeures, le problème du gradient de fuite (vanishing gradient) et le problème d'explosion du gradient (exploding gradient).
Le premier problème est que les signaux de gradient très faibles deviennent très petits et l'apprentissage est très, très lent en raison de la défaillance de la mémoire à long terme. Et le deuxième problème, il survient après que la valeur de la matrice de poids des neurones est trop élevée, ce qui entraîne un apprentissage instable, c'est-à-dire qu'il va diverger.

Une solution à ces problèmes est de remplacer les neurones par des cellules mémoire qui contiennent un ensemble d'opérations qui activent la mémoire dans la cellule, une idée de Hochreiter and Schmidhuber (1997). Cette unité dite à pas de temps unique (single time step cell) a une structure plus complexe, bien que sa composition soit similaire à celle d'un neurone avec une entrée, un état caché et une sortie. La complexité réside dans le traitement de l'information à l'intérieur de la cellule au niveau de l'état de la cellule, qui est régulé par des mécanismes appelés portes, au nombre de 3, la porte d'oubli (forget gate), la porte d'entrée (input gate) et la porte de sortie (output gate).



(a) Architecture d'un LSTM (b) Architecture d'un GRU

FIGURE 4.4 – Deux versions de RNN

Sur la figure 4.4a illustre l'ensemble du processus, où C_{t-1},C_t,h_{t-1}, h_t sont respectivement la cellule précédente, la cellule actuelle, la sortie de la cellule précédente et la sortie de la cellule actuelle, la bande qui ressemble à une ligne de transport entre C_{t-1} et C_t est appelée l'état de la cellule, elle transporte les informations de la cellule antérieure vers la présente et ainsi de suite. Nous avons également

- la porte oublier (forget gate) f_t qui constitue la première étape, en effet à ce niveau est prise la décision de mémoriser ou d'oublier certaines informations de cellule précédente

$$f_t = \sigma(W_{xf}X_t + W_{hf}h_{t-1} + b_f) \tag{4.25}$$

- la porte entrée (input gate) i_t vient ensuite comme la seconde étape qui est décomposée en deux parties, d'abord elle décide des valeurs à mettre à jour et par la suite créer de nouvelles valeurs candidates $\tilde{C}_t$

$$i_t = \sigma(W_{xi}X_t + W_{hi}h_{t-1} + b_i) \tag{4.26}$$

$$\tilde{C}_t = tanh(W_{xc}X_t + W_{hc}h_{t-1} + b_c) \tag{4.27}$$

Ensuite, la cellule C_{t-1} précédente est mise à jour dans la cellule C_t est faite comme :

$$C_t = C_{t-1} \otimes f_t \oplus i_t \otimes \tilde{C}_t \tag{4.28}$$

- la porte sortie (output gate) o_t est l'étape finale du processus, elle décide quelles parties de l'état de la cellule vont être produites en sortie et aussi permet à l'état de la cellule d'avoir ou non un effet sur d'autres neurones.

$$o_t = \sigma(W_{xo}X_t + W_{ho}h_{t-1} + b_o) \tag{4.29}$$

$$h_t = o_t \otimes C_t \tag{4.30}$$

Pour simplifier le calcul et surtout réduire le temps d'apprentissage pour des performances similaires, Cho et al. (2014b) ont proposé la Gated Recurrent Unit (GRU). C'est une variante de LSTM avec une unité de conditionnement de flux d'informations et aucune unité de stockage séparée, évitant ainsi la redondance des informations. Comparé au modèle LSTM, le modèle GRU a également la capacité de capturer les dépendances à long terme, en utilisant également le mécanisme de portes (gates).

Cependant, il existe des différences importantes à connaître, notamment au niveau de la gestion de la mémoire. En fait, le GRU expose entièrement son contenu mémoire sans aucun contrôle à chaque pas de temps, tout en trouvant un équilibre entre l'ancien et le plus récent contenu de la mémoire. De plus, il n'a que deux portes :

- la porte de mise à jour (update gate) qui régule la quantité d'informations précédentes à communiquer à l'étape suivante en donnant au réseau le choix de reconnaître toutes les informations passées, elle met à jour l'état caché (hidden state)
- la porte réinitialiser (reset gate) qui permet d'oublier certaines quantités d'informations.

L'architecture réseau du GRU est présentée à la figure 4.4b et est définie par les équations suivantes :

$$z_t = \sigma(W_{xz}X_t + W_{xh}h_{t-1} + b_z) \tag{4.31}$$

$$r_t = \sigma(W_{xr}X_t + W_{rh}h_{t-1} + b_r) \tag{4.32}$$

$$\tilde{h}_t = tanh(W_{xh}X_t + W_{rh}(h_{t-1} \odot h_{t-1})) \tag{4.33}$$

$$h_t = z_t\tilde{h}_t + (1 - z_t)h_{t-1} \tag{4.34}$$

où h_t est la sortie du GRU, z_t est la porte de mise à jour (update gate), r_t est la porte réinitialiser (reset gate) et $\tilde{h}_t$ est la sortie candidate. Cependant, à l'exception des réseaux de neurones récurrents, il existe d'autres architectures de réseaux de neurones telles que le Convolutional Neural Network (CNN). Traditionnellement connu pour son efficacité et sa performance sur des problématiques de classification d'images, de vision par ordinateur (segmentation, détection de contours, reconstruction d'images) surtout pour voitures autonomes et de traitement du langage naturel. Cette capacité à apprendre des caractéristiques discriminantes et informatives peut dans le cadre de séries temporelles financières aider à saisir les variations et les tendances pour de meilleures prédictions. Le Convolutional Neural Network de base est composé de 5 principaux éléments comme nous pouvons le voir à figure 4.5 ce sont :

- une couche de convolution (CONV) qui constitue les blocs de base du CNN.
  Ces blocs sont conçus pour détecter et extraire les caractéristiques pertinentes à différentes échelles temporelles dans les séries

- une fonction d'activation (ACTIV)
  Elle permet d'introduire de la non-linéarité dans le modèle et cela après chaque opération de convolution
- une couche de pooling (POOL)
  Elle permet de réduire respectivement la taille spatiale des données tout en préservant les caractéristiques pertinentes et la complexité du modèle. Ce qui donne le moyen de contrôler le surapprentissage (overfitting) et de rendre le modèle robuste.
- une couche entièrement connectée (Fully Connected) (FC)
  Similaire à un réseau neuronal classique, elle permet d'apprendre des combinaisons linéaires des caractéristiques extraites après les opérations des couches de convolution et de pooling et faire des prédictions
- une fonction d'activation (ACTIV)
  Elle est utilisée pour la couche de sortie dépend du type de problème, vu que dans notre cas nous avons un problème de régression alors elle sera une fonction linéaire

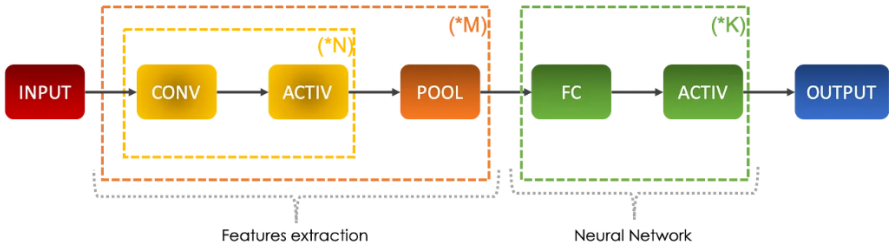


FIGURE 4.5 – Architecture du CNN

Comme toutes les architectures précédemment citées, celle-ci a son lot d'imperfections. Il convient de souligner sa nécessité d'un grand volume de données pour être efficace, une sensibilité aux hyperparamètres et un réglage des hyperparamètres qui peut s'avérer être une tâche laborieuse. Mais la plus importante est son manque de prise en compte des dépendances temporelles à long terme. De nombreuses architectures ont développé dans le but d'intégrer la causalité temporelle comme la convolution causale (causal convolution), la convolution dilatée (dilated convolution) et le réseau résiduel (residual network) (voir LeCun et al., 1998; Ding et al., 2015; Yu et al., 2017b; Navon and Keller, 2017; Oreshkin et al., 2019).

Bai, Kolter, and Koltun (2018), Yu, Yin, and Zhu (2017a) de par leurs travaux de recherche sont parvenus à obtenir une architecture bien plus complète en combinant les structures mentionnées précédemment pour former le Temporal Convolution Network dont on peut voir l'architecture à la figure 4.6. Il peut être considéré comme une combinaison d'une convolution causale(Waibel et al., 1989), une convolution dilatée(Yu and Koltun, 2015) et réseau résiduel(He et al., 2016). Plus concrètement on pourrait le formuler de la manière suivante.

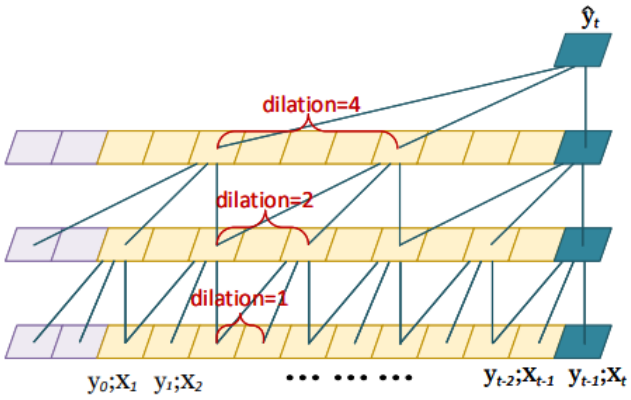


FIGURE 4.6 – Architecture du TCN

Soit $(x_0, \ldots, x_N)$ une séquence de données d'entrée du modèle de prévision d'une série temporelle financière. Sur la base des données historiques observées, l'objectif de la prédiction du modèle est de produire des valeurs de prédiction $(\hat{y}_0, \ldots, \hat{y}_N)$ pour une période future. Notre modèle de prédiction consiste à obtenir la fonction $\mathcal{F}$ et à lui faire produire la cartographie suivante :

$$\hat{y}_0, \ldots, \hat{y}_N = \mathcal{F}(x_0, \ldots, x_N) \tag{4.35}$$

Le processus d'apprentissage d'un modèle consiste à optimiser en continu la fonction de perte pour minimiser les écarts entre les valeurs réelles et prédites. Vu que cette architecture composée des plusieurs structures, elle bénéficie de l'ensemble des propriétés. En effet, la convolution causale lui permet de respecter l'ordre chronologique des événements et capture les relations causales dans les séries temporelles. Ce qui évite l'apparition de "fuites"(leakage) d'informations.

Elle adopte une forme d'un réseau convolution unidimensionnel entièrement connecté, utile pour capturer des caractéristiques temporelles complexes sur des séquences de longueur variable. Le réseau de convolution 1D a aussi l'avantage d'utiliser des filtres partagés sur toute une séquence qui par conséquent réduit le nombre de paramètres et le temps d'apprentissage du modèle. On peut le définir comme :

Soit une séquence unidimensionnelle $\mathbf{x} \in \mathbb{R}^n$, avec un noyau de convolution $f\ \{0, \ldots, k-1\} \rightarrow \mathbb{R}$, l'opération de convolution F sur les éléments x_t de la séquence est définie comme :

$$(F * \mathbf{x})(x_t) = \sum_{k=1}^{K} f_k(x_t(K-k)) \tag{4.36}$$

où K est le nombre de noyaux de convolution.
Ensuite, la convolution dilatée, quant à elle permet d'élargir le champ réceptif des neurones sans augmenter le nombre du modèle. Ce qui améliore la stabilité de l'apprentissage et réduit le surajustement. Par analogie avec la définition précédente, on peut la définir comme

$$(F_d * \mathbf{x})(x_t) = \sum_{k=1}^{K} f_k(x_t(K-k)d) \tag{4.37}$$

où d est le facteur de dilatation.
Elle peut être assimilée à une généralisation de convolution de base.
Pour terminer, le réseau résiduel qu'on peut généralement être représentée par $\mathcal{F}$ et qui effectue une série de transformations sur l'input $\mathbf{x}$ à laquelle on ajoute $\mathbf{x}$ pour les opérations de sortie.

$$\text{output} = \sigma(\mathbf{x} + \mathcal{F}(\mathbf{x})) \tag{4.38}$$

où σ est une fonction d'activation.
Il est utilisé pour optimiser le flux d'informations et faciliter l'optimisation du réseau.

4.2.3 Les modèles hybrides

Les séries temporelles peuvent présenter des caractéristiques de tendance linéaire donc il serait incorrect de vouloir les modéliser avec uniquement des modèles non linéaires. De plus, des auteurs comme Srivastava et al. (2015), Yosinski et al. (2015) ont mené des études empiriques sur les limites des réseaux de neurones sur des tâches linéaires. Ils sont arrivés à mettre en évidence leurs difficultés à capturer les relations linéaires. C'est la raison pour laquelle, l'approche hybride de Zhang (2003) qui consiste à combiner le modèle ARIMA (AutoRegressive Integrated Moving Average) avec les réseaux de neurones nous semble la plus adaptée aux séries temporelles financières. Selon Zhang (2003), un modèle hybride est la somme d'une composante linéaire et d'une composante non linéaire qui peut être défini comme

$$y_t = L_t + N_t \tag{4.39}$$

où L_t désigne la composante linéaire et N_t la composante non linéaire.
La modélisation commence par la composante linéaire avec le modèle ARIMA, puis les résidus du modèle linéaire sont donnés au modèle non linéaire pour apprendre son comportement non linéaire, appelé e_t, nous avons :

$$e_t = y_t - \hat{L}_t \tag{4.40}$$

$$e_t = f(e_{t-1}, e_{t-2}, \ldots) \tag{4.41}$$

où f est une fonction non linéaire appartenant aux modèles avancés, $\hat{N}_t = f(e_{t-1}, e_{t-2}, \ldots)$
Dans la littérature, les potentiels candidats pour l'approximation de la fonction f, nous avons MLP, RNN, LSTM (GRU), SVM, GP, TCN avec de nombreuses applications comme dans les articles de Pai and Lin (2005); Moghadam and Keshmirpour (2011); Asadi et al. (2012); Kumar and Thenmozhi (2014); Zhang et al. (2016); Choi (2018); Fathi (2019); Ordóñez et al. (2019); Rajalakshmi et al. (2022); Temur et al. (2019); Xu et al. (2022).

Dans le cadre de notre analyse, nous mettrons en évidence la nécessité de la distribution de Student dans le processus de perturbation aléatoire du modèle ARIMA. Cela revient à supposer que le terme d'erreur (ϵ) dans l'équation 4.1 suit une distribution de Student. Nous aurons alors les modèles ARIMA N-ANN et ARIMA T-ANN pour les distributions Normale et Student, respectivement. Cela nous permettra de comparer leurs performances respectives. En ce qui concerne la modélisation de la composante non linéaire, nous considérons les modèles d'apprentissage automatique GP, MLP, LSTM (et GRU), SVM et TCN.

4.2.4 Les métriques d'évaluation

Afin de comparer quantitativement les performances des modèles prédictifs, il est essentiel d'utiliser des métriques appropriées qui mesurent l'adéquation du modèle aux données et sa capacité à généraliser. Notre choix s'est porté sur trois mesures de performance :

- Root Mean Square Error (RMSE)
  Elle quantifie la précision globale d'un modèle et indique la propagation de l'erreur. Des valeurs RMSE inférieures indiquent de meilleures performances du modèle.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}}$$

- Mean Absolute Percentage Error (MAPE)
  Elle mesure de l'erreur absolue moyenne en pourcentage entre les valeurs prédites et les valeurs réelles. Ce qui la rend facile à interpréter et à comprendre pour les praticiens car elle exprime le pourcentage d'erreur relative entre la valeur prédite et la valeur réelle. En finance, les erreurs relatives ont tendance à avoir une signification plus pertinente que les erreurs absolues.

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100$$

- Nash-Sutcliffe Efficiency (NSE)
  Introduit initialement par Nash and Sutcliffe (1970) pour évaluer la puissance des modèles en hydrologie, ce coefficient est une forme générale de l'indice du coefficient de détermination (R^2) avec des valeurs allant de $-\infty$ à 1. Lorsque le modèle s'adapte parfaitement la valeur du coefficient vaut 1 ($NSE = 1$) mais lorsque sa valeur est proche de 0 ($NSE \approx 0$) cela signifie que les prédictions du modèle ne sont pas meilleures que les données moyennes prédites par le modèle. Elle pénalise les erreurs de prédiction plus importantes de manière disproportionnée, ce qui peut être pertinent dans le domaine financier où les erreurs de prédiction à la baisse peuvent avoir des conséquences significatives.

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

où y_i, $\hat{y}_i$ et $\bar{y}$ sont respectivement la valeur réelle, prédite à l'instant i et la moyenne des valeurs réelles.

4.3 Résultats empiriques

Dans le cadre de notre analyse empirique, nous avons choisi d'évaluer les performances de différents modèles de prévision sur 4 actifs appartenant à différentes classes d'actifs. Le premier actif est le SP500, l'un des indices boursiers les plus connus. Comme tous les indices boursiers, ses variations reflètent l'évolution des actifs sous-jacents. Ensuite en second, une action [3] de l'entreprise TotalEnergies (TTE.PA). Ses variations peuvent être liés à ses facteurs spécifiques, tels que la mauvaise gestion, les scandales, les dividendes potentiellement élevés de la croissance et, d'autre part, des facteurs inhérents aux marchés financiers.

3. une action est un titre de propriété d'une fraction du capital d'une entreprise

Les deux derniers sont les actifs EURUSD et JNUG qui sont respectivement une paire de devise(currencies) et un ETF (Exchange-Traded Fund). L'EURUSD est l'une des paires de devises les plus négociées sur le marché des changes (forex).

De nombreux facteurs économiques, politiques et financiers tels que les taux d'intérêt, les indicateurs économiques, la politique monétaire, les événements géopolitiques sont à l'origines de ces variations. "Direxion Daily Junior Gold Miners Index Bull 3X Shares" (JNUG) est un ETF dont l'objectif est de réaliser des performances trois fois supérieures à celles de l'actif sous-jacent, qui est lié aux actions des petites sociétés minières d'or. Dans le cas de cet actif, en plus des variations dues aux actifs sous-jacents, une autre source serait liée à son objectif de levier et de la volatilité. À la lumière des facteurs susceptibles d'affectés les cours de ces actifs financiers, nous avons jugé préférable de faire des prédictions à un pas (One-step ahead forecasting). Elles réalisent des prédictions précises et détaillées à court terme et sont plus flexibles aux changements récents de tendance et aux évènements inattendus.

Le tableau 4.1 répertorie certaines informations sur les actifs que nous avons sélectionné. Les données utilisées dans cet article sont les cours de clôture quotidiens pour trois périodes, couvrant la période du 2016-01-04 au 2021-04-30, du 2017-01-03 au 2022-04-29 et du 2018-01-02 au 2023-04-28 extraites du Datastream de Yahoo `https://fr.finance.yahoo.com/`. Dans la suite, pour parler des différentes périodes, nous utiliserons les termes period 1, period 2, et period 3 respectivement.

nº	Asset	Ticker	Asset class
1	SP500	ĜSPC	Index
3	TOTALENERGIES	TTE.PA	Stock
2	EURO/USD	EUR/USD	Currency
4	Direxion Daily Junior Gold Miners Index Bull	JNUG	ETF

TABLE 4.1 – Les actifs financiers sélectionnés

periods	assets	min	mean	max	std	Jarque-Bera	ADF
period 1	SP500	1829.1	2549.8	3240	333.47	49.693	1.7658
	TTE.PA	22.662	33.452	41.838	4.2065	41.768	0.73931
	EURUSD	1.039	1.1346	1.251	0.044382	32.356	0.13472
	JNUG	305.9	1213.4	6202.1	1167.6	2251.3	-1.5011
period 2	SP500	2237.4	2834.5	3756.1	338.67	50.84	1.1876
	TTE.PA	17.282	33.598	41.838	4.4917	65.967	-0.43465
	EURUSD	1.042	1.1434	1.251	0.045354	20.476	1.0819
	JNUG	32.831	616.91	2417.7	404.25	681.87	-1.8111
period 3	SP500	2237.4	3291.8	4793.1	638.5	136.62	1.6493
	TTE.PA	17.282	34.316	41.838	4.62	131.73	0.16857
	EURUSD	1.0657	1.1567	1.251	0.042236	42.934	-0.52357
	JNUG	32.831	366.31	992.38	276.34	94.692	-1.9364

TABLE 4.2 – Statistiques descriptives des prix

periods	assets	min	mean (%)	max	std (%)	skewness	kurtosis	Jarque-Bera	ADF
period 1	SP500	-0.038259	0.043346	0.033759	0.71423	-0.82547	7.9593	1129.3	-32.462
	EURUSD	-0.020638	0.0065923	0.01709	0.38979	0.13245	5.267	215.32	-32.877
	TTE.PA	-0.072858	0.042104	0.072995	1.1604	-0.25059	7.7857	957.04	-31.54
	JNUG	-0.37271	-0.01437	0.31466	6.712	-0.33072	6.8298	624.35	-33.659
period 2	SP500	-0.099945	0.048916	0.089683	1.1017	-0.2481	22.985	16518	-38.429
	EURUSD	-0.028144	0.0098999	0.014467	0.38805	-0.4134	7.0829	717.29	-31.65
	TTE.PA	-0.16025	-0.0057421	0.13976	1.5774	-0.81047	30.154	30586	-29.106
	JNUG	-1.0843	-0.32552	0.46082	8.0692	-5.6052	74.13	2.1432e+05	-27.404
period 3	SP500	-0.099945	0.057186	0.089683	1.1433	-0.25249	19.888	11858	-37.898
	EURUSD	-0.028144	-0.0065445	0.014467	0.36421	-0.58391	7.9138	1059.7	-31.149
	TTE.PA	-0.16025	0.018933	0.13976	1.6596	-0.7047	24.827	19875	-29.435
	JNUG	-1.0843	-0.35709	0.46082	7.7771	-6.2347	84.826	2.846e+05	-26.422

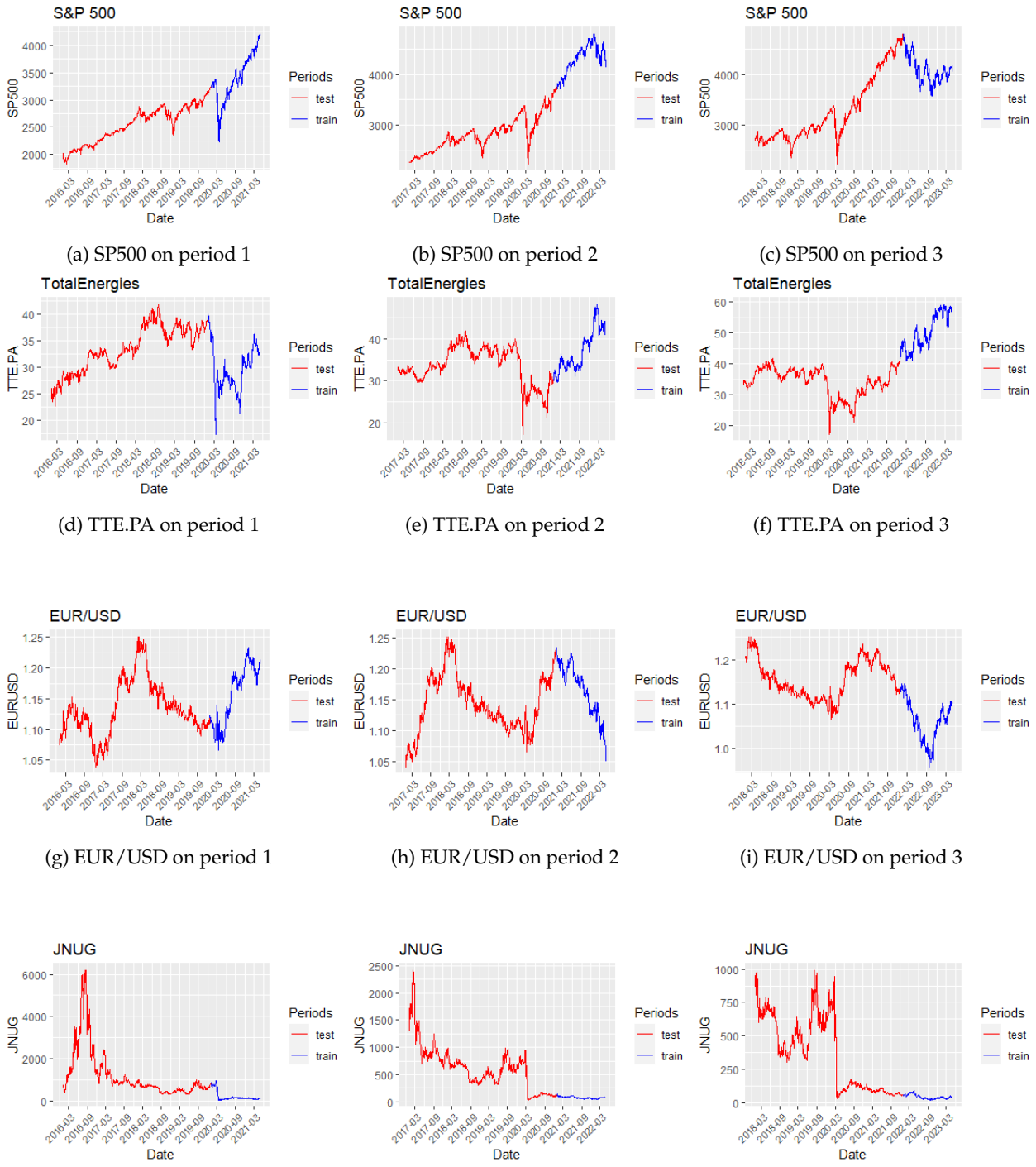
TABLE 4.3 – Statistiques descriptives des rendements

Les statistiques descriptives usuelles ainsi que les statistiques de normalité [4] et du test de stationnarité [5] sur les 3 périodes pour nos séries temporelles sont résumées dans le tableau 4.3. Lorsque la statistique de Jarque-Bera est calculée, elle est comparée à une valeur critique basée sur une distribution de référence. Cette valeur critique vaut 5.99 au-delà de laquelle l'hypothèse nulle de normalité est rejetée. Comme nous pouvons le voir aucune des séries n'est distribuée selon une loi normale. L'hypothèse nulle du test de Dickey-Fuller augmenté est que la série temporelle contient une racine unitaire, ce qui signifie qu'elle n'est pas stationnaire. Lorsque la statistique de test est inférieure à la valeur critique -1.9416, on peut rejeter l'hypothèse nulle.

Alors nous sommes en mesure de dire qu'aucune des séries n'est stationnaire. La stationnarité est une propriété statistique importante pour de modèles économétriques, car elle permet de supposer des propriétés statistiques, telles que la moyenne et la variance, restent constantes dans le temps. Ainsi, pour ces modèles, nous sommes aménés à convertir les séries brutes de prix des actifs considérés en séries des rendements par différence logarithmique. L'observation de la figure 4.7 montre que certains actifs semblent présenter une tendance haussière comme le SP500 et une tendance baissière comme l'EURUSD. La différenciation permet aussi d'éliminer la tendance et des effets saisonniers. La figure 4.8 montre les séries temporelles des rendements logarithmiques qui oscillent autour la valeur moyenne de 0 et est stationnaire.

La stationnarité des séries est vérifiable grâce aux statistiques de test de Dickey-Fuller (ADF) augmenté dans la table 4.3. Les différentes séries des rendements présentent des caractéristiques propres aux séries temporelles financières une asymétrie négative et un excès de kurtosis du fait des valeurs de skewness négatives et de kurtosis supérieures à 3. Elles ont donc des queues de distributions plus épaisses que des distributions gaussiennes. En plus de cela, le test de normalité de Jarque-Bera nous permet de rejeter l'hypothèse de normalité des rendements. Nous pouvons également constater des valeurs d'écart-type supérieures aux valeurs des moyennes ce qui traduit une forte volatilité et laisse supposer la présence d'hétéroscédasticité. Pour évaluer les performances des différents modèles, chaque ensemble de données est divisé en deux parties : les 996 premières observations (environ 75%) pour l'apprentissage et les 332 observations restantes (environ 25%) pour les tests.

4. Jarque-Bera
5. Augmented Dickey-Fuller (ADF)



(a) SP500 on period 1 (b) SP500 on period 2 (c) SP500 on period 3

(d) TTE.PA on period 1 (e) TTE.PA on period 2 (f) TTE.PA on period 3

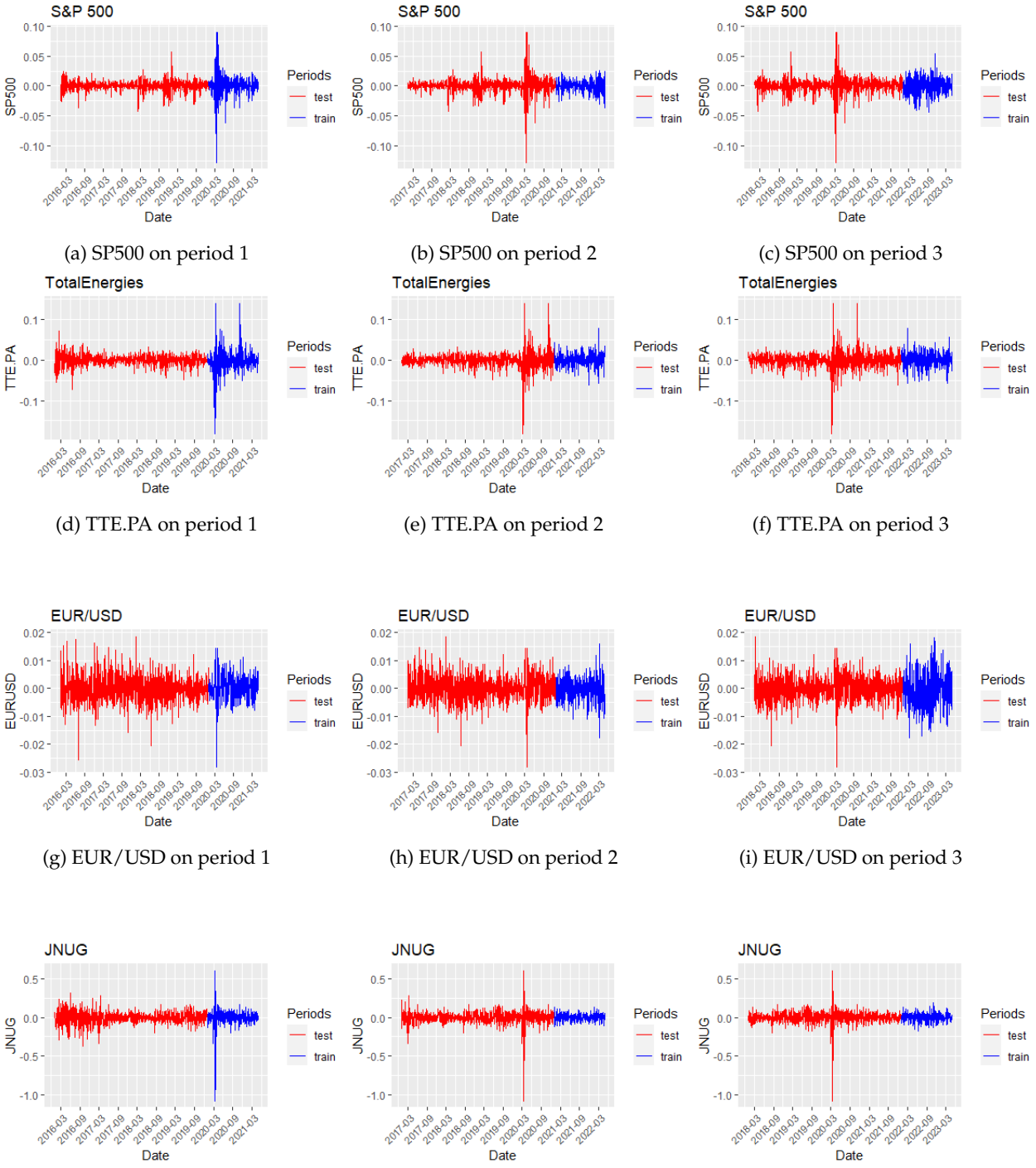
(g) EUR/USD on period 1 (h) EUR/USD on period 2 (i) EUR/USD on period 3

(j) JNUG on period 1 (k) JNUG on period 2 (l) JNUG on period 3

FIGURE 4.7 – Les séries brutes

L'implémentation des modèles de machine learning et hybrides requiert un choix de paramètres particuliers que nous estimerons en subdivisant en plusieurs sous-échantillons l'échantillon d'apprentissage selon la méthode de validation croisée (cross validation [6]). Notre premier objectif est de montrer que dans le cadre de la prévision des séries temporelles financières à l'aide de processus ARIMA, l'estimation des paramètres doit se faire sous l'hypothèse que les erreurs sont distribuées selon la loi de Student.

6. pour plus de détails (voir Kohavi et al., 1995)



(a) SP500 on period 1 (b) SP500 on period 2 (c) SP500 on period 3

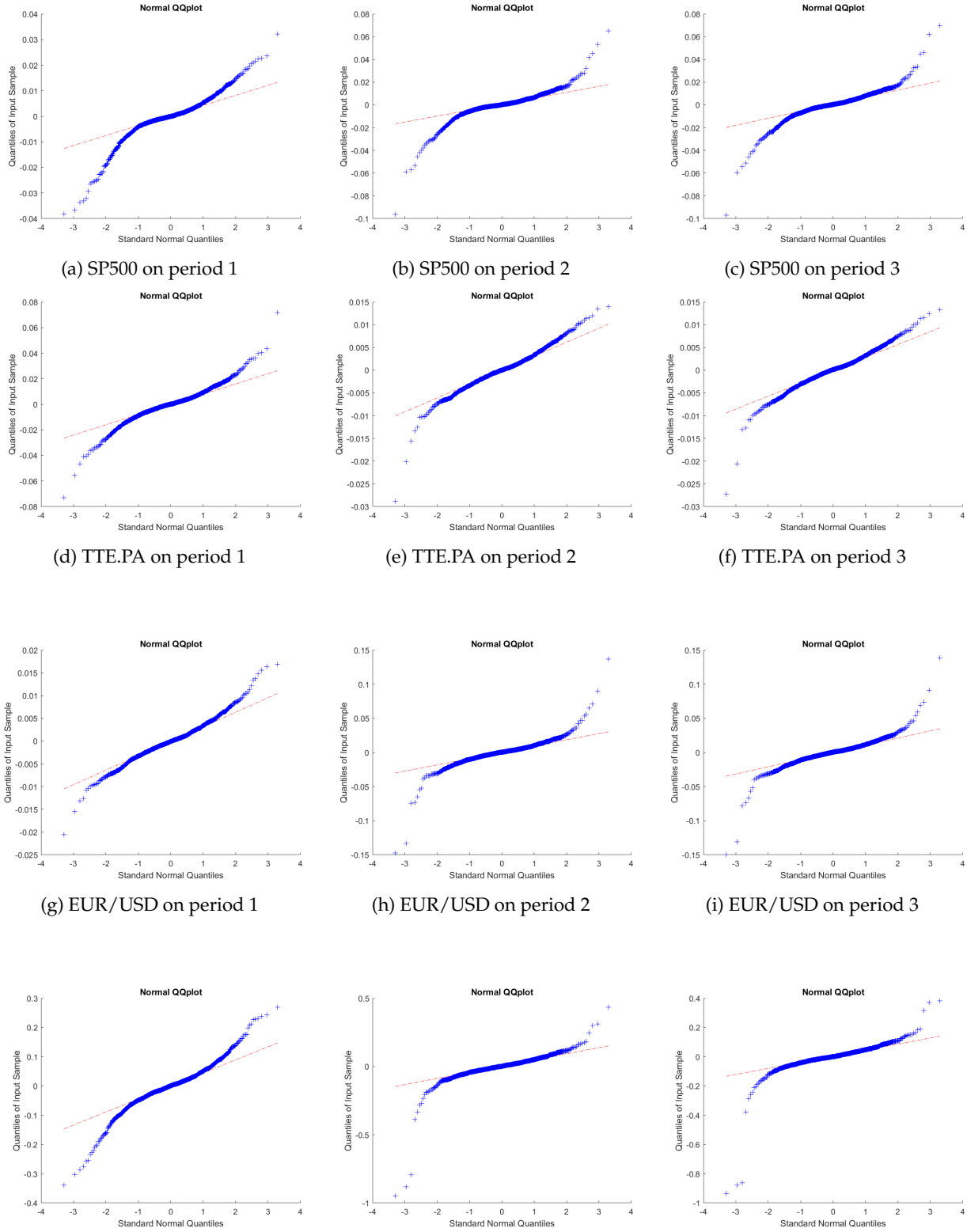
(d) TTE.PA on period 1 (e) TTE.PA on period 2 (f) TTE.PA on period 3

(g) EUR/USD on period 1 (h) EUR/USD on period 2 (i) EUR/USD on period 3

(j) JNUG on period 1 (k) JNUG on period 2 (l) JNUG on period 3

FIGURE 4.8 – Les séries des rendements

Pour trouver des preuves empiriques à cet égard, nous avons recherché les meilleurs paramètres du modèle ARIMA pour chaque série et réalisé des graphiques quantiles-quantiles (QQ-plots) pour évaluer la similitude entre la distribution empirique des erreurs et les distributions normale et Student. C'est sans grande surprise que sur la figure 4.9, nous voyons que la distribution des erreurs a des queues plus épaisses que la distribution normale. Cependant, quand on regarde la figure 4.10, si elles ne montrent pas toutes une similitude exacte avec la distribution de Student, nous ne sommes pas loin. Nous n'allons pas nous contenter uniquement de ces résultats.



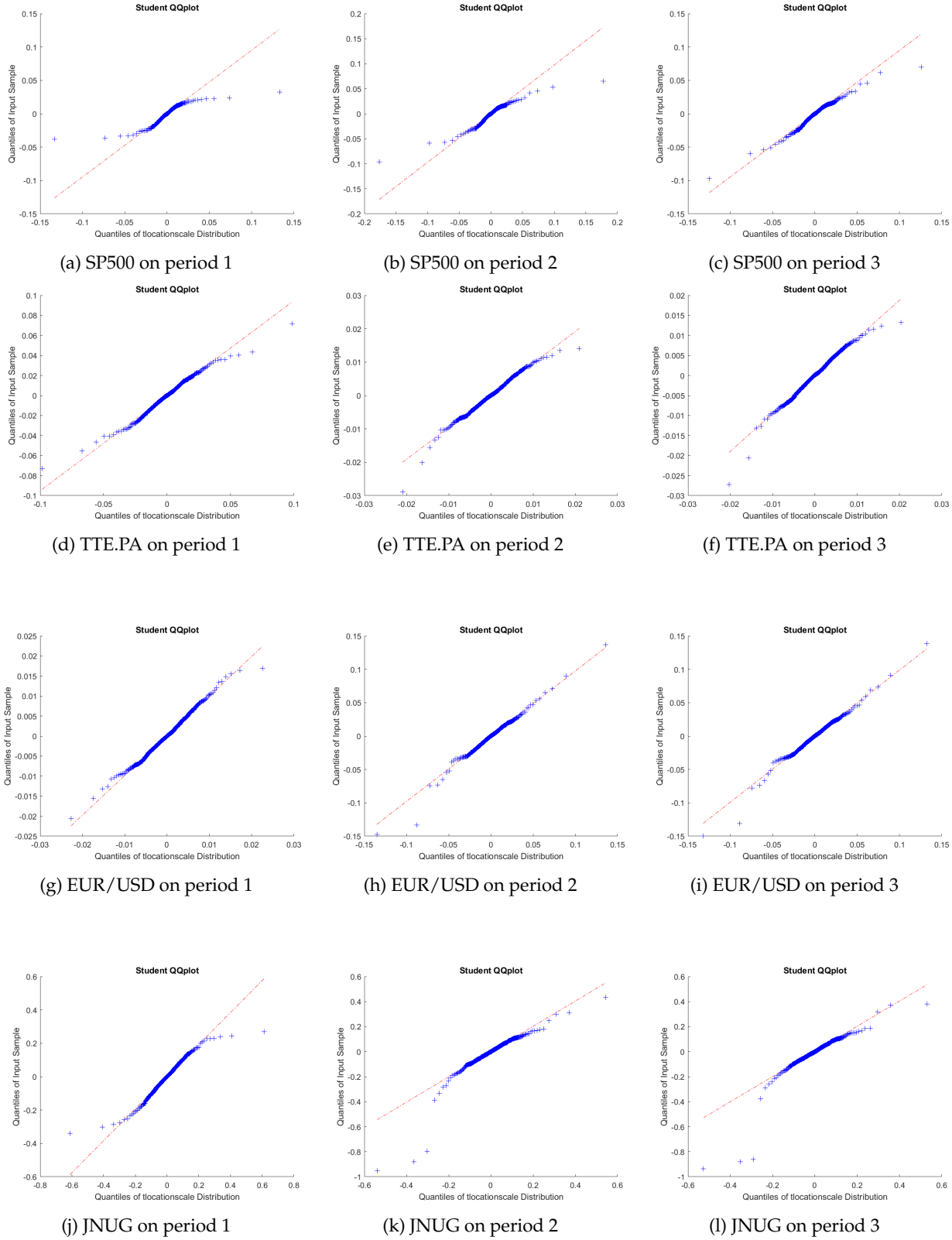
(a) SP500 on period 1 (b) SP500 on period 2 (c) SP500 on period 3

(d) TTE.PA on period 1 (e) TTE.PA on period 2 (f) TTE.PA on period 3

(g) EUR/USD on period 1 (h) EUR/USD on period 2 (i) EUR/USD on period 3

(j) JNUG on period 1 (k) JNUG on period 2 (l) JNUG on period 3

FIGURE 4.9 – Les diagrammes quantile-quantile avec une loi normale



(a) SP500 on period 1 (b) SP500 on period 2 (c) SP500 on period 3

(d) TTE.PA on period 1 (e) TTE.PA on period 2 (f) TTE.PA on period 3

(g) EUR/USD on period 1 (h) EUR/USD on period 2 (i) EUR/USD on period 3

(j) JNUG on period 1 (k) JNUG on period 2 (l) JNUG on period 3

FIGURE 4.10 – Les diagrammes quantile-quantile avec une loi de Student

C'est pourquoi dans la suite, nous allons estimer les paramètres des modèles ARIMA (resp. ARIMA-ANN) avec l'hypothèse que la perturbation aléatoire suit une loi normale ou une loi de Student (ARIMA N et ARIMA T) (resp ARIMA N-ANN et ARIMA T-ANN). Ainsi sur la base des données utilisées dans ce papier, nous allons comparer leurs performances.

Ensuite, le second objectif est de montrer la pertinence de l'utilisation des modèles hybrides qui permettre de profiter des avantages des modèles ARIMA et de machine learning. Pour ce faire, la première étape consiste à évaluer la présence de non-linéarité dans les résidus (erreurs) du modèle ARIMA. À la table 4.4, nous avons les statistiques de test et les p-values du test de non-linéarité de Terasvirta(Teräsvirta et al., 1993) dont l'hypothèse nulle (H_0) est que la série est linéaire et l'hypothèse alternative est que la série présente une non-linéarité significative. Lorsque la p-value est inférieure au seuil α (5%), on peut rejeter l'hypothèse nulle (H_0) et conclure qu'il y a une non-linéarité significative dans la série.

Sur la période 1, nous pouvons accepter l'hypothèse H_0 tandis que sur les périodes 2 et 3 excepté l'actif EURUSD, les p-values faibles une preuve solide de non-linéarité significative dans les résidus. Nous sommes passés outre cette condition de non-linéarité en appliquant tout de même les modèles hybrides à l'actif EURUSD. Sinon nous n'aurons pas eu suffisamment d'observations pour montrer si les échantillons des RMSE des modèles hybrides et de machine learning proviennent de la même population ou pas. L'ensemble des résultats de prédiction sont répertoriés aux tables C.1, C.2 et C.3. Mais pour une meilleure synthèse des résultats, nous avons trié par ordre décroissant selon leurs performances des modèles et enregistré les 10 meilleurs dans les tables 4.5, 4.6 et 4.7.

	Period 1				Period 2				Period 3			
	SP500	TTE.PA	EURUSD	JNUG	SP500	TTE.PA	EURUSD	JNUG	SP500	TTE.PA	EURUSD	JNUG
statistic	1.40	0.18	1.48	3.35	55.31	23.95	0.28	11.11	53.49	22.83	1.05	14.58
P-value	0.50	0.91	0.48	0.19	0.00	0.00	0.87	0.00	0.00	0.00	0.59	0.00

TABLE 4.4 – test de non linéarité de Terasvirta

Models	RMSE	MAPE	NSE
TCN	23.7150	0.6038	0.9968
GRU	30.2211	0.7768	0.9948
LSTM	30.5332	0.8123	0.9947
MLP	31.9272	0.8209	0.9942
GP	32.2475	0.8295	0.9940
SVR	35.2081	0.9215	0.9929
ARIMA(3,1,0) N gjr-GARCH(1,1)	55.1519	1.1775	0.9826
ARIMA(2,1,1) T gjr-GARCH(1,1)	55.2326	1.1806	0.9825
ARIMA(3,1,0) N GARCH(1,1)	55.2656	1.1780	0.9825
ARIMA(2,1,1) T GARCH(1,1)	55.5243	1.1835	0.9824

(a) SP500

Models	RMSE	MAPE	NSE
LSTM	0.5154	1.5277	0.9858
TCN	0.5230	1.5843	0.9854
GRU	0.5245	1.5527	0.9853
SVR	0.5745	1.7186	0.9824
GP	0.5745	1.7186	0.9824
MLP	0.7639	1.8683	0.9689
ARIMA(4,1,5) T GARCH(1,1)	0.7642	1.8769	0.9689
ARIMA(2,1,0) N GARCH(1,1)	0.7643	1.8701	0.9689
ARIMA(2,1,0) N gjr-GARCH(1,1)	0.7646	1.8695	0.9688
ARIMA(4,1,5) T gjr-GARCH(1,1)	0.7655	1.8770	0.9688

(b) TTE.PA

Models	RMSE	MAPE	NSE
TCN	0.0039	0.2907	0.9931
LSTM	0.0044	0.3238	0.9912
GRU	0.0046	0.3481	0.9904
ARIMA(2,1,3) T TCN	0.0046	0.3543	0.9902
ARIMA(3,1,0) N TCN	0.0047	0.3552	0.9900
SVR	0.0050	0.3733	0.9887
GP	0.0052	0.3485	0.9878
MLP	0.0052	0.3505	0.9877
ARIMA(3,1,0) N gjr-GARCH(1,1)	0.0052	0.3524	0.9876
ARIMA(3,1,0) N GARCH(1,1)	0.0052	0.3529	0.9876

(c) EUR/USD

Models	RMSE	MAPE	NSE
GP	28.4639	7.5424	0.9834
ARIMA(4,1,5) N GARCH(1,1)	28.4803	7.2518	0.9834
ARIMA(4,1,4) T gjr-GARCH(1,1)	28.6244	7.1204	0.9832
ARIMA(4,1,4) T SVR	28.6661	7.0874	0.9832
ARIMA(4,1,5) N gjr-GARCH(1,1)	28.7547	7.1091	0.9831
ARIMA(4,1,4) T GP	28.8505	7.1566	0.9830
ARIMA(4,1,4) T MLP	28.9143	7.0991	0.9829
ARIMA(4,1,4) T TCN	28.9826	21.8245	0.9828
ARIMA(4,1,5) N SVR	29.0494	7.4029	0.9827
ARIMA(4,1,4) T	29.0794	7.1304	0.9827

(d) JNUG

TABLE 4.5 – Les 10 meilleurs modèles de prédiction sur la période 1

Models	RMSE	MAPE	NSE
ARIMA(4,1,4) T TCN	27.9655	0.3201	0.9894
ARIMA(4,1,4) T LSTM	30.1613	0.3814	0.9876
ARIMA(4,1,4) T GRU	32.7457	0.4279	0.9854
ARIMA(3,1,5) N LSTM	33.5177	0.4474	0.9847
MLP	34.1043	0.5054	0.9842
ARIMA(3,1,5) N GP	34.4903	0.4685	0.9838
ARIMA(4,1,4) T GP	34.7230	0.4556	0.9836
GP	35.5304	0.5484	0.9828
ARIMA(4,1,4) T gjr-GARCH(1,1)	35.5479	0.5344	0.9828
ARIMA(4,1,4) T SVR	38.8534	0.5922	0.9794

(a) SP500

Models	RMSE	MAPE	NSE
TCN	0.3110	0.7396	0.9958
ARIMA(5,1,5) T LSTM	0.4227	0.8610	0.9923
LSTM	0.5064	1.1958	0.9890
GRU	0.5731	0.9841	0.9859
ARIMA(5,1,5) T TCN	0.6314	1.3166	0.9829
ARIMA(5,1,5) T GARCH(1,1)	0.6322	1.2360	0.9828
ARIMA(5,1,5) N gjr-GARCH(1,1)	0.6325	1.2427	0.9828
ARIMA(5,1,5) T gjr-GARCH(1,1)	0.6325	1.2427	0.9828
MLP	0.6334	1.2411	0.9828
ARIMA(5,1,5) N GARCH(1,1)	0.6348	1.2315	0.9827

(b) TTE.PA

Models	RMSE	MAPE	NSE
TCN	0.0019	0.1359	0.9978
LSTM	0.0023	0.1711	0.9967
GRU	0.0024	0.1800	0.9965
MLP	0.0040	0.2941	0.9904
ARIMA(2,1,3) T TCN	0.0047	0.3091	0.9867
ARIMA(2,1,3) T GARCH(1,1)	0.0047	0.3102	0.9866
ARIMA(3,1,0) N GP	0.0047	0.3101	0.9866
ARIMA(2,1,3) T SVR	0.0047	0.3128	0.9866
ARIMA(3,1,0) N gjr-GARCH(1,1)	0.0047	0.3120	0.9865
ARIMA(3,1,0) N	0.0047	0.3105	0.9865

(c) EUR/USD

Models	RMSE	MAPE	NSE
ARIMA(1,1,5) T GARCH(1,1)	0.2050	0.2306	0.9999
ARIMA(4,1,5) N TCN	0.4555	0.4967	0.9993
ARIMA(1,1,5) T TCN	0.4581	0.5044	0.9993
ARIMA(1,1,5) T	0.4746	0.5233	0.9993
TCN	1.4747	1.6646	0.9929
ARIMA(1,1,5) T LSTM	3.6472	3.5077	0.9566
ARIMA(1,1,5) T SVR	3.6662	3.5405	0.9562
ARIMA(4,1,5) N GARCH(1,1)	3.6776	3.4910	0.9559
ARIMA(1,1,5) T gjr-GARCH(1,1)	3.6785	3.5267	0.9559
ARIMA(1,1,5) T GRU	3.6868	3.5684	0.9557

(d) JNUG

TABLE 4.6 – Les 10 meilleurs modèles de prédiction sur la période 2

Models	RMSE	MAPE	NSE
ARIMA(3,1,2) T TCN	23.7824	0.5117	0.9916
ARIMA(4,1,2) N TCN	25.7229	0.5486	0.9902
ARIMA(3,1,2) T GRU	27.0831	0.5844	0.9891
ARIMA(3,1,2) T gjr-GARCH(1,1)	27.4110	0.5823	0.9888
ARIMA(3,1,2) T SVR	27.5931	0.5925	0.9887
TCN	43.6003	0.9180	0.9717
GP	57.4471	1.1097	0.9509
MLP	57.4717	1.1081	0.9509
ARIMA(3,1,2) T LSTM	57.9272	1.1154	0.9501
ARIMA(3,1,2) T	57.9391	1.1165	0.9501

(a) SP500

Models	RMSE	MAPE	NSE
ARIMA(1,1,2) T LSTM	0.5097	0.8824	0.9913
ARIMA(1,1,2) T TCN	0.5311	0.9291	0.9905
ARIMA(1,1,2) T GRU	0.5410	0.9473	0.9902
ARIMA(1,1,2) T GP	0.5913	1.0268	0.9882
ARIMA(1,1,2) T SVR	0.6135	1.0661	0.9873
TCN	0.6614	1.1613	0.9853
LSTM	0.8328	1.4567	0.9767
ARIMA(1,1,2) T GARCH(1,1)	0.9423	1.4941	0.9701
ARIMA(1,1,2) T gjr-GARCH(1,1)	0.9425	1.4950	0.9701
ARIMA(1,1,2) T	0.9434	1.4944	0.9701

(b) TTE.PA

Models	RMSE	MAPE	NSE
TCN	0.0024	0.1974	0.9972
LSTM	0.0025	0.2045	0.9969
GRU	0.0025	0.2071	0.9969
SVR	0.0029	0.2391	0.9958
MLP	0.0030	0.2495	0.9955
GP	0.0030	0.2517	0.9955
ARIMA(4,1,5) T TCN	0.0038	0.3139	0.9930
ARIMA(4,1,5) T GRU	0.0040	0.3301	0.9924
ARIMA(5,1,2) N GRU	0.0064	0.4833	0.9801
ARIMA(5,1,2) N GP	0.0064	0.4827	0.9801

(c) EUR/USD

Models	RMSE	MAPE	NSE
ARIMA(5,1,1) T TCN	1.1082	2.4595	0.9954
ARIMA(5,1,1) T LSTM	1.3347	2.9219	0.9933
ARIMA(4,1,1) N GRU	1.4028	3.0685	0.9926
ARIMA(5,1,1) T SVR	1.4105	3.0641	0.9926
ARIMA(5,1,1) T MLP	1.4234	3.1097	0.9924
ARIMA(4,1,1) N GP	1.5522	3.4321	0.9910
ARIMA(5,1,1) T GRU	1.5722	3.4058	0.9908
ARIMA(5,1,1) T GP	1.6084	3.4330	0.9903
ARIMA(4,1,1) N LSTM	2.4312	4.2681	0.9779
ARIMA(4,1,1) N gjr-GARCH(1,1)	2.4549	4.2808	0.9775

(d) JNUG

TABLE 4.7 – Les 10 meilleurs modèles de prédiction sur la période 3

Les figures 4.11 et 4.12 montrent les prédictions des meilleurs modèles par rapport aux séries originales. La période 1 est dominée par les modèles de machine learning même si au niveau de l'actif JNUG, hormis le fait que le GPR soit le plus performant, les 9 autres sont des modèles hybrides. Les résultats sont plus hétérogènes sur la période 2. Par exemple, les modèles hybrides sont hyper-représentés parmi les meilleurs modèles du SP500 et du JNUG. Indépendamment du fait que le TCN ait surperformé sur l'action TTE.PA, les modèles hybrides figurent parmi ceux qui ce sont les mieux illustrés. Les modèles de machine learning ont totalement surclassé les autres sur la série de la paire de devise EURUSD sur les périodes 2 et 3. Sur la période 3, ce sont les modèles hybrides qui ont réalisé les meilleures performances sur les actifs SP500, TTE.PA et du JNUG.

4.4 Discussions

Premièrement, pour obtenir des preuves statistiques solides pour ou contre l'hypothèse selon laquelle les modèles hybrides réalisent de meilleures performances lorsque les processus de perturbation aléatoire suivent la distribution de Student contrairement à la distribution normale. Pour cela, nous avons choisi le test unilatéral de Wilcoxon (Mann-Whitney).
$\mathbf{H}_0$: Il n'y a pas de différence entre les médianes des deux échantillons de RMSE (ARIMA T ANN et ARIMA N ANN)
$\mathbf{H}_1$: La médiane de l'échantillon des RMSE de l'ARIMA T ANN est inférieure à la médiane de l'échantillon des RMSE de l'ARIMA N ANN. Les résultats du test sont contenus dans la table 4.8.

Sur la période 1, nous notons qu'il n'y a pas de différence significative entre les modèles hybrides sous l'hypothèse que la perturbation aléatoire suit une loi gaussienne ou de Student. Lorsqu'on observe les figures 4.10a, 4.10d, 4.10g, 4.10j, nous constatons que les queues de distributions sont bien plus épaisses que celle de la distribution de Student. Les distributions Normale et de Student sont toutes les deux inadaptées, les évènements extrêmes observés sur cette période suggèrent que d'autres distributions comme la distribution de Student asymétrique ou (Generalized Student's Distribution) ou distribution de Burr. Alors que sur la période 2, les modèles hybrides sous l'hypothèse que la perturbation aléatoire suit une loi de Student ont des performances statistiquement supérieures par rapport au cadre gaussien pour le SP500 et le JNUG. Pour l'action TTE.PA nous ne sommes pas loin du même résultat. Une explication serait qu'au regard des figures 4.10b,4.10e,4.10h et 4.10k, nous voyons qu'il n'y a plus de similarité entre la distribution de Student et des résidus du processus ARIMA.

	Period 1				Period 2				Period 3			
	SP500	TTE.PA	EURUSD	JNUG	SP500	TTE.PA	EURUSD	JNUG	SP500	TTE.PA	EURUSD	JNUG
statistic	13	13	15	11	6	8	9	7	9	1	17	6
P-value	0.2424	0.2424	0.3496	0.1548	0.0325	0.0660	0.0898	0.0465	0.0898	0.0022	0.4686	0.0325

TABLE 4.8 – Test de Wilcoxon unilatéral des RMSE entre les modèles ARIMA N et ARIMA T

	Period 1				Period 2				Period 3			
	SP500	TTE.PA	EURUSD	JNUG	SP500	TTE.PA	EURUSD	JNUG	SP500	TTE.PA	EURUSD	JNUG
statistic	72	72	66	13	24	53	55	11	24	22	72	0.00
P-value	1	0.9996	0.9989	0.0159	0.1406	0.9487	0.9666	0.0091	0.1454	0.1064	1.0000	0.0001

TABLE 4.9 – Test de Wilcoxon unilatéral des RMSE entre les modèles hybrides et machine learning

L'amplitude des mouvements de baisse observés sur les prix de ces actifs financiers commencent à s'atténuer ce qui entraine une baisse des événements extrêmes. Nous voyons sur les figures 4.8b, 4.8e et 4.8k que les rendements sont caractérisés par de basse volatilité contrairement à la devise EURUSD (fig. 4.8h).
Nous parvenons à des résultats similaires sur la période 3, sauf que cette fois, nous n'avons pas de différence significative entre les performances de ces 2 approches de modèles hybrides sur les séries des actifs SP500 et EURUSD. Subséquemment à cette série de tests, une nouvelle a été faire similairement à la précédente. Mais cette fois nous avons cherché à trouver des preuves statistiques soutenant ou non la supériorité des performances des modèles hybrides par rapport aux modèles de machine learning.

Le test unilatéral de test de Wilcoxon est défini comme suit.
$\mathbf{H}_0$: Il n'y a pas de différence entre les médianes des deux échantillons de RMSE (modèles hybrides et ML)
$\mathbf{H}_1$: La médiane de l'échantillon des RMSE modèles hybrides est inférieure à la médiane de l'échantillon des modèles de machine learning.

Les résultats du test sont contenus dans la table 4.9.

À l'exception des séries de l'actif JNUG sur lesquelles les modèles hybrides ont des performances significativement supérieures aux modèles de machine learning, nous n'avons pas obtenu de différence significative entre ces 2 types de modèles sur les autres séries. Les échantillons d'apprentissage des 4 actifs sur la période 1 ne sont marquées d'aucunes variations inhabituelles. Alors que les séries de test, elles sont toutes marquées par mouvements abrupts liés à pandémie de COVID-19. En regardant les figures 4.11 et 4.12, on peut observer autour du 20 février 2020, les premières baisses sur les marchés financiers. Puis vint le lundi 9 mars 2020, une journée noire pour les marchés financiers, où les baisses se sont intensifiées. Et comme après chaque baisse, un rebond qui dans le cas échéant était lié aux mesures de relance économique annoncées par les gouvernements et les banques centrales. Ces mouvements abrupts peuvent susciter des réactions rapides et parfois excessives de la part des opérateurs de marchés et engendrer d'autres mouvements. Ils font partie des arguments en faveur de l'impossibilité de prévision à long terme des séries temporelles financières.

C'est la raison pour laquelle, il n'y a que les architectures particulières des réseaux de neurones artificiels qui ont pu s'adapter à ces changements soudains de dépendances temporelles et de tendances. La modélisation adéquate des dépendances temporelles dans les données nécessitait des niveaux de non-linéarité polynomiale ou fonctionnelle qui sont plus élevés que le niveau de non-linéarité quadratique des modèles GARCH. Même si l'ajout de terme non quadratique comme dans le modèle GJR-GARCH pour permet d'améliorer la capacité des modèles GARCH pour capturer l'asymétrie, le niveau de non-linéarité de ces modèles reste quadratique. Alors que sur la période 2, les échantillons d'apprentissage contiennent la fenêtre temporelle avec les variations brutales des cours des marchés financiers liées à la crise sanitaire. La plupart des échantillons de test présentent des mouvements relativement calmes en dehors de la série de la paire de devises EURUSD (figs. 4.7h et 4.8h). La tendance à la baisse amorcée par l'EURUSD engendre une hausse de la volatilité. Et comme à chaque fois que nous sommes en présence de forte volatilité, les modèles de machine learning notamment le TCN et le LSTM qui ont surclassé les autres modèles.

Une remarque similaire peut être faite en observant les légers pics de volatilités sur la série des rendements de TTE.PA (fig. 4.8e). Sur cet actif, le modèle TCN affiche des résultats remarquables suivit du modèle hybride ARIMA(5,1,5) T LSTM. Concernant les actifs SP500 et JNUG, ce sont respectivement les modèles ARIMA (4,1,4) T TCN, et ARIMA (1,1,5) T GARCH (1,1) qui se sont démarqués L'augmentation de niveaux de non-linéarité peut également augmenter la complexité du modèle qui peut s'avérer inappropriée pour la modélisation des dépendances temporelles. Les performances du modèle ARIMA (1,1,5) T GARCH (1,1) sur l'etf JNUG montrent qu'un niveau de non-linéarité peut largement et donc les modèles traditionnels ne sont pas à oublier aussi facilement.

En effet, l'etf JNUG est conçu de façon à fournir des rendements amplifiés par rapport à ses principaux sous-jacents qui sont des actions de sociétés généralement impliquées dans l'exploration, le développement et l'extraction d'or de petite capitalisation. Il vise à fournir un effet de levier 3x et est donc plus sensible aux mouvements de ses sous-jacents sachant que cet effet de levier fonctionne à la fois à la hausse et à la baisse. Voilà pourquoi parmi nos actifs, il est celui qui a perdu plus 95% de valeur durant les fortes baisses. Lorsque que nous observons l'évolution de sa chronique, nous voyons une tendance linéaire horizontale qui est très peu affectée par de fortes variations. C'est ce qui explique les meilleures performances du modèle ARIMA (1,1,5) T GARCH (1,1) suivi des modèles hybrides ARIMA (4,1,5) N TCN, ARIMA(1,1,5) T TCN et même du ARIMA(1,1,5) T.

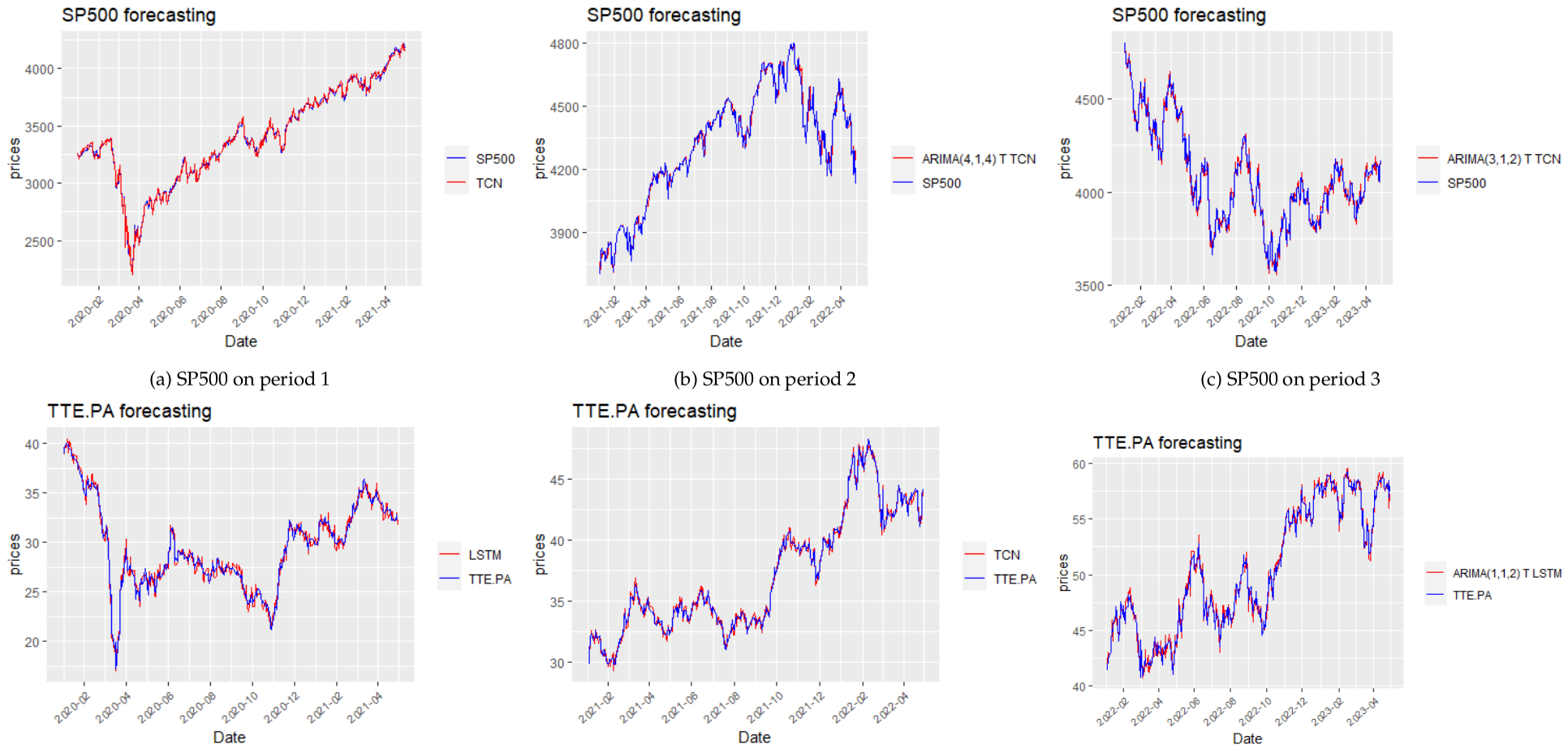
Sur la période 3, la série temporelle EURUSD est la seule sur laquelle les modèles de machine learning ont réalisé les meilleures performances. Sur les autres actifs, les modèles hybrides ont obtenu les meilleurs résultats par rapport aux 3 indices de performances. Cette période couvre le début de la crise entre l'Ukraine et la Russie. Le seul actif le plus influencé par cette crise est l'EURUSD comme nous pouvons le voir sur les figures 4.7i et 4.8i. En effet depuis le déclenchement de la crise russo-ukrainienne, les pays de la zone Euro font face à un choc économique et énergétique majeur qui se traduit par un choc inflationniste. Elle a engendré des externalités négatives sur la croissance économique de ces pays en raison de leur proximité géographique avec la zone de conflit mais surtout la forte dépendance de leurs économies aux hydrocarbures russes. Ce climat d'incertitude provoque une défiance envers l'euro au profit du dollar. La figure 4.12c montre clairement 2 tendances, une tendance baissière suivie d'une tendance haussière. Il n'y a que les modèles de machine learning qui ont pu capturer ce changement structurel dans la chronique de cette série temporelle. Pour les autres actifs, malgré que leurs chroniques soient marquées par des motifs qui se répètent on distingue clairement soit une tendance à la hausse ou à la baisse mais pas les deux à la fois. Ces caractéristiques ont été parfaitement modélisé avec par les modèles ARIMA T-TCN et ARIMA T-LSTM.

4.5 Conclusion

Dans cet article, nous avons proposé de faire l'hypothèse que le processus de perturbation aléatoire suit une loi de Student dans les modèles hybrides afin de certains faits stylisés souvent observés lors de l'analyse des séries temporelles des rendements d'actifs financiers. Les résultats empiriques nous fournissent des preuves statistiques qui nous permettent d'affirmer que cette approche est plus performante que celle dans un cadre gaussien habituel. Ils apportent aussi des preuves statistiques solides soutenant qu'il n'a pas de différence significative entre les modèles hybrides et les modèles de machine learning. Après avoir après avoir étudié un grand nombre d'architectures de réseaux de neurones, Greff et al. (2016) avaient affirmé qu'aucune architecture ne pouvait être significativement supérieure à l'architecture LSTM, au regard de nos résultats, nous sommes en mesure de prétendre que le TCN est un très bon concurrent. Et cela aussi bien en tant que modèle de machine learning que pour les modèles hybrides pour les tâches de modélisation de séquences comme dans notre cas.

Bien que nous ne sommes parvenus à montrer une différence significative entre ces 2 classes de modèles, les modèles hybrides ont très souvent été les plus performants par rapport aux modèles de machine learning. En restant dans la logique de chercheurs comme Granger (1989) et Krogh and Vedelsby (1994) qui avaient démontré théoriquement et empiriquement que la combinaison de modèles complètement différents réduirait les erreurs de prédiction, une solution serait d'opter pour d'autres modèles hybrides en combinant plusieurs modèles de deep learning comme CNN-LSTM comme le suggèrent la littérature récente.

D'autre part, étant conscient que la dynamique des cours boursiers est affectée par de nombreux facteurs externes, il serait judicieux d'intégrer l'analyse des sentiments des données des médias sociaux via le Natural Language Processing (NLP) dans des modèles prédictifs en supposant une corrélation entre les mouvements des cours des actions et les médias sociaux. Comme ça été avec les cours des actions GameStop (GME), AMC Entertainment Holdings (AMC), Hertz Global Holdings (HTZ) et même de la crypto-monnaie Bitcoin (BTC). Cette alternative pourrait aussi être considérée comme une forme d'hybridation au résultat tout aussi prometteur.



(a) SP500 on period 1

(b) SP500 on period 2

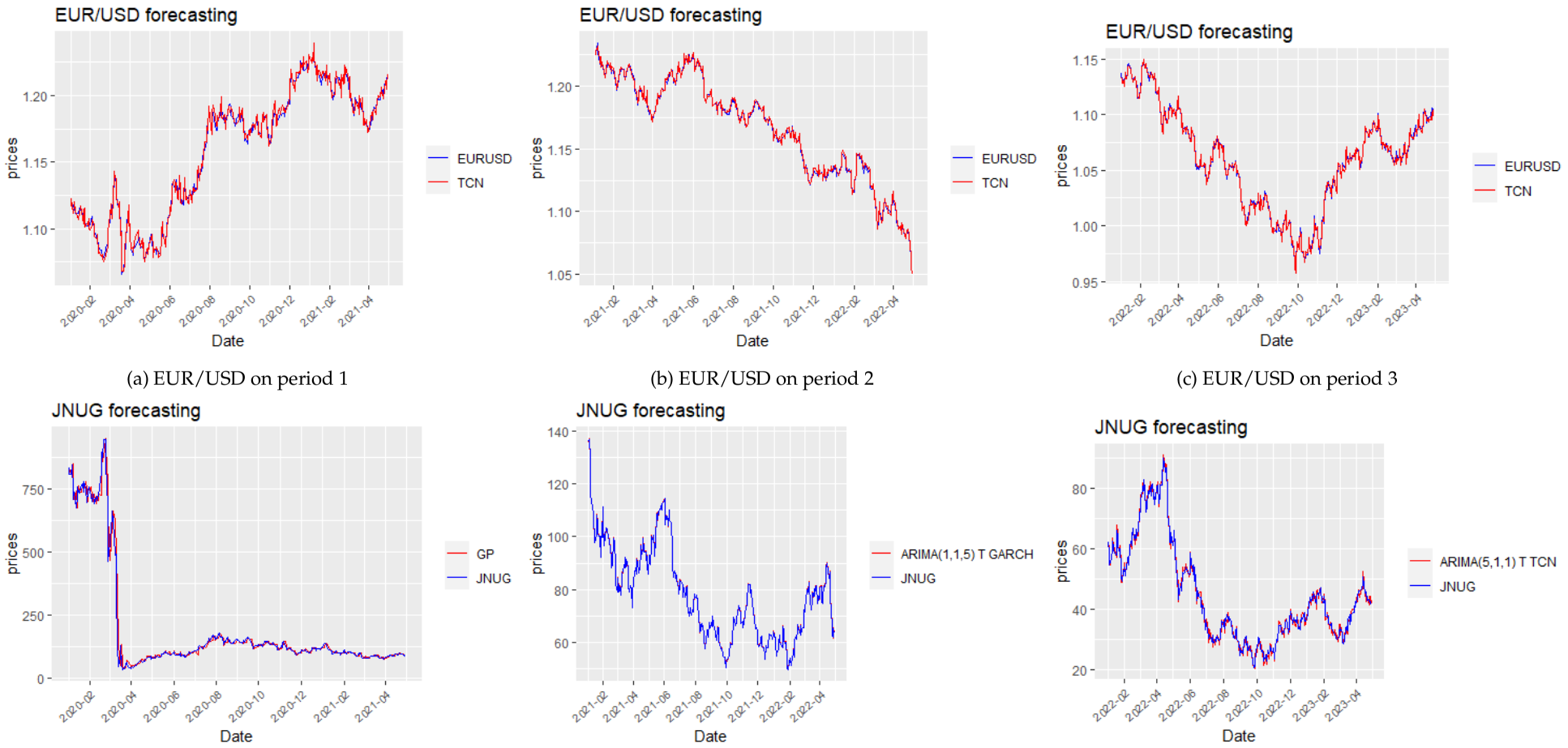
(c) SP500 on period 3

(d) TTE.PA on period 1

(e) TTE.PA on period 2

(f) TTE.PA on period 3

FIGURE 4.11 – Les meilleurs modèles pour la prévision des séries SP500 et TTE.PA



(a) EUR/USD on period 1

(b) EUR/USD on period 2

(c) EUR/USD on period 3

(d) JNUG on period 1

(e) JNUG on period 2

(f) JNUG on period 3

FIGURE 4.12 – Les meilleurs modèles pour la prévision des séries EURUSD et JNUG

Conclusion

Discussion et conclusion

L'objectif de cette thèse est d'analyser l'intégration des critères environnementaux, sociaux et de gouvernance (ESG) dans l'allocation des portefeuilles d'actifs financiers. En fait, au cours des dernières années, nous avons assisté à une croissance phénoménale de l'investissement socialement responsable, qui peut s'expliquer par des investisseurs institutionnels et même des investisseurs individuels cherchant à aligner leurs objectifs financiers sur des objectifs de développement durable. La véritable motivation de ces investisseurs peut être de bénéficier de la pérennité et de la stabilité possibles des rendements financiers à long terme de tels investissements d'une part, et de se protéger des risques extra-financiers potentiels d'autre part (scandales, catastrophes écologiques, questions fiscales). Contrairement à l'investissement traditionnel, l'investissement ESG peut être animé par une noble intention qui permet des changements profonds dans nos habitudes qui peuvent réduire considérablement notre impact environnemental en luttant contre la pollution des sols et le changement climatique. Parmi les différentes catégories d'Investissement Socialement Responsable (ISR) disponibles, nous avons choisi de nous limiter à l'investissement ESG.

Ainsi, pour analyser l'impact de la prise en compte des critères ESG sur la rentabilité et le risque des portefeuilles d'actifs financiers, nous menons deux tests empiriques. Dans le premier article correspondant au chapitre 2, nous avons comparé les performances de trois portefeuilles, chacun composé de 39 actions du marché américain. Chacun des trois portefeuilles est composé à l'aide de filtres ESG. Ainsi, le premier portefeuille est construit à partir des actions des entreprises les mieux notées ESG (AAA et AA), dites best in class. Quant au second portefeuille, il est composé d'actions aux notes moyennes (A et BBB), tandis que le troisième portefeuille est composé d'actions aux plus mauvaises notes ESG (BB, B et CCC). Nos résultats empiriques montrent que les meilleurs portefeuilles d'actions de leur catégorie ne surperforment pas les autres portefeuilles. Cela rejoint les conclusions de Utz et al. (2014), Halbritter and Dorfleitner (2015) qui ont montré qu'il n'y a pas de différence significative entre les rendements des entreprises les mieux notées ESG et ceux des autres entreprises.

En revanche, lorsque la structure de dépendance est considérée à l'aide de modèles plus complexes tels que le modèle regime-switching vine copula, nos résultats montrent que le portefeuille composé des actions les moins bien notées ESG surperforme les deux autres portefeuilles. Ainsi, le portefeuille idéal semble être composé d'actions avec des notes ESG moyennes. Quels que soient le modèle d'optimisation et la structure de dépendance utilisés, ces portefeuilles ont obtenu les meilleurs résultats en termes de rentabilité et de risque tout au long de la période d'étude. En effet, le portefeuille reste le moins risqué, malgré la période d'étude marquée par la crise sanitaire du Covid 19 et le début du conflit russo-ukrainien à semer la pagaille sur les marchés financiers.

Concernant la structure de dépendance, le cadre robuste de modélisation mean-LPM (Lower Partial Moment) avec copules nous donne une grande satisfaction par rapport au modèle traditionnel de Markowitz (1952a). C'est ce qui a motivé notre décision de choisir un cadre de modélisation robuste pour le deuxième test correspondant au chapitre 3. Dans le chapitre 3, nous analysons la performance de deux portefeuilles constitués d'actions d'entreprises du marché de la zone euro dans le cadre du cadre de modélisation défini par Zhu and Fukushima (2009) et étendu par Kakouris and Rustem (2014). Cette fois, nous avons d'un côté un portefeuille de 30 actions best-in-class (AAA et AA) et de 30 actions de différentes notations (A, BBB, BB et B) de l'autre. Contrairement au chapitre 2, la période de recherche a débuté quelques jours avant les fortes turbulences des marchés financiers liées à la crise du Covid 19. Nos résultats montrent que le portefeuille best-in-class sous-performe les autres portefeuilles. Il semble insatisfaisant que les portefeuilles dits ESG soient loin d'être surperformants voire comparables aux portefeuilles traditionnels.

Ainsi, le concept de rentabilité financière semble difficile à relier au concept d'investissement durable. Une autre façon de voir les choses est que le concept de durabilité remet en question les mesures de performance traditionnelles, même si l'investissement ESG est un investissement à long terme et que les tendances sont réversibles. Notre dernier essai correspondant au chapitre 4 s'inscrit dans une volonté d'étendre nos recherches à des modèles de gestion de portefeuille plus complexes mais avec des contraintes plus réalistes. Ce chapitre nous permet de souligner la puissance d'outils tels que les réseaux de neurones artificiels (Neural networks) et la valeur ajoutée qu'ils peuvent apporter aux méthodes économétriques traditionnelles. L'un des plus grands avantages qu'une personne puisse avoir en finance est le privilège d'avoir des informations futures.

C'est pourquoi la communauté scientifique a déployé d'énormes efforts pour développer des méthodes permettant de faire des prédictions avec la plus grande précision possible. Le modèle de réseau de neurones se distingue naturellement par son excellente qualité prédictive. Cependant, nous avons également exploré la possibilité d'utiliser des modèles hybrides qui peuvent tirer parti à la fois des modèles avancés et traditionnels. Bien que nous nous limitions au cas univarié, des résultats avec des performances similaires peuvent être obtenus dans un cadre multivarié, permettant une transition parfaite vers des modèles de gestion de portefeuille plus réalistes.

Contribution

Cette thèse contribue à la littérature empirique qui s'est intéressée à l'analyse de l'impact de l'intégration des considérations ESG sur la performance des portefeuilles. Dans la littérature, il existe deux grandes approches empiriques :

- Une approche qui a fait le choix de l'extension du modèle de Markowitz comme Hirschberger et al. (2013) et Gasser et al. (2017)
- Une autre approche qui a fait le choix de l'approche multi-objectifs comme Utz et al. (2014)

Nous différencions les deux approches en construisant des portefeuilles à l'aide de filtres ESG. L'optimisation de ces portefeuilles est effectuée par des méthodes d'optimisation robustes telles que la mean-LPM et le worst case mixture copula mean-CVaR. Nos résultats empiriques constituent une contribution empirique au débat scientifique sur l'impact de la prise en compte des critères ESG dans la construction de portefeuille.

D'un point de vue théorique, nous avons apporté une légère modification au modèle Regime switching vine copula de Stöber and Czado (2014) et repris par Fink et al. (2017). En effet, considérer que la structure de dépendance pourrait être modélisée par un changement de régime entre deux structure de dépendance de copule de vines serait inadapté. Cela peut se justifier par le fait que les actions ne restent pas la même durée dans un régime. Cette contribution théorique s'est basée sur la flexibilité qu'offre le théorème de Sklar qui nous a permis de modéliser les marginales avec les modèles de Markow Regime Switching univarié et la structure de dépendance avec les copules de vines.

Limites

Notre thèse présente certaines limites :

- Au niveau des notations ESG
  Les notations utilisées sont celles de Morgan Stanley Capital International ESG rating[7]. Le problème est qu'il n'y a pas de convergence entre les notations des différentes agences Standard & Poor's, Moody's, Fitch Ratings et Bloomberg avec celles de MSCI. L'autre problème concernant les notations a été soulevé par Wimmer (2013) en faisant le constat que les notations ESG ne se conservaient pas sur le long terme ce qui a pour conséquence un rééquilibrage fréquent du portefeuille.
- Au niveau des coûts de transactions
  Nous rééquilibrons notre portefeuille très fréquemment, de sorte que nos résultats peuvent changer à mesure que les coûts de transaction augmentent.

Travaux futurs

Le modèle de Markowitz et la plupart de ses extensions décrivent un processus statique, que nous appelons un problème d'optimisation à cycle unique. En d'autres termes, la distribution reste constante jusqu'à la fin de la période d'investissement, de sorte que la distribution ne peut pas être ajustée pour s'adapter aux changements supposés des marchés financiers.

7. https://www.msci.com/our-solutions/esg-investing/esg-ratings-climate-search-tool/issuer/

Pour résoudre ce problème, l'amélioration de modèles plus réalistes avec allocation dynamique est envisagée, mais la mise en œuvre implique de nombreux problèmes complexes. C'est la raison pour laquelle nous avons adopté une stratégie "rolling buy-and-hold" où nous résolvons le problème d'optimisation en réoptimisant à plusieurs reprises. Cette stratégie consiste à réaliser une séquence de stratégies de buy-and-hold afin d'incorporer les nouvelles informations du marché.

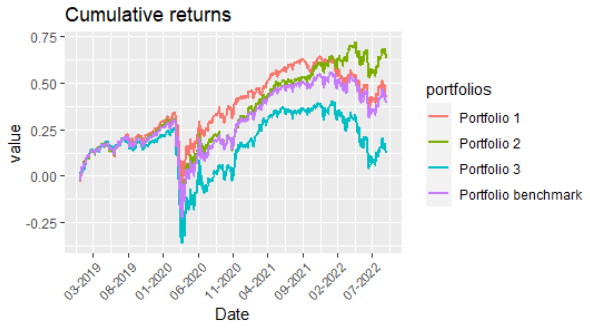
Une solution plus pérenne est la programmation dynamique, mais il existe encore de nombreux freins, dont la malédiction de la dimensionnalité, qui la rendent difficile à mettre en place pour les grands portefeuilles. Ce problème a été résolu avec l'utilisation des réseaux de neurones artificiels qui ont conduit à l'apprentissage par renforcement avec des algorithmes comme le Proximal Policy Optimization (PPO) et le Deep Deterministic Policy Gradient (DDPG). Ces algorithmes, qui ont déjà fait leur preuve dans d'autres domaines, commencent d'être utilisés en finance. Ils seraient de bonnes alternatives aux modèles de gestion de portefeuille déjà existantes.

ANNEXE A

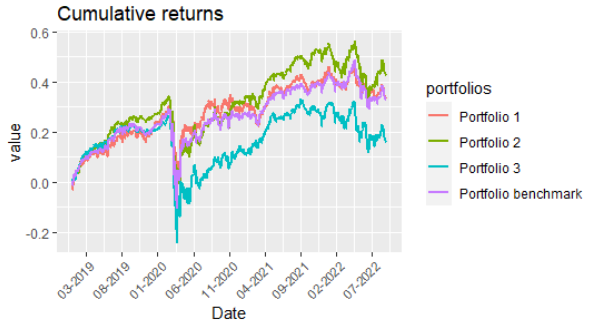
Annexe du chapitre 2

A.1 Model performance

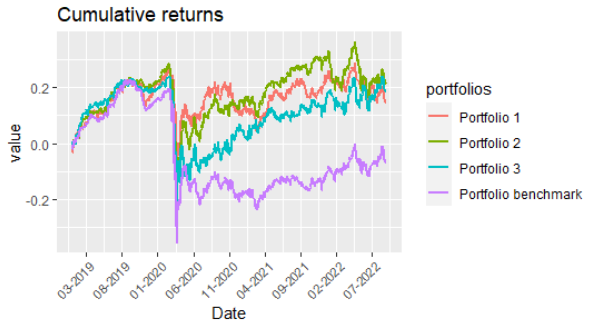
It is important to note that when you see the words “model (1)” and “model (2)”, this refers respectively to whether short selling is allowed or not.



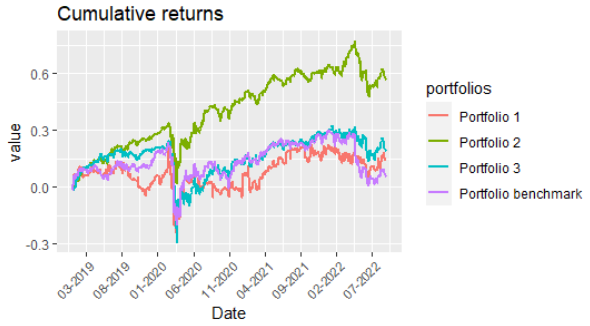
(a) $1/n$ model



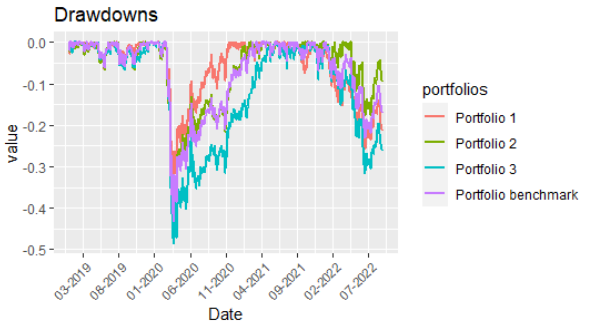
(b) mean-var model (1)



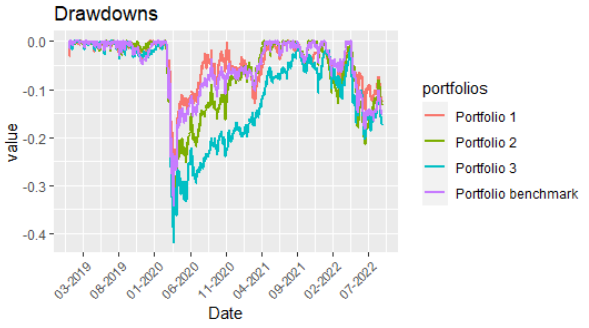
(c) mean-var model (2)



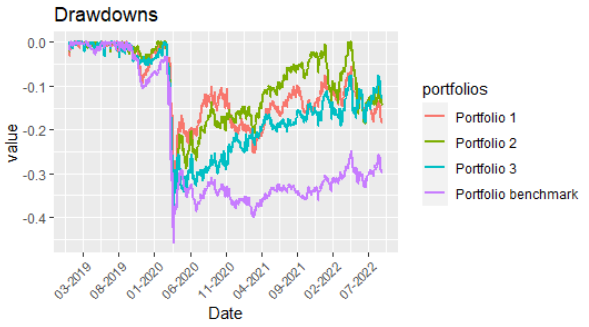
(d) mean-var cop GARCH model (1)



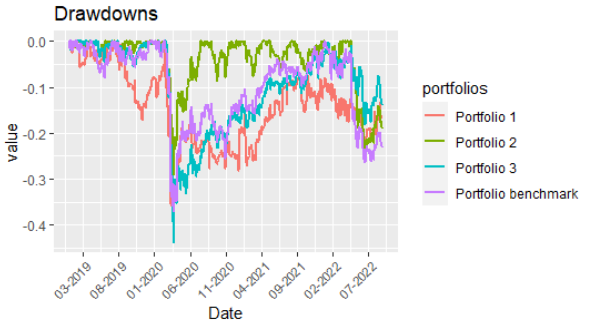
(e) $1/n$ model



(f) mean-var model (1)

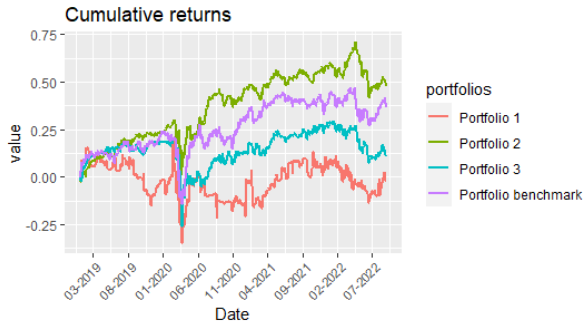


(g) mean-var model (2)

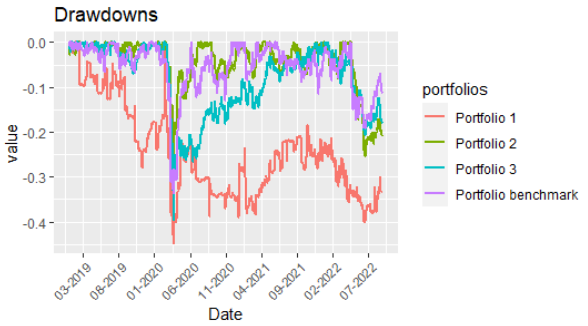


(h) mean-var cop GARCH model (1)

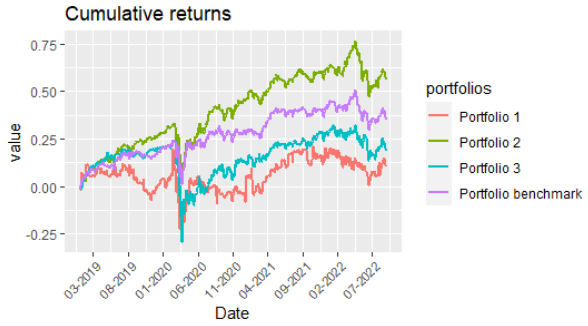
FIGURE A.1 – Model performances (1)



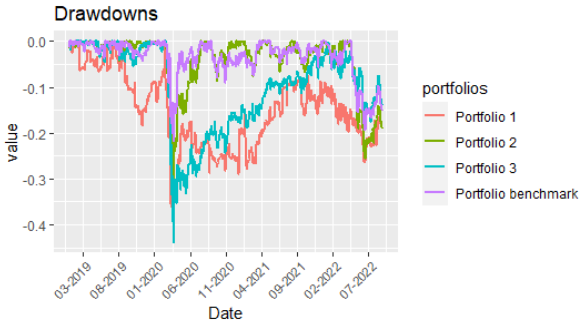
(a) mean-var cop GARCH model (2)



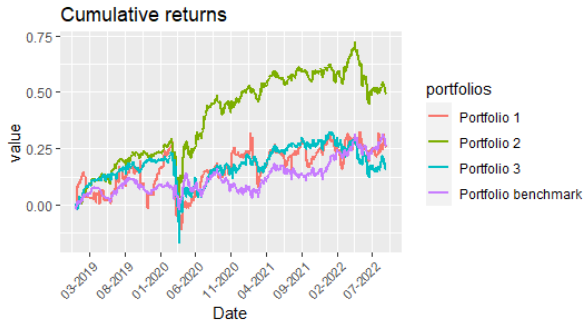
(e) mean-var cop GARCH model (2)



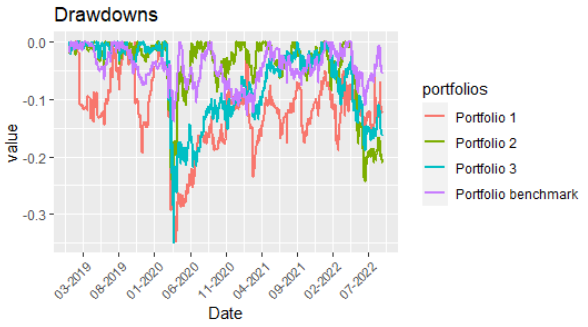
(b) mean-var DCC-GARCH model (1)



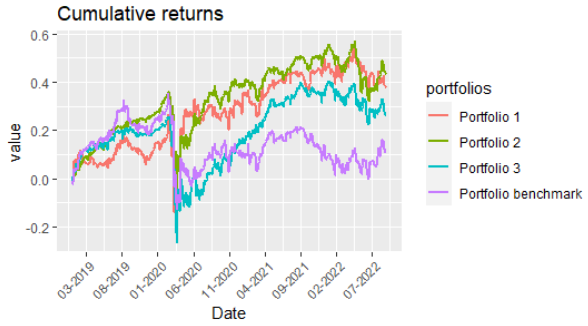
(f) mean-var DCC GARCH model (1)



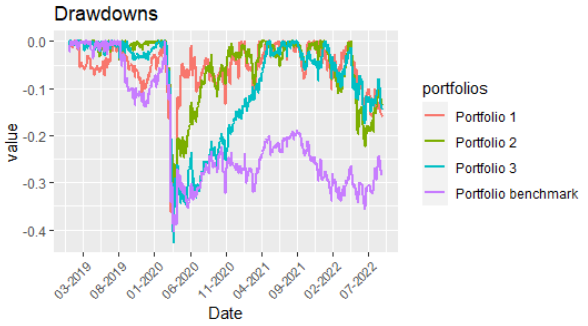
(c) mean-var DCC-GARCH model (2)



(g) mean-var DCC GARCH model (2)

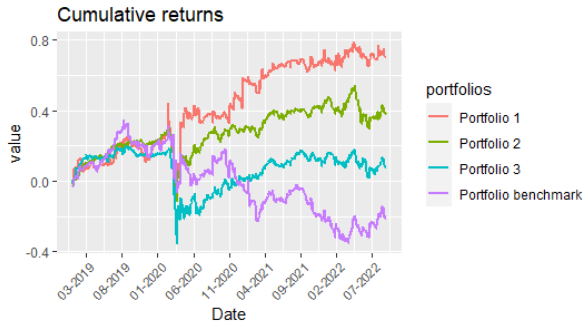


(d) mean-var PCC model (1)

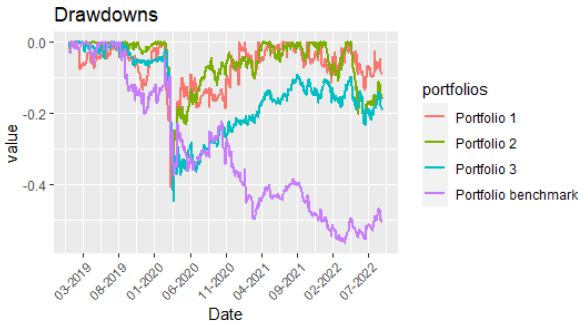


(h) mean-var PCC model (1)

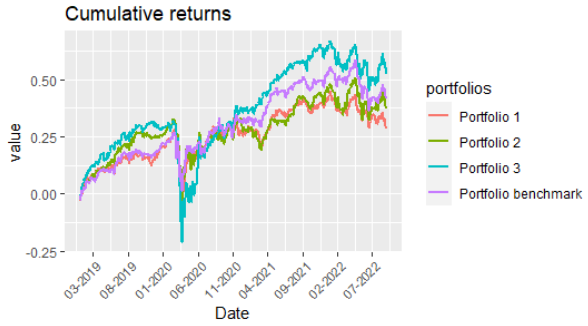
FIGURE A.2 – Model performances (2)



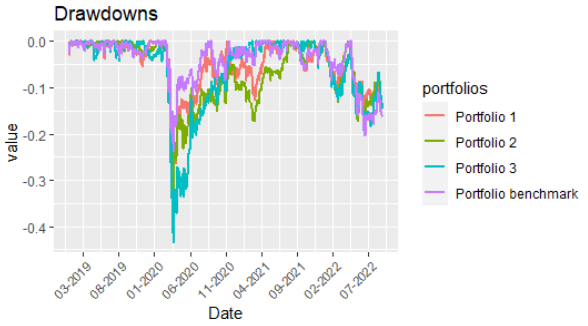
(a) mean-var PCC model (2)



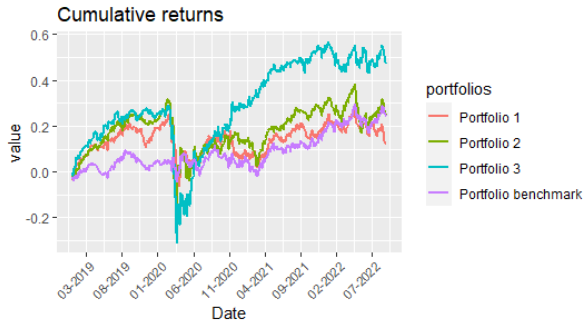
(e) mean-var PCC model (2)



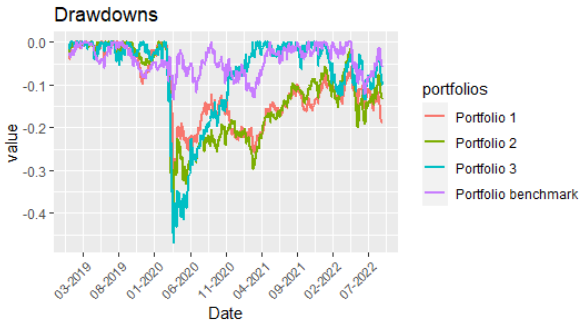
(b) mean-var SR PCC model (1)



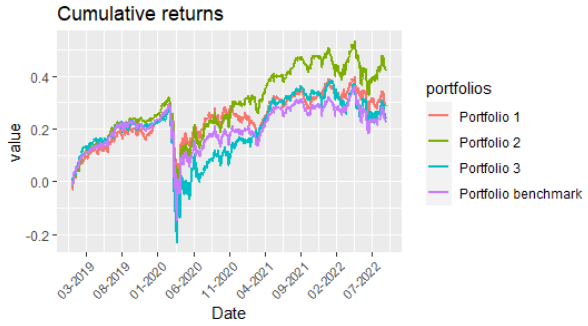
(f) mean-var SR PCC model (1)



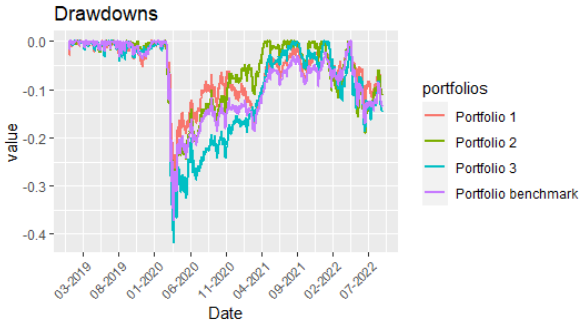
(c) mean-var SR PCC model (2)



(g) mean-var SR PCC model (2)

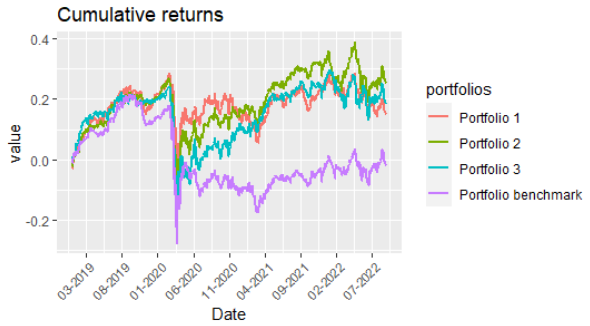


(d) mean-$LPM_{\tau=0,a=1}$ model (1)

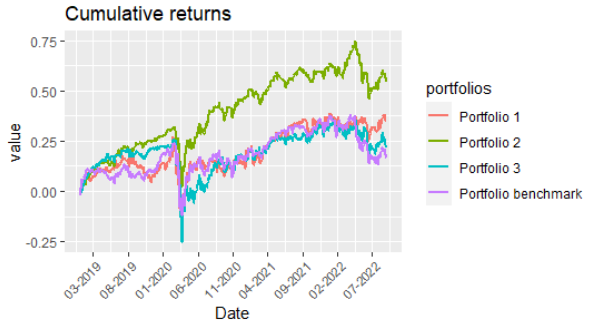


(h) mean-$LPM_{\tau=0,a=1}$ model (1)

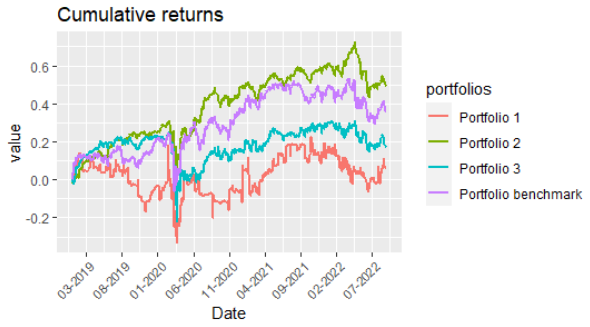
FIGURE A.3 – Model performances (3)



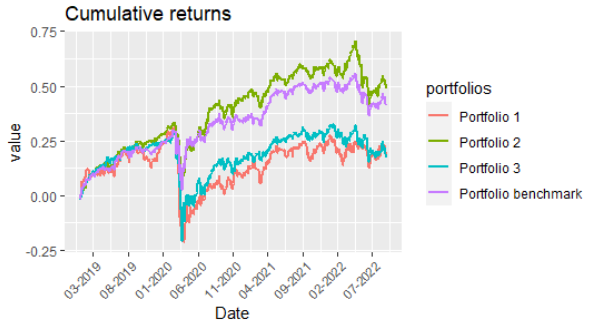
(a) mean-$LPM_{\tau=0,a=1}$ model (2)



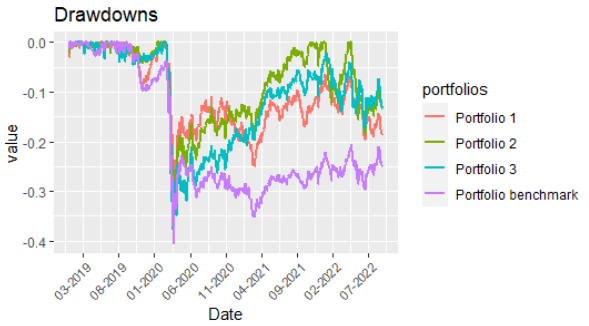
(b) mean-$LPM_{\tau=0,a=1}$ Copula-GARCH model(1)



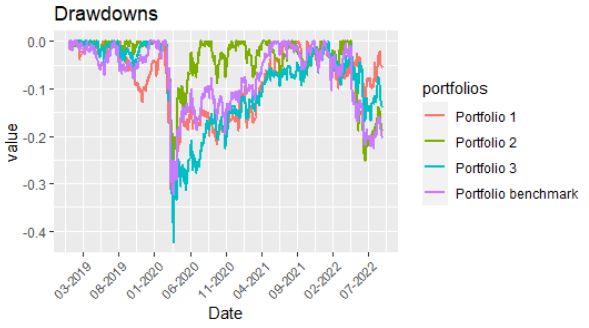
(c) mean-$LPM_{\tau=0,a=1}$ Copula-GARCH model (2)



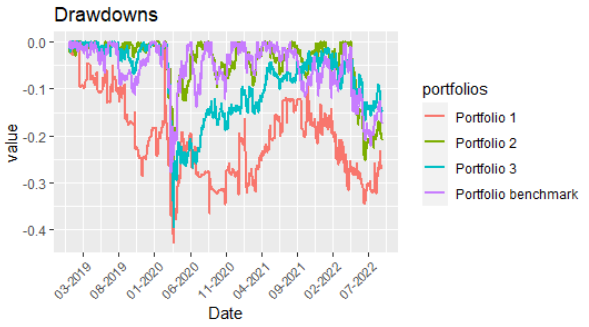
(d) mean-$LPM_{\tau=0,a=1}$ DCC-GARCH model (1)



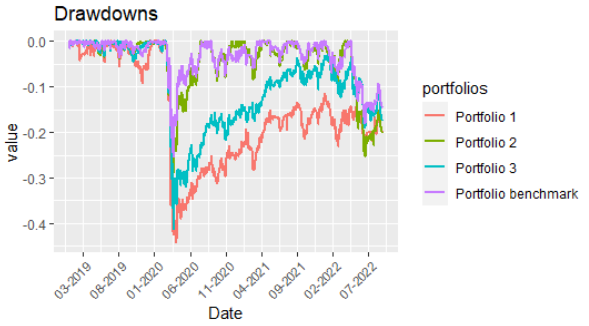
(e) mean-$LPM_{\tau=0,a=1}$ model (2)



(f) mean-$LPM_{\tau=0,a=1}$ Copula-GARCH model (1)

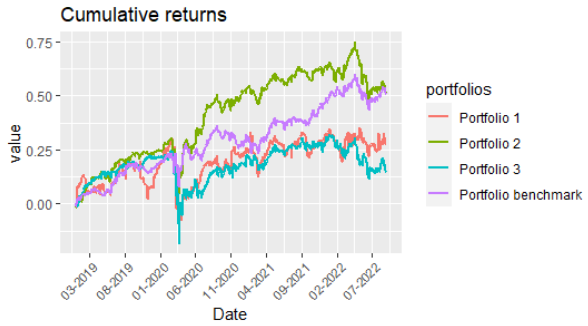


(g) mean-$LPM_{\tau=0,a=1}$ Copula-GARCH model (2)

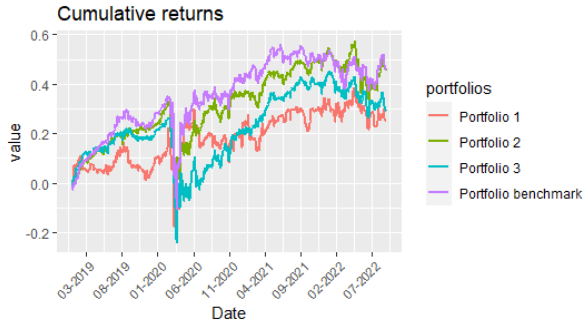


(h) mean-$LPM_{\tau=0,a=1}$ DCC-GARCH model (1)

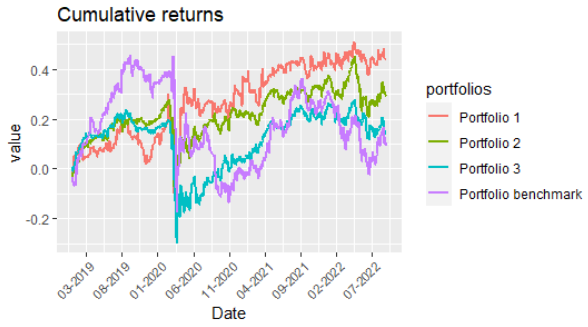
FIGURE A.4 – Model performances (4)



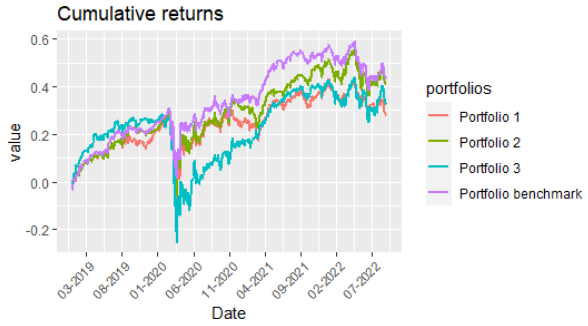
(a) mean-$LPM_{\tau=0,a=1}$ DCC-GARCH model (2)



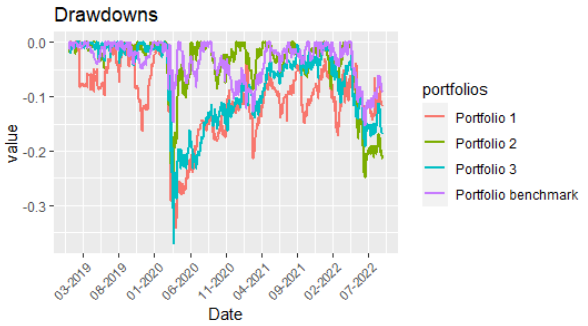
(b) mean-$LPM_{\tau=0,a=1}$ PCC model(1)



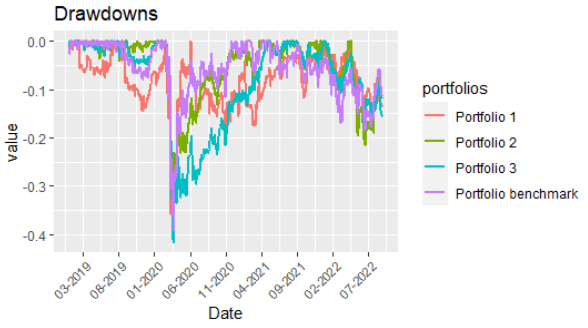
(c) mean-$LPM_{\tau=0,a=1}$ PCC model (2)



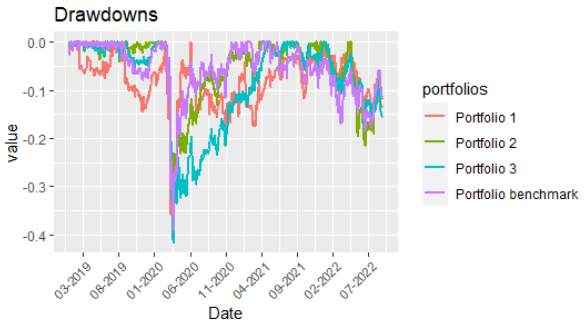
(d) mean-$LPM_{\tau=0,a=1}$ RS PCC model (1)



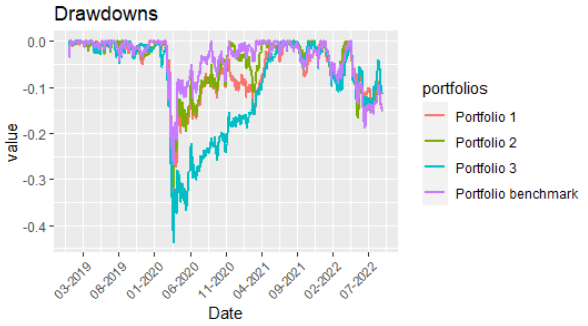
(e) mean-$LPM_{\tau=0,a=1}$ DCC-GARCH model (2)



(f) mean-$LPM_{\tau=0,a=1}$ PCC model (1)

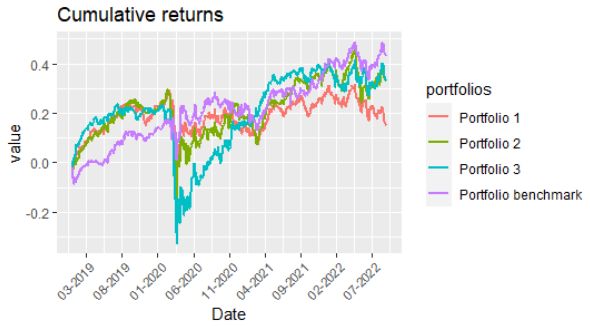


(g) mean-$LPM_{\tau=0,a=1}$ PCC model (2)

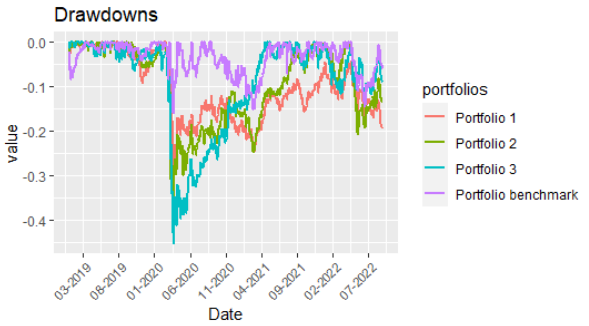


(h) mean-$LPM_{\tau=0,a=1}$ RS PCC model (1)

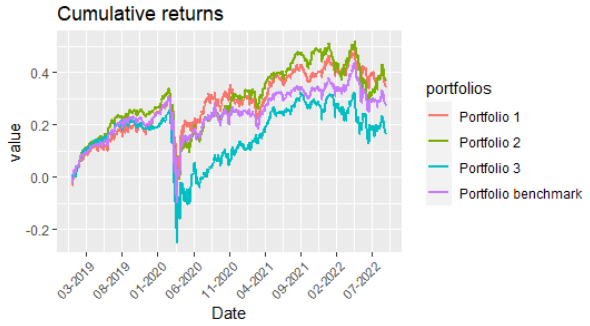
FIGURE A.5 – Model performances (5)



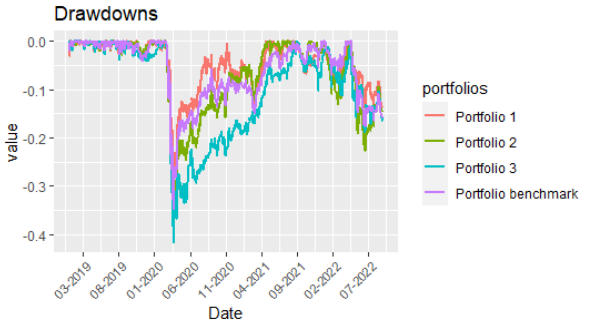
(a) mean-$LPM_{\tau=0,a=1}$ RS PCC model (2)



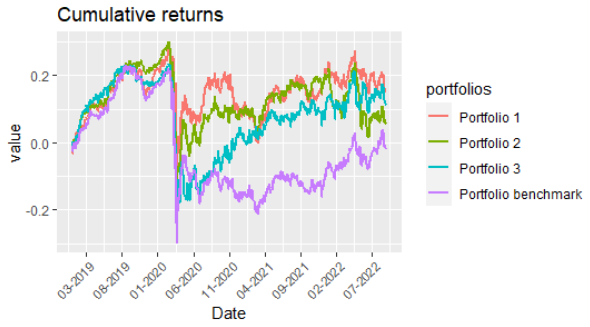
(e) mean-$LPM_{\tau=0,a=1}$ RS PCC model (2)



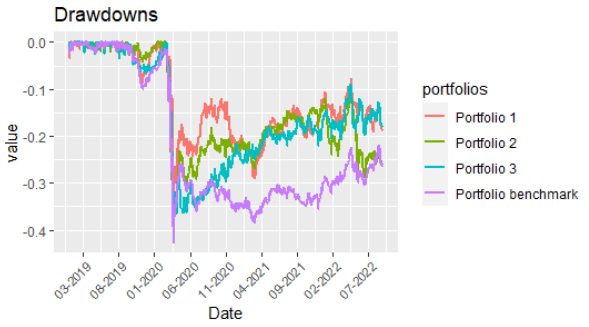
(b) mean-$LPM_{\tau=0,a=2}$ model(1)



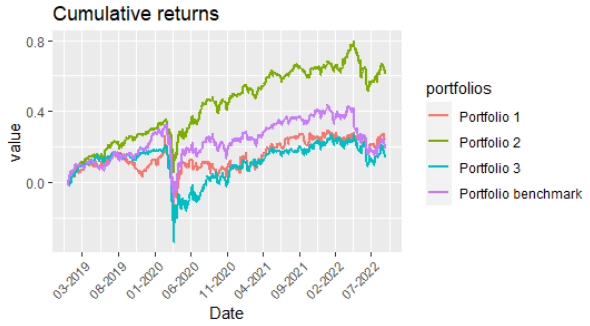
(f) mean-$LPM_{\tau=0,a=2}$ model (1)



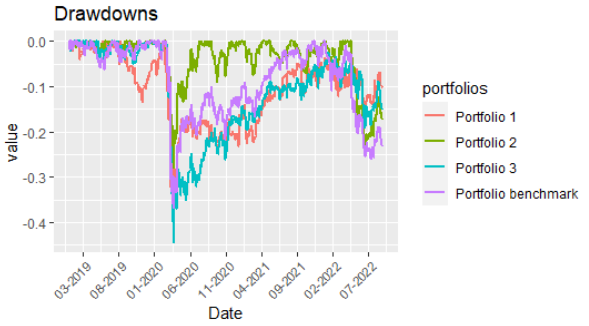
(c) mean-$LPM_{\tau=0,a=2}$ model (2)



(g) mean-$LPM_{\tau=0,a=2}$ model (2)

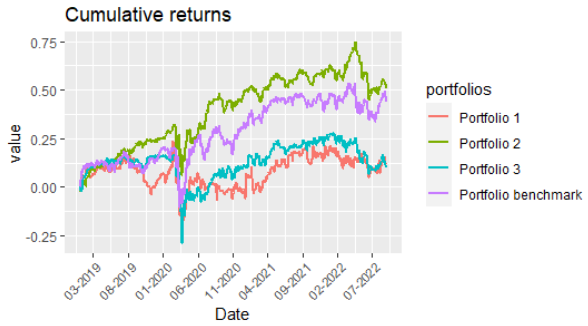


(d) mean-$LPM_{\tau=0,a=2}$ Copula-GARCH model (1)

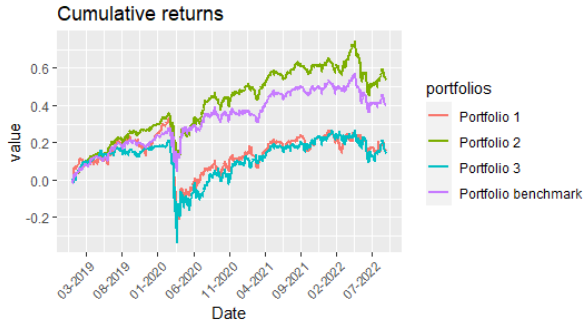


(h) mean-$LPM_{\tau=0,a=2}$ Copula-GARCH model (1)

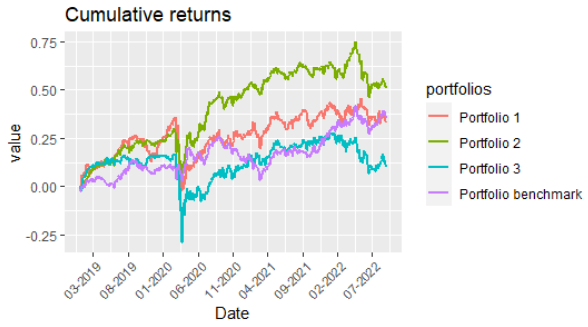
FIGURE A.6 – Model performances (6)



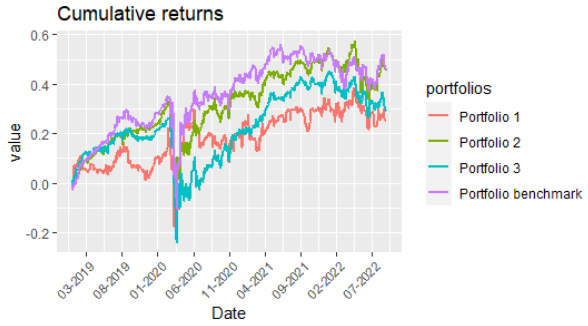
(a) mean-$LPM_{\tau=0,a=2}$ Copula-GARCH model (2)



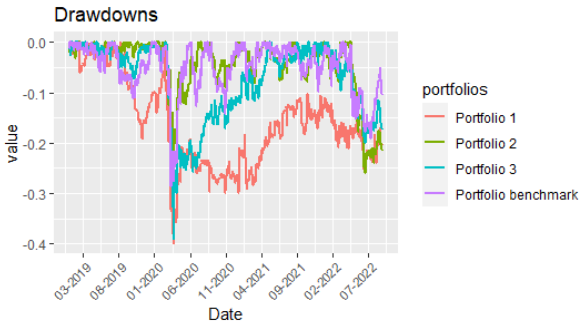
(b) mean-$LPM_{\tau=0,a=2}$ DCC GARCH model(1)



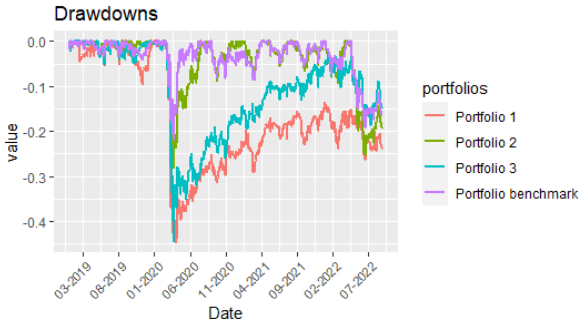
(c) mean-$LPM_{\tau=0,a=2}$ DCC GARCH model (2)



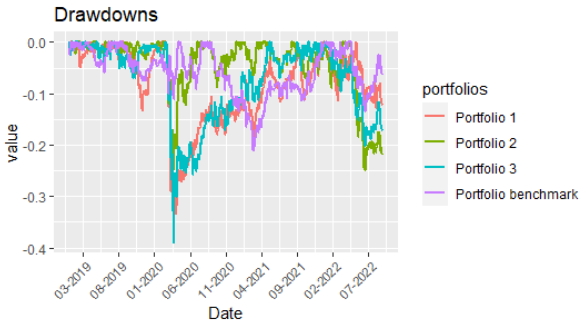
(d) mean-$LPM_{\tau=0,a=2}$ PCC GARCH model (1)



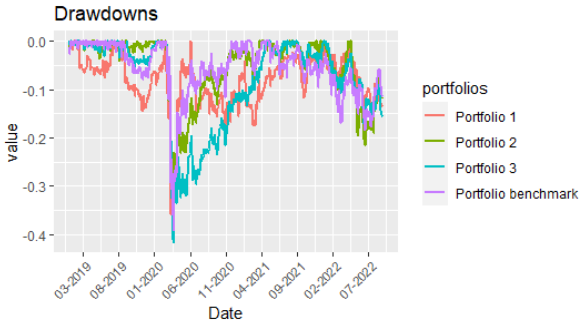
(e) mean-$LPM_{\tau=0,a=2}$ Copula-GARCH model (2)



(f) mean-$LPM_{\tau=0,a=2}$ DCC GARCH model (1)

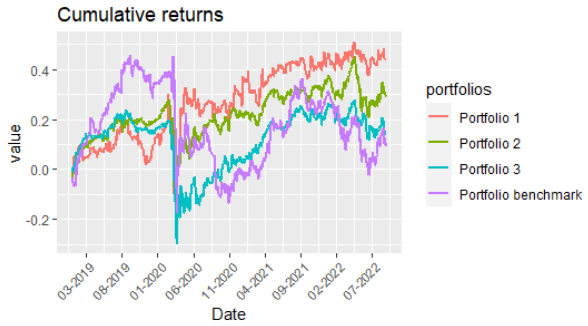


(g) mean-$LPM_{\tau=0,a=2}$ DCC GARCH model (2)

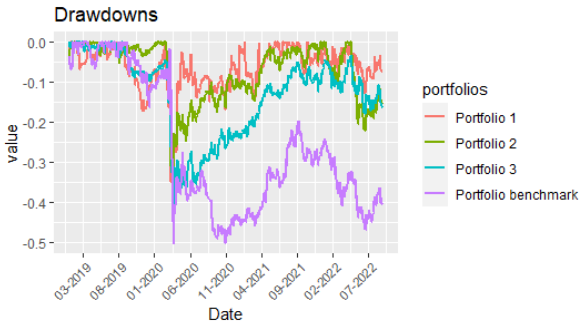


(h) mean-$LPM_{\tau=0,a=2}$ PCC GARCH model (1)

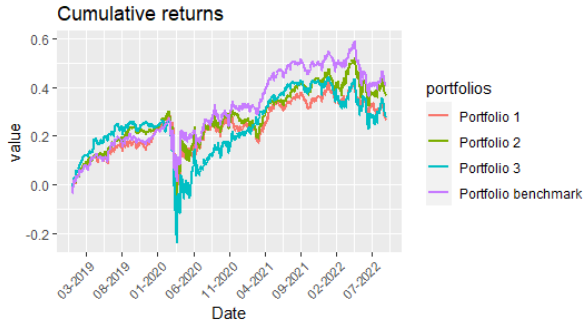
FIGURE A.7 – Model performances (7)



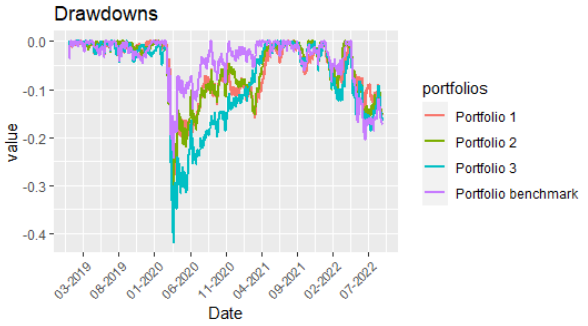
(a) mean-$LPM_{\tau=0,a=2}$ PCC GARCH model (2)



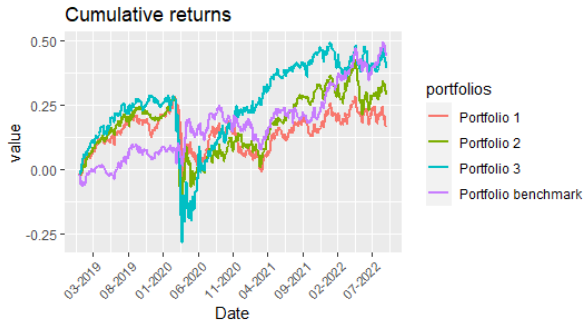
(d) mean-$LPM_{\tau=0,a=2}$ PCC GARCH model (2)



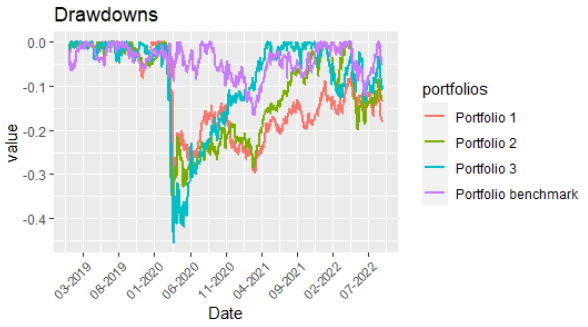
(b) mean-$LPM_{\tau=0,a=2}$ RS PCC GARCH model (1)



(e) mean-$LPM_{\tau=0,a=2}$ RS PCC GARCH model (1)



(c) mean-$LPM_{\tau=0,a=2}$ RS PCC GARCH model (2)



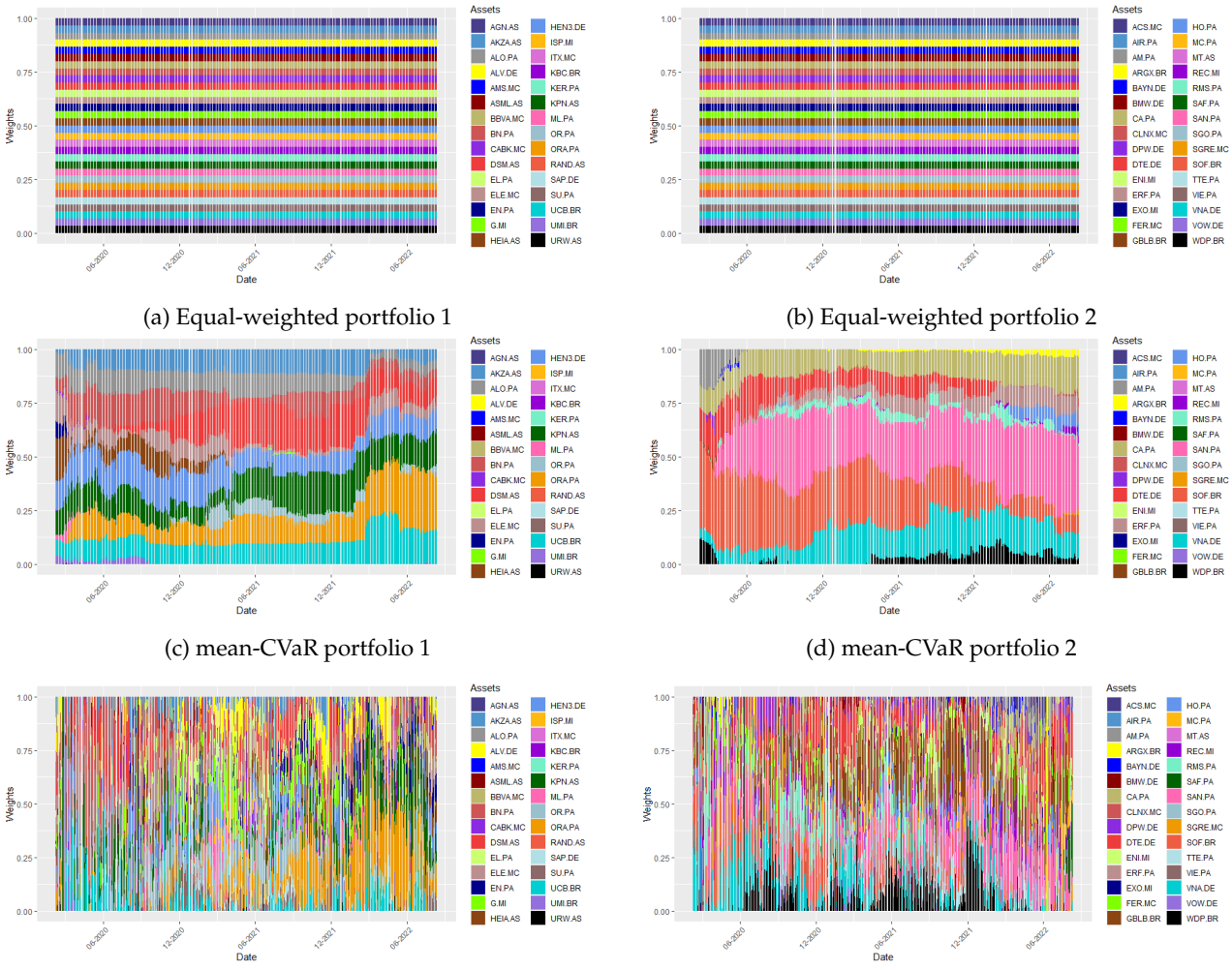
(f) mean-$LPM_{\tau=0,a=2}$ RS PCC GARCH model (2)

FIGURE A.8 – Model performances (8)

ANNEXE B

Annexe du chapitre 3

B.1 Plots of weights



(a) Equal-weighted portfolio 1

(b) Equal-weighted portfolio 2

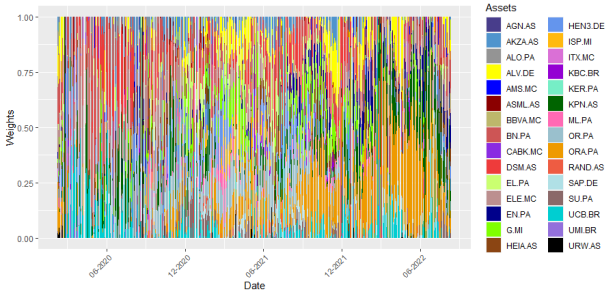
(c) mean-CVaR portfolio 1

(d) mean-CVaR portfolio 2

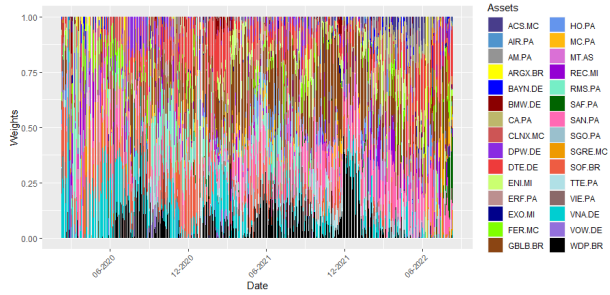
(e) gaussian copula mean-CVaR portfolio 1

(f) gaussian copula mean-CVaR portfolio 2

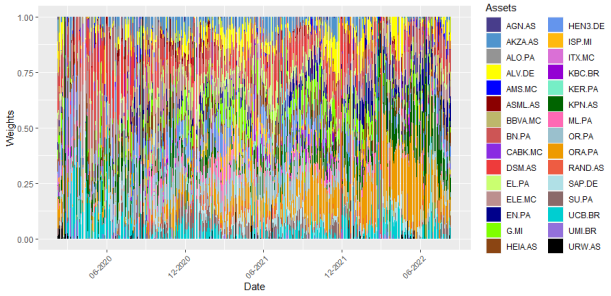
FIGURE B.1 – Les pondérations



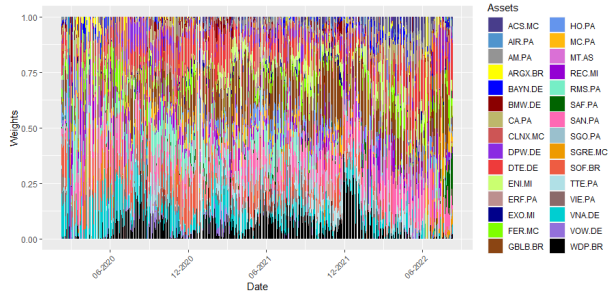
(a) worst case copula-CVaR 1 portfolio 1



(b) worst case copula-CVaR 1 portfolio 2



(c) worst case copula-CVaR 2 portfolio 1



(d) worst case copula-CVaR 2 portfolio 2

FIGURE B.2 – Les pondérations (suite)

ANNEXE C

Annexe du chapitre 4

C.1 Model performance

Models	RMSE	MAPE	NSE
ARIMA(3,1,0) N	55.8196	1.1888	0.9822
ARIMA(2,1,1) T	55.8352	1.1884	0.9822
ARIMA(3,1,0) N GARCH(1,1)	55.2656	1.1780	0.9825
ARIMA(2,1,1) T GARCH(1,1)	55.5243	1.1835	0.9824
ARIMA(3,1,0) N gjr-GARCH(1,1)	55.1519	1.1775	0.9826
ARIMA(2,1,1) T gjr-GARCH(1,1)	55.2326	1.1806	0.9825
SVR	35.2081	0.9215	0.9929
GP	32.2475	0.8295	0.9940
MLP	31.9272	0.8209	0.9942
LSTM	30.5332	0.8123	0.9947
GRU	30.2211	0.7768	0.9948
TCN	23.7150	0.6038	0.9968
ARIMA(3,1,0) N SVR	56.1839	1.1977	0.9819
ARIMA(2,1,1) T SVR	56.0953	1.1969	0.9820
ARIMA(3,1,0) N MLP	55.8586	1.1896	0.9821
ARIMA(2,1,1) T MLP	55.9401	1.1931	0.9821
ARIMA(3,1,0) N LSTM	56.0912	1.2024	0.9820
ARIMA(2,1,1) T LSTM	56.0621	1.2060	0.9820
ARIMA(3,1,0) N GRU	55.9633	1.1944	0.9821
ARIMA(2,1,1) T GRU	55.9575	1.1985	0.9821
ARIMA(3,1,0) N GP	56.1951	1.1918	0.9819
ARIMA(2,1,1) T GP	58.6480	1.2237	0.9803
ARIMA(3,1,0) N TCN	56.9778	1.1947	0.9814
ARIMA(2,1,1) T TCN	55.6544	1.1770	0.9823

(a) SP500

Models	RMSE	MAPE	NSE
ARIMA(2,1,0) N	0.7664	1.8723	0.9687
ARIMA(4,1,5) T	0.7665	1.8755	0.9687
ARIMA(2,1,0) N GARCH(1,1)	0.7643	1.8701	0.9689
ARIMA(4,1,5) T GARCH(1,1)	0.7642	1.8769	0.9689
ARIMA(2,1,0) N gjr-GARCH(1,1)	0.7646	1.8695	0.9688
ARIMA(4,1,5) T gjr-GARCH(1,1)	0.7655	1.8770	0.9688
SVR	0.5745	1.7186	0.9824
GP	0.5745	1.7186	0.9824
MLP	0.7639	1.8683	0.9689
LSTM	0.5154	1.5277	0.9858
GRU	0.5245	1.5527	0.9853
TCN	0.5230	1.5843	0.9854
ARIMA(2,1,0) N SVR	0.7709	1.8960	0.9683
ARIMA(4,1,5) T SVR	0.7679	1.8715	0.9686
ARIMA(2,1,0) N MLP	0.7664	1.8724	0.9687
ARIMA(4,1,5) T MLP	0.7670	1.8754	0.9687
ARIMA(2,1,0) N LSTM	0.7716	1.8740	0.9683
ARIMA(4,1,5) T LSTM	0.7688	1.8762	0.9685
ARIMA(2,1,0) N GRU	0.7758	1.8838	0.9679
ARIMA(4,1,5) T GRU	0.7682	1.8771	0.9686
ARIMA(2,1,0) N GP	0.7844	1.9129	0.9672
ARIMA(4,1,5) T GP	0.7749	1.8901	0.9680
ARIMA(2,1,0) N TCN	0.7892	1.9729	0.9668
ARIMA(4,1,5) T TCN	0.8111	2.0212	0.9649

(b) TTE.PA

Models	RMSE	MAPE	NSE
ARIMA(3,1,0) N	0.0052	0.3527	0.9875
ARIMA(2,1,3) T	0.0052	0.3518	0.9876
ARIMA(3,1,0) N GARCH(1,1)	0.0052	0.3529	0.9876
ARIMA(2,1,3) T GARCH(1,1)	0.0052	0.3533	0.9875
ARIMA(3,1,0) N gjr-GARCH(1,1)	0.0052	0.3524	0.9876
ARIMA(2,1,3) T gjr-GARCH(1,1)	0.0052	0.3526	0.9875
SVR	0.0050	0.3733	0.9887
GP	0.0052	0.3485	0.9878
MLP	0.0052	0.3505	0.9877
LSTM	0.0044	0.3238	0.9912
GRU	0.0046	0.3481	0.9904
TCN	0.0039	0.2907	0.9931
ARIMA(3,1,0) N SVR	0.0052	0.3536	0.9875
ARIMA(2,1,3) T SVR	0.0054	0.3657	0.9869
ARIMA(3,1,0) N MLP	0.0052	0.3527	0.9875
ARIMA(2,1,3) T MLP	0.0052	0.3523	0.9875
ARIMA(3,1,0) N LSTM	0.0052	0.3526	0.9875
ARIMA(2,1,3) T LSTM	0.0052	0.3523	0.9875
ARIMA(3,1,0) N GRU	0.0052	0.3530	0.9875
ARIMA(2,1,3) T GRU	0.0052	0.3522	0.9876
ARIMA(3,1,0) N GP	0.0052	0.3527	0.9875
ARIMA(2,1,3) T GP	0.0052	0.3518	0.9876
ARIMA(3,1,0) N TCN	0.0047	0.3552	0.9900
ARIMA(2,1,3) T TCN	0.0046	0.3543	0.9902

(c) EUR/USD

Models	RMSE	MAPE	NSE
ARIMA(4,1,5) N	29.2527	7.4667	0.9825
ARIMA(4,1,4) T	29.0794	7.1304	0.9827
ARIMA(4,1,5) N GARCH(1,1)	28.4803	7.2518	0.9834
ARIMA(4,1,4) T GARCH(1,1)	29.0794	7.1304	0.9827
ARIMA(4,1,5) N gjr-GARCH(1,1)	28.7547	7.1091	0.9831
ARIMA(4,1,4) T gjr-GARCH(1,1)	28.6244	7.1204	0.9832
SVR	271.2142	90.6921	-0.5057
GP	28.4639	7.5424	0.9834
MLP	30.4323	12.5026	0.9810
LSTM	36.6435	28.4023	0.9725
GRU	36.5489	28.5311	0.9727
TCN	33.8611	25.6001	0.9765
ARIMA(4,1,5) N SVR	29.0494	7.4029	0.9827
ARIMA(4,1,4) T SVR	28.6661	7.0874	0.9832
ARIMA(4,1,5) N MLP	29.2134	7.4642	0.9825
ARIMA(4,1,4) T MLP	28.9143	7.0991	0.9829
ARIMA(4,1,5) N LSTM	30.3065	8.2178	0.9812
ARIMA(4,1,4) T LSTM	31.5963	25.2595	0.9796
ARIMA(4,1,5) N GRU	29.3752	7.5237	0.9823
ARIMA(4,1,4) T GRU	29.7176	7.3208	0.9819
ARIMA(4,1,5) N GP	29.2660	7.4786	0.9825
ARIMA(4,1,4) T GP	28.8505	7.1566	0.9830
ARIMA(4,1,5) N TCN	29.3231	7.4560	0.9824
ARIMA(4,1,4) T TCN	28.9826	21.8245	0.9828

(d) JNUG

TABLE C.1 – L'ensemble des modèles de prédiction sur la période 1

Models	RMSE	MAPE	NSE
ARIMA(3,1,5) N	44.4264	0.7875	0.9731
ARIMA(4,1,4) T	43.5414	0.7617	0.9742
ARIMA(3,1,5) N GARCH(1,1)	44.4941	0.7756	0.9730
ARIMA(4,1,4) T GARCH(1,1)	39.1280	0.5989	0.9792
ARIMA(3,1,5) N gjr-GARCH(1,1)	44.1350	0.7735	0.9735
ARIMA(4,1,4) T gjr-GARCH(1,1)	35.5479	0.5344	0.9828
SVR	88.1714	1.1767	0.8942
GP	35.5304	0.5484	0.9828
MLP	34.1043	0.5054	0.9842
LSTM	38.9033	0.6012	0.9794
GRU	80.2651	1.1510	0.9123
TCN	38.9033	0.6012	0.9794
ARIMA(3,1,5) N SVR	44.9100	0.7968	0.9725
ARIMA(4,1,4) T SVR	38.8534	0.5922	0.9794
ARIMA(3,1,5) N MLP	44.4510	0.7891	0.9731
ARIMA(4,1,4) T MLP	43.8994	0.7747	0.9738
ARIMA(3,1,5) N LSTM	33.5177	0.4474	0.9847
ARIMA(4,1,4) T LSTM	30.1613	0.3814	0.9876
ARIMA(3,1,5) N GRU	44.4123	0.7877	0.9731
ARIMA(4,1,4) T GRU	32.7457	0.4279	0.9854
ARIMA(3,1,5) N GP	34.4903	0.4685	0.9838
ARIMA(4,1,4) T GP	34.7230	0.4556	0.9836
ARIMA(3,1,5) N TCN	48.5099	0.8431	0.9680
ARIMA(4,1,4) T TCN	27.9655	0.3201	0.9894

(a) SP500

Models	RMSE	MAPE	NSE
ARIMA(5,1,5) N	0.6439	1.2692	0.9822
ARIMA(5,1,5) T	0.6419	1.2623	0.9823
ARIMA(5,1,5) N GARCH(1,1)	0.6348	1.2315	0.9827
ARIMA(5,1,5) T GARCH(1,1)	0.6322	1.2360	0.9828
ARIMA(5,1,5) N gjr-GARCH(1,1)	0.6325	1.2427	0.9828
ARIMA(5,1,5) T gjr-GARCH(1,1)	0.6325	1.2427	0.9828
SVR	1.7551	1.9070	0.8677
GP	0.6396	1.2542	0.9824
MLP	0.6334	1.2411	0.9828
LSTM	0.5064	1.1958	0.9890
GRU	0.5731	0.9841	0.9859
TCN	0.3110	0.7396	0.9958
ARIMA(5,1,5) N SVR	0.6604	1.2908	0.9813
ARIMA(5,1,5) T SVR	0.6359	1.2530	0.9826
ARIMA(5,1,5) N MLP	0.6437	1.2689	0.9822
ARIMA(5,1,5) T MLP	0.6454	1.2706	0.9821
ARIMA(5,1,5) N LSTM	0.6452	1.2704	0.9821
ARIMA(5,1,5) T LSTM	0.4227	0.8610	0.9923
ARIMA(5,1,5) N GRU	0.6451	1.2698	0.9821
ARIMA(5,1,5) T GRU	0.6458	1.2701	0.9821
ARIMA(5,1,5) N GP	0.6670	1.3053	0.9809
ARIMA(5,1,5) T GP	0.6422	1.2657	0.9823
ARIMA(5,1,5) N TCN	0.6431	1.4204	0.9822
ARIMA(5,1,5) T TCN	0.6314	1.3166	0.9829

(b) TTE.PA

Models	RMSE	MAPE	NSE
ARIMA(3,1,0) N	0.0047	0.3105	0.9865
ARIMA(2,1,3) T	0.0047	0.3123	0.9865
ARIMA(3,1,0) N GARCH(1,1)	0.0047	0.3120	0.9865
ARIMA(2,1,3) T GARCH(1,1)	0.0047	0.3102	0.9866
ARIMA(3,1,0) N gjr-GARCH(1,1)	0.0047	0.3120	0.9865
ARIMA(2,1,3) T gjr-GARCH(1,1)	0.0047	0.3110	0.9865
SVR	0.0067	0.3426	0.9725
GP	0.0047	0.3113	0.9865
MLP	0.0040	0.2941	0.9904
LSTM	0.0023	0.1711	0.9967
GRU	0.0024	0.1800	0.9965
TCN	0.0019	0.1359	0.9978
ARIMA(3,1,0) N SVR	0.0049	0.3213	0.9855
ARIMA(2,1,3) T SVR	0.0047	0.3128	0.9866
ARIMA(3,1,0) N MLP	0.0047	0.3106	0.9865
ARIMA(2,1,3) T MLP	0.0047	0.3126	0.9865
ARIMA(3,1,0) N LSTM	0.0047	0.3123	0.9863
ARIMA(2,1,3) T LSTM	0.0047	0.3124	0.9865
ARIMA(3,1,0) N GRU	0.0048	0.3127	0.9862
ARIMA(2,1,3) T GRU	0.0047	0.3127	0.9865
ARIMA(3,1,0) N GP	0.0047	0.3101	0.9866
ARIMA(2,1,3) T GP	0.0047	0.3118	0.9865
ARIMA(3,1,0) N TCN	0.0070	0.5067	0.9705
ARIMA(2,1,3) T TCN	0.0047	0.3091	0.9867

(c) EUR/USD

Models	RMSE	MAPE	NSE
ARIMA(4,1,5) N	3.7286	3.5255	0.9547
ARIMA(1,1,5) T	0.4746	0.5233	0.9993
ARIMA(4,1,5) N GARCH(1,1)	3.6776	3.4910	0.9559
ARIMA(1,1,5) T GARCH(1,1)	0.2050	0.2306	0.9999
ARIMA(4,1,5) N gjr-GARCH(1,1)	3.6868	3.4948	0.9557
ARIMA(1,1,5) T gjr-GARCH(1,1)	3.6785	3.5267	0.9559
SVR	25.2148	32.2094	-1.0721
GP	3.7611	3.7052	0.9539
MLP	5.4177	6.0292	0.9043
LSTM	13.0677	5.3607	0.4435
GRU	11.0497	4.0949	0.6021
TCN	1.4747	1.6646	0.9929
ARIMA(4,1,5) N SVR	3.7329	3.5606	0.9546
ARIMA(1,1,5) T SVR	3.6662	3.5405	0.9562
ARIMA(4,1,5) N MLP	3.7300	3.5264	0.9547
ARIMA(1,1,5) T MLP	3.6926	3.5196	0.9556
ARIMA(4,1,5) N LSTM	3.6976	3.5438	0.9554
ARIMA(1,1,5) T LSTM	3.6472	3.5077	0.9566
ARIMA(4,1,5) N GRU	3.8924	3.5879	0.9506
ARIMA(1,1,5) T GRU	3.6868	3.5684	0.9557
ARIMA(4,1,5) N GP	3.7419	3.5507	0.9544
ARIMA(1,1,5) T GP	3.7214	3.5978	0.9549
ARIMA(4,1,5) N TCN	0.4555	0.4967	0.9993
ARIMA(1,1,5) T TCN	0.4581	0.5044	0.9993

(d) JNUG

TABLE C.2 – L'ensemble des modèles de prédiction sur la période 2

Models	RMSE	MAPE	NSE
ARIMA(4,1,2) N	58.3394	1.1344	0.9494
ARIMA(3,1,2) T	57.9391	1.1165	0.9501
ARIMA(4,1,2) N GARCH(1,1)	58.3413	1.1320	0.9494
ARIMA(3,1,2) T GARCH(1,1)	58.6261	1.1390	0.9489
ARIMA(4,1,2) N gjr-GARCH(1,1)	58.2430	1.1279	0.9496
ARIMA(3,1,2) T gjr-GARCH(1,1)	27.4110	0.5823	0.9888
SVR	59.8852	1.1559	0.9467
GP	57.4471	1.1097	0.9509
MLP	57.4717	1.1081	0.9509
LSTM	86.2838	1.3287	0.8893
GRU	77.9964	1.2645	0.9095
TCN	43.6003	0.9180	0.9717
ARIMA(4,1,2) N SVR	59.3442	1.1379	0.9476
ARIMA(3,1,2) T SVR	27.5931	0.5925	0.9887
ARIMA(4,1,2) N MLP	58.3340	1.1345	0.9494
ARIMA(3,1,2) T MLP	58.0641	1.1231	0.9499
ARIMA(4,1,2) N LSTM	58.5111	1.1384	0.9491
ARIMA(3,1,2) T LSTM	57.9272	1.1154	0.9501
ARIMA(4,1,2) N GRU	58.4873	1.1368	0.9491
ARIMA(3,1,2) T GRU	27.0831	0.5844	0.9891
ARIMA(4,1,2) N GP	59.2353	1.1545	0.9478
ARIMA(3,1,2) T GP	59.2797	1.1375	0.9477
ARIMA(4,1,2) N TCN	25.7229	0.5486	0.9902
ARIMA(3,1,2) T TCN	23.7824	0.5117	0.9916

(a) SP500

Models	RMSE	MAPE	NSE
ARIMA(1,1,4) N	0.9516	1.5160	0.9695
ARIMA(1,1,2) T	0.9434	1.4944	0.9701
ARIMA(1,1,4) N GARCH(1,1)	0.9454	1.4997	0.9699
ARIMA(1,1,2) T GARCH(1,1)	0.9423	1.4941	0.9701
ARIMA(1,1,4) N gjr-GARCH(1,1)	0.9454	1.4979	0.9699
ARIMA(1,1,2) T gjr-GARCH(1,1)	0.9425	1.4950	0.9701
SVR	2.6751	3.0495	0.7592
GP	0.9532	1.5231	0.9694
MLP	0.9505	1.5167	0.9696
LSTM	0.8328	1.4567	0.9767
GRU	2.3101	3.5183	0.8205
TCN	0.6614	1.1613	0.9853
ARIMA(1,1,4) N SVR	0.9655	1.5324	0.9686
ARIMA(1,1,2) T SVR	0.6135	1.0661	0.9873
ARIMA(1,1,4) N MLP	0.9518	1.5163	0.9695
ARIMA(1,1,2) T MLP	0.9500	1.5063	0.9696
ARIMA(1,1,4) N LSTM	0.9522	1.5084	0.9695
ARIMA(1,1,2) T LSTM	0.5097	0.8824	0.9913
ARIMA(1,1,4) N GRU	0.9491	1.5062	0.9697
ARIMA(1,1,2) T GRU	0.5410	0.9473	0.9902
ARIMA(1,1,4) N GP	0.9645	1.5307	0.9687
ARIMA(1,1,2) T GP	0.5913	1.0268	0.9882
ARIMA(1,1,4) N TCN	1.0463	1.6463	0.9632
ARIMA(1,1,2) T TCN	0.5311	0.9291	0.9905

(b) TTE.PA

Models	RMSE	MAPE	NSE
ARIMA(5,1,2) N	0.0064	0.4834	0.9801
ARIMA(4,1,5) T	0.0065	0.4857	0.9798
ARIMA(5,1,2) N GARCH(1,1)	0.0065	0.4860	0.9798
ARIMA(4,1,5) T GARCH(1,1)	0.0065	0.4882	0.9797
ARIMA(5,1,2) N gjr-GARCH(1,1)	0.0065	0.4848	0.9799
ARIMA(4,1,5) T gjr-GARCH(1,1)	0.0065	0.4869	0.9799
SVR	0.0029	0.2391	0.9958
GP	0.0030	0.2517	0.9955
MLP	0.0030	0.2495	0.9955
LSTM	0.0025	0.2045	0.9969
GRU	0.0025	0.2071	0.9969
TCN	0.0024	0.1974	0.9972
ARIMA(5,1,2) N SVR	0.0065	0.4888	0.9798
ARIMA(4,1,5) T SVR	0.0066	0.4968	0.9790
ARIMA(5,1,2) N MLP	0.0064	0.4834	0.9801
ARIMA(4,1,5) T MLP	0.0065	0.4857	0.9798
ARIMA(5,1,2) N LSTM	0.0064	0.4832	0.9801
ARIMA(4,1,5) T LSTM	0.0065	0.4867	0.9798
ARIMA(5,1,2) N GRU	0.0064	0.4833	0.9801
ARIMA(4,1,5) T GRU	0.0040	0.3301	0.9924
ARIMA(5,1,2) N GP	0.0064	0.4827	0.9801
ARIMA(4,1,5) T GP	0.0065	0.4857	0.9798
ARIMA(5,1,2) N TCN	0.0090	0.7064	0.9607
ARIMA(4,1,5) T TCN	0.0038	0.3139	0.9930

(c) EUR/USD

Models	RMSE	MAPE	NSE
ARIMA(4,1,1) N	2.4935	4.3519	0.9768
ARIMA(5,1,1) T	2.4663	4.2850	0.9773
ARIMA(4,1,1) N GARCH(1,1)	2.4572	4.2931	0.9774
ARIMA(5,1,1) T GARCH(1,1)	2.4663	4.2850	0.9773
ARIMA(4,1,1) N gjr-GARCH(1,1)	2.4549	4.2808	0.9775
ARIMA(5,1,1) T gjr-GARCH(1,1)	2.4553	4.2723	0.9775
SVR	4.2483	9.1418	0.9326
GP	2.9438	5.9032	0.9676
MLP	2.8564	5.3825	0.9695
LSTM	2.5806	5.6253	0.9751
GRU	2.5675	5.5493	0.9754
TCN	2.5615	5.5141	0.9755
ARIMA(4,1,1) N SVR	2.4808	4.3376	0.9770
ARIMA(5,1,1) T SVR	1.4105	3.0641	0.9926
ARIMA(4,1,1) N MLP	2.4928	4.3504	0.9768
ARIMA(5,1,1) T MLP	1.4234	3.1097	0.9924
ARIMA(4,1,1) N LSTM	2.4312	4.2681	0.9779
ARIMA(5,1,1) T LSTM	1.3347	2.9219	0.9933
ARIMA(4,1,1) N GRU	1.4028	3.0685	0.9926
ARIMA(5,1,1) T GRU	1.5722	3.4058	0.9908
ARIMA(4,1,1) N GP	1.5522	3.4321	0.9910
ARIMA(5,1,1) T GP	1.6084	3.4330	0.9903
ARIMA(4,1,1) N TCN	2.4972	4.3594	0.9767
ARIMA(5,1,1) T TCN	1.1082	2.4595	0.9954

(d) JNUG

TABLE C.3 – L'ensemble des modèles de prédiction sur la période 3

Bibliographie

K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance : Mathematics and economics*, 44(2) :182–198, 2009.

L. Abramson and D. Chung. Socially responsible investing : Viable for value investors ? *The Journal of Investing*, 9(3) :73–80, 2000.

M.-V. Achim and S. N. Borlea. Developing of esg score to assess the non-financial performances in romanian companies. *Procedia Economics and Finance*, 32 :1209–1224, 2015.

Z. Adams, R. Füss, and T. Glück. Are correlations constant ? empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance*, 84 :9–24, 2017.

A. F. M. Agarap. A neural network architecture combining gated recurrent unit (gru) and support vector machine (svm) for intrusion detection in network traffic data. In *Proceedings of the 2018 10th International Conference on Machine Learning and Computing*, pages 26–30, 2018.

C. Almeida and C. Czado. Efficient bayesian inference for stochastic time-varying copula models. *Computational Statistics & Data Analysis*, 56(6) :1511–1527, 2012.

A. Amel-Zadeh and G. Serafeim. Why and how investors use esg information : Evidence from a global survey. *Financial Analysts Journal*, 74(3) :87–103, 2018.

K. P. Anagnostopoulos and G. Mamanis. A portfolio optimization model with three objectives and discrete variables. *Computers & Operations Research*, 37(7) :1285–1297, 2010.

L. Anastasakis and N. Mort. Neural network-based prediction of the usd/gbp exchange rate : the utilisation of data compression techniques for input dimension reduction. In *Proceedings of the Nostradamus 2000 : 3rd Conference on prediction, synergetic, behavior of dynamic systems and its application in control theory, physics, chemistry and economy*, pages 3–9, 2000.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realized stock return volatility. *Journal of financial economics*, 61(1) :43–76, 2001.

F. Andersson, H. Mausser, D. Rosen, and S. Uryasev. Credit risk optimization with conditional value-at-risk criterion. *Mathematical programming*, 89(2) :273–291, 2001.

A. Ang and G. Bekaert. International asset allocation with regime shifts. *The review of financial studies*, 15(4) : 1137–1187, 2002.

A. Ang and J. Chen. Asymmetric correlations of equity portfolios. *Journal of financial Economics*, 63(3) :443–494, 2002.

A. Ansar, B. Caldecott, and J. Tilbury. Stranded assets and the fossil fuel divestment campaign : what does divestment mean for the valuation of fossil fuel assets ? 2013.

D.-L. Arjaliès. Institutional change in the making-the case of socially responsible investment. *Available at SSRN 1932295*, 2010.

J. Arreola Hernandez, S. Hammoudeh, D. K. Nguyen, M. A. Al Janabi, and J. C. Reboredo. Global financial crisis and dependence risk analysis of sector portfolios : a vine copula approach. *Applied Economics*, 49(25) : 2409–2427, 2017.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3) :203–228, 1999.

S. Asadi, A. Tavakoli, and S. R. Hejazi. A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization. *Expert Systems with Applications*, 39(5) :5332–5337, 2012.

M. C. Ausin and H. F. Lopes. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis*, 54(11) :2383–2399, 2010.

E. M. Azoff. *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.

F. B Sabino da Silva and F. Ziegelman. Robust portfolio optimization with multivariate copulas : A worst-case cvar approach. *Flávio, Robust Portfolio Optimization with Multivariate Copulas : A Worst-Case CVaR Approach (November 22, 2017)*, 2017.

S. Babaei, M. M. Sepehri, and E. Babaei. Multi-objective portfolio optimization considering the dependence structure of asset returns. *European Journal of Operational Research*, 244(2) :525–539, 2015.

S. Bai, J. Z. Kolter, and V. Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv :1803.01271*, 2018.

Z. Bai, H. Liu, and W.-K. Wong. Enhancement of the applicability of markowitz's portfolio optimization by utilizing random matrix theory. *Mathematical Finance : An International Journal of Mathematics, Statistics and Financial Economics*, 19(4) :639–667, 2009.

E. Ballestero, M. Bravo, B. Pérez-Gladish, M. Arenas-Parra, and D. Pla-Santamaria. Socially responsible investment : A multicriteria approach to portfolio selection combining ethical and financial objectives. *European Journal of Operational Research*, 216(2) :487–494, 2012.

N. Barberis. Investing for the long run when returns are predictable. *The Journal of Finance*, 55(1) :225–264, 2000.

M. L. Barnett and R. M. Salomon. Beyond dichotomy : The curvilinear relationship between social responsibility and financial performance. *Strategic management journal*, 27(11) :1101–1122, 2006.

R. Bauer and D. Hann. Corporate environmental management and credit risk. *Available at SSRN 1660470*, 2010.

L. Bauwens, B. De Backer, and A. Dufays. A bayesian method of change-point estimation with recurrent regimes : Application to garch models. *Journal of Empirical Finance*, 29 :207–229, 2014.

V. S. Bawa. Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2(1) :95–121, 1975.

V. S. Bawa. Safety-first, stochastic dominance, and optimal portfolio choice. *Journal of Financial and Quantitative Analysis*, 13(2) :255–271, 1978.

V. S. Bawa and E. B. Lindenberg. Capital market equilibrium in a mean-lower partial moment framework. *Journal of financial economics*, 5(2) :189–200, 1977.

K. Bax, Ö. Sahin, C. Czado, and S. Paterlini. Esg, risk, and (tail) dependence. *Available at SSRN 3846739*, 2021.

T. Bedford and R. M. Cooke. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, 32(1) :245–268, 2001.

T. Bedford and R. M. Cooke. Vines–a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4) :1031–1068, 2002.

A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations research letters*, 25 (1) :1–13, 1999.

S. Benati. The optimal portfolio problem with coherent risk measure constraints. *European Journal of Operational Research*, 150(3) :572–584, 2003.

C. Bengtsson and J. Holst. *On portfolio selection : Improved covariance matrix estimation for Swedish asset returns*. Citeseer, 2002.

A. BenSaïda. The contagion effect in european sovereign debt markets : A regime-switching vine copula approach. *International Review of Financial Analysis*, 58 :153–165, 2018.

D. Berg and K. Aas. Models for construction of multivariate dependence : A comparison study. *The European Journal of Finance*, 15(7) :639–659, 2009.

F. Berg, J. F. Koelbel, and R. Rigobon. Aggregate confusion : The divergence of esg ratings. *Review of Finance*, 26(6) :1315–1344, 2022.

A. Bernal, S. Fok, and R. Pidaparthi. Financial market time series prediction with recurrent neural networks. *State College : Citeseer.[Google Scholar]*, 2012.

J. Bernardo, J. Berger, A. Dawid, A. Smith, et al. Regression and classification using gaussian process priors. *Bayesian statistics*, 6 :475, 1998.

A. Bilbao-Terol, M. Arenas-Parra, and V. Cañal-Fernández. A fuzzy multi-objective approach for sustainable investments. *Expert Systems with Applications*, 39(12) :10904–10915, 2012.

M. Billio, M. Costola, I. Hristova, C. Latino, and L. Pelizzon. Inside the esg ratings :(dis) agreement and performance. *Corporate Social Responsibility and Environmental Management*, 28(5) :1426–1445, 2021.

R. C. Bishop, K. J. Boyle, R. T. Carson, D. Chapman, W. M. Hanemann, B. Kanninen, R. J. Kopp, J. A. Krosnick, J. List, N. Meade, et al. Putting a value on injuries to natural assets : The bp oil spill. *Science*, 356(6335) : 253–254, 2017.

F. Black and R. Litterman. Global portfolio optimization. *Financial analysts journal*, 48(5) :28–43, 1992.

T. Bodnar, S. Mazur, and K. Podgórski. Singular inverse wishart distribution and its application to portfolio theory. *Journal of Multivariate Analysis*, 143 :314–326, 2016.

T. Bodnar, S. Mazur, and Y. Okhrin. Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research*, 256(1) :292–307, 2017a.

T. Bodnar, S. Mazur, and K. Podgórski. A test for the global minimum variance portfolio for small sample and singular covariance. *AStA Advances in Statistical Analysis*, 101 :253–265, 2017b.

T. Bodnar, S. Mazur, K. Podgórski, and J. Tyrcha. Tangency portfolio weights for singular covariance matrix in small and large dimensions : Estimation and test theory. *Journal of Statistical Planning and Inference*, 201 : 40–57, 2019.

T. Bodnar, H. Dette, N. Parolya, and E. Thorsén. Sampling distributions of optimal portfolio weights and characteristics in small and large dimensions. *Random Matrices : Theory and Applications*, 11(01) :2250008, 2022.

M. T. Bohl, P. Kaufmann, and P. M. Stephan. From hero to zero : Evidence of performance reversal and speculative bubbles in german renewable energy stocks. *Energy Economics*, 37 :40–51, 2013.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3) :307–327, 1986.

T. Bollerslev. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, pages 542–547, 1987.

T. Bollerslev. Modelling the coherence in short-run nominal exchange rates : a multivariate generalized arch model. *The review of economics and statistics*, pages 498–505, 1990.

B. Bolton. *Sustainable Financial Investments : Maximizing Corporate Profits and Long-Term Economic Value Creation*. Springer, 2016.

F. Boutin-Dufresne and P. Savaria. Corporate social responsibility and financial risk. *The Journal of investing*, 13(1) :57–66, 2004.

H. Bowen. Social responsibilities of the businessman. harper & row : N. y. *Y, USA*, 1953.

H. R. Bowen. *Social responsibilities of the businessman*. University of Iowa Press, 2013.

G. E. Box and G. M. Jenkins. Time series analysis. forecasting and control. *tsaf*, 1976.

L. Brandstetter and O. M. Lehner. Opening the market for impact investments : The need for adapted portfolio tools. *Entrepreneurship Research Journal*, 5(2) :87–107, 2015.

E. Brechmann, C. Czado, and S. Paterlini. Flexible dependence modeling of operational risk losses and its impact on total capital requirements. *Journal of Banking & Finance*, 40 :271–285, 2014.

M. Broadie. Computing efficient frontiers using estimated parameters. *Annals of operations research*, 45(1) : 21–58, 1993.

Z. Cai and X. Wang. Selection of mixed copula model via penalized likelihood. *Journal of the American Statistical Association*, 109(506) :788–801, 2014.

L. Cao and F. E. Tay. Financial forecasting using support vector machines. *Neural Computing & Applications*, 10(2) :184–192, 2001.

Y. Cao, Y. Ding, M. Jia, and R. Tian. A novel temporal convolutional network with residual self-attention mechanism for remaining useful life prediction of rolling bearings. *Reliability Engineering & System Safety*, 215 :107813, 2021.

G. Capelle-Blancard and A. Petit. Every little helps ? esg news and stock market reaction. *Journal of Business Ethics*, 157(2) :543–565, 2019.

L. Cappiello, R. F. Engle, and K. Sheppard. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial econometrics*, 4(4) :537–572, 2006.

A. B. Carroll. A three-dimensional conceptual model of corporate performance. *Academy of management review*, 4(4) :497–505, 1979.

S. Chava. Environmental externalities and cost of capital. *Management science*, 60(9) :2223–2247, 2014.

A. H. Chen, F. J. Fabozzi, and D. Huang. Portfolio revision under mean-variance and mean-cvar with transaction costs. *Review of Quantitative Finance and Accounting*, 39(4) :509–526, 2012.

L. Chen, L. Zhang, J. Huang, H. Xiao, and Z. Zhou. Social responsibility portfolio optimization incorporating esg criteria. *Journal of Management Science and Engineering*, 6(1) :75–85, 2021.

P. Y. Chen, M. Smithson, P. M. Popovich, C. ter Y, et al. *Correlation : Parametric and nonparametric measures*. Number 139. Sage, 2002.

K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio. On the properties of neural machine translation : Encoder-decoder approaches. *arXiv preprint arXiv :1409.1259*, 2014a.

K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv :1406.1078*, 2014b.

H. K. Choi. Stock price correlation coefficient prediction with arima-lstm hybrid model. *arXiv preprint arXiv :1808.01560*, 2018.

L. Chollete, V. H. de la Pena, and C.-C. Lu. Comovement of international financial markets. *Available at SSRN 675382*, 2005.

L. Chollete, A. Heinen, and A. Valdesogo. Modeling international financial returns with a multivariate regime-switching copula. *Journal of financial econometrics*, 7(4) :437–480, 2009.

V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2) :6–11, 1993.

V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. In *Handbook of the fundamentals of financial decision making : Part I*, pages 365–373. World Scientific, 2013.

K. V. Chow and K. C. Denning. On variance and lower partial moment betas the equivalence of systematic risk measures. *Journal of Business Finance & Accounting*, 21(2) :231–241, 1994.

J. Chung, C. Gulcehre, K. Cho, and Y. Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv :1412.3555*, 2014.

A. Clare, R. O'Brien, S. Thomas, and M. Wickens. Macroeconomic shocks and the capm : Evidence from the uk stockmarket. *International Journal of Finance & Economics*, 3(2) :111–126, 1998.

D. G. Clayton. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1) :141–151, 1978.

R. Clow. Money that grows on trees. *INSTITUTIONAL INVESTOR-NEW YORK-*, 33 :212–216, 1999.

G. Constantinides and A. Malliaris. Portfolio theory, finance ed ra jarrow, v. maksimovic and wt ziemba, 1995.

G. Cornuejols and R. Tütüncü. *Optimization methods in finance*, volume 5. Cambridge University Press, 2006.

A. Costinot, T. Roncalli, and J. Teiletche. Revisiting the dependence between financial markets with copulas. *Available at SSRN 1032535*, 2000.

D. Cumova and D. Nawrocki. A symmetric lpm model for heuristic mean–semivariance analysis. *Journal of Economics and Business*, 63(3) :217–236, 2011.

G. Cybenko. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(4) :359–366, 1989.

C. Czado. Pair-copula constructions of multivariate copulas. In *Copula theory and its applications*, pages 93–109. Springer, 2010.

C. Czado, E. C. Brechmann, and L. Gruber. Selection of vine copulas. In *Copulae in Mathematical and Quantitative Finance*, pages 17–37. Springer, 2013.

L. D'Antonio, T. Johnsen, and R. B. Hutton. Socially responsible investing and asset allocation. *The Journal of Investing*, 9(3) :65–72, 2000.

M. L. DeFond, R. N. Hann, and X. Hu. Does the market value financial expertise on audit committees of boards of directors ? *Journal of accounting research*, 43(2) :153–193, 2005.

R. Della Croce, F. Stewart, and J. Yermo. Promoting longer-term investment by institutional investors : selected issues and policies. *OECD Journal : Financial Market Trends*, 2011(1) :145–164, 2011.

S. Demarta and A. J. McNeil. The t copula and related copulas. *International statistical review*, 73(1) :111–129, 2005.

E. Demers, J. Hendrikse, P. Joos, and B. Lev. Esg did not immunize stocks during the covid-19 crisis, but investments in intangible assets did. *Journal of Business Finance & Accounting*, 48(3-4) :433–462, 2021.

V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification : How inefficient is the 1/n portfolio strategy ? *The review of Financial studies*, 22(5) :1915–1953, 2009.

X. Deng and Y. Liang. Robust portfolio optimization based on semi-parametric arma-tgarch-evt model with mixed copula using wcvar. *Computational Economics*, pages 1–28, 2021.

A. Dhara, B. Das, and K. Natarajan. Worst-case expected shortfall with univariate and bivariate marginals. *INFORMS Journal on Computing*, 33(1) :370–389, 2021.

V. Díaz, D. Ibrushi, and J. Zhao. Reconsidering systematic factors during the covid-19 pandemic–the rising importance of esg. *Finance Research Letters*, 38 :101870, 2021.

X. Ding, Y. Zhang, T. Liu, and J. Duan. Deep learning for event-driven stock prediction. In *Twenty-fourth international joint conference on artificial intelligence*, 2015.

J. Dissmann, E. C. Brechmann, C. Czado, and D. Kurowicka. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, 59 :52–69, 2013.

J. F. Dißmann. Statistical inference for regular vines and application. 2010.

N. Dobrovolskienė and R. Tamošiūnienė. Sustainability-oriented financial resource allocation in a project portfolio through multi-criteria decision-making. *Sustainability*, 8(5) :485, 2016.

M. J. Dueker. Markov switching in garch processes and mean-reverting stock-market volatility. *Journal of Business & Economic Statistics*, 15(1) :26–34, 1997.

M. Ehrgott, K. Klamroth, and C. Schwehm. An mcdm approach to portfolio optimization. *European Journal of Operational Research*, 155(3) :752–770, 2004.

D. Ellsberg. Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669, 1961.

E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann. *Modern portfolio theory and investment analysis*. John Wiley & Sons, 2009.

P. Embrechts, E. Mcneil, and D. Straumann. Correlation : pitfalls and alternatives. In *Risk Magazine*. Citeseer, 1999.

P. Embrechts, F. Lindskog, and A. McNeil. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*, 14, 2001.

P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management : properties and pitfalls. *Risk management : value at risk and beyond*, 1 :176–223, 2002.

R. Engle. Dynamic conditional correlation : A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3) :339–350, 2002.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica : Journal of the Econometric Society*, pages 987–1007, 1982.

R. F. Engle III and K. Sheppard. Theoretical and empirical properties of dynamic conditional correlation multivariate garch, 2001.

M. Escobar-Anel. Multivariate risk aversion utility, application to esg investments. *The North American Journal of Economics and Finance*, 63 :101790, 2022.

F. J. Fabozzi, S. T. Rachev, and C. Menn. *Fat-tailed and skewed asset return distributions : implications for risk management, portfolio selection, and option pricing*. John Wiley & Sons, 2005.

F. J. Fabozzi, S. M. Focardi, P. N. Kolm, and D. A. Pachamanova. *Robust portfolio optimization and management*. John Wiley & Sons, 2007.

F. J. Fabozzi, K. Ma, and B. J. Oliphant. Sin stock returns. *The Journal of Portfolio Management*, 35(1) :82–94, 2008.

E. F. Fama. Mandelbrot and the stable paretian hypothesis. *The journal of business*, 36(4) :420–429, 1963.

E. F. Fama. The behavior of stock-market prices. *The journal of Business*, 38(1) :34–105, 1965.

E. F. Fama. Random walks in stock market prices. *Financial analysts journal*, 51(1) :75–80, 1995.

E. F. Fama and K. R. French. The cross-section of expected stock returns. *the Journal of Finance*, 47(2) :427–465, 1992.

Y. Fang, K. K. Lai, and S.-Y. Wang. Portfolio rebalancing model with transaction costs based on fuzzy decision theory. *European Journal of Operational Research*, 175(2) :879–893, 2006.

O. Fathi. Time series forecasting using a hybrid arima and lstm model. In *Velvet Consulting*. 2019.

F. Fei, A.-M. Fuertes, and E. Kalotychou. Dependence in credit default swap and equity markets : Dynamic copula with markov-switching. *International Journal of Forecasting*, 33(3) :662–678, 2017.

A. Fertis, M. Baes, and H.-J. Lüthi. Robust risk management. *European Journal of Operational Research*, 222(3) : 663–672, 2012.

H. Fink, Y. Klimova, C. Czado, and J. Stöber. Regime switching vine copula models for global equity and volatility indices. *Econometrics*, 5(1) :3, 2017.

T. Fischer and C. Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2) :654–669, 2018.

P. C. Fishburn. Mean-risk analysis with risk associated with below-target returns. *The American Economic Review*, 67(2) :116–126, 1977.

C. J. Fombrun, N. A. Gardberg, and M. L. Barnett. Opportunity platforms and safety nets : Corporate citizenship and reputational risk. *Business and society review*, 105(1), 2000.

T. Foo and S. Eng. Asset allocation in a downside risk framework. *Journal of Real Estate Portfolio Management*, 6(3) :213–223, 2000.

R. Freeman. Strategic management : A stakeholder approach pitman publishing inc. *Marshfield, MA*, 1984.

E. W. Frees and E. A. Valdez. Understanding relationships using copulas. *North American actuarial journal*, 2 (1) :1–25, 1998.

W. R. Freudenburg and R. Gramling. *Blowout in the Gulf : The BP oil spill disaster and the future of energy in America*. MIT Press, 2011.

R. Frey and A. J. McNeil. Modelling dependent defaults. Technical report, ETH Zurich, 2001.

M. Friedman. The social responsibility of business is to increase its profits. In *Corporate ethics and corporate governance*, pages 173–178. Springer, 2007.

Y. Funabashi and K. Kitazawa. Fukushima in review : A complex disaster, a disastrous response. *Bulletin of the Atomic Scientists*, 68(2) :9–21, 2012.

J. Gao, K. Zhou, D. Li, and X. Cao. Dynamic mean-lpm and mean-cvar portfolio optimization in continuous-time. *SIAM Journal on Control and Optimization*, 55(3) :1377–1397, 2017.

F. García, J. González-Bueno, J. Oliver, and N. Riley. Selecting socially responsible portfolios : A fuzzy multi-criteria approach. *Sustainability*, 11(9) :2496, 2019.

A. Garel and A. Petit-Romec. Investor rewards to environmental responsibility : Evidence from the covid-19 crisis. *Journal of Corporate Finance*, 68 :101948, 2021.

L. Garlappi, R. Uppal, and T. Wang. Portfolio selection with parameter and model uncertainty : A multi-prior approach. *The Review of Financial Studies*, 20(1) :41–81, 2007.

S. M. Gasser, M. Rammerstorfer, and K. Weinmayer. Markowitz revisited : Social portfolio engineering. *European Journal of Operational Research*, 258(3) :1181–1190, 2017.

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas : A review and a power study. *Insurance : Mathematics and economics*, 44(2) :199–213, 2009.

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5) :1779–1801, 1993.

P. C. Godfrey, C. B. Merrill, and J. M. Hansen. The relationship between corporate social responsibility and shareholder value : An empirical test of the risk management hypothesis. *Strategic management journal*, 30 (4) :425–445, 2009.

P. Gompers, J. Ishii, and A. Metrick. Corporate governance and equity prices. *The quarterly journal of economics*, 118(1) :107–156, 2003.

J.-P. Gond, N. Kang, and J. Moon. The government of self-regulation : On the comparative dynamics of corporate social responsibility. *Economy and society*, 40(4) :640–671, 2011.

G. Gordon. Oil, water and law don't mix : Environmental liability for offshore oil and gas operations in the uk : Part i : Liability in the law of tort/delict and under the petroleum licence. *Environmental Law and Management*, 25(1) :3–11, 2013.

P.-O. Gourinchas and A. Tornell. Exchange rate puzzles and distorted beliefs. *Journal of International Economics*, 64(2) :303–333, 2004.

C. Granger. Combining forecasts-twenty years later. *Journal of Forecasting*, 8 :167–173, 1989.

C. W. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica : journal of the Econometric Society*, pages 424–438, 1969.

C. W. Granger. Forecasting stock market prices : Lessons for forecasters. *International Journal of Forecasting*, 8 (1) :3–13, 1992.

C. W. J. Granger. *Modelling economic series : readings in econometric methodology*. Oxford University Press, 1991.

S. F. Gray. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1) :27–62, 1996.

J. Green, J. R. Hand, and X. F. Zhang. The supraview of return predictive signals. *Review of Accounting Studies*, 18(3) :692–730, 2013.

R. C. Green and B. Hollifield. When will mean-variance efficient portfolios be well diversified ? *The Journal of Finance*, 47(5) :1785–1809, 1992.

K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber. Lstm : A search space odyssey. *IEEE transactions on neural networks and learning systems*, 28(10) :2222–2232, 2016.

P. Grundke and S. Polle. Crisis and risk dependencies. *European Journal of Operational Research*, 223(2) :518–528, 2012.

M. Gulliksson and S. Mazur. An iterative approach to ill-conditioned optimal portfolio selection. *Computational Economics*, 56(4) :773–794, 2020.

E. J. Gumbel. Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292) :698–707, 1960.

H. Guo, X. Liu, and Z. Sun. Multivariate time series prediction using a hybridization of varma models and bayesian networks. *Journal of Applied Statistics*, 43(16) :2897–2909, 2016.

A. S. Hadi, A. A. El Naggar, and M. N. Abdel Bary. New model and method for portfolios selection. *Applied Mathematical Sciences*, 10(6) :263–288, 2016.

C. M. Hafner and O. Reznikova. Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics & Data Analysis*, 54(11) :2609–2627, 2010.

G. Halbritter and G. Dorfleitner. The wages of social responsibility—where are they ? a critical review of esg investing. *Review of Financial Economics*, 26 :25–35, 2015.

W. Hallerbach, H. Ning, A. Soppe, and J. Spronk. A framework for managing a portfolio of socially responsible investments. *European Journal of Operational Research*, 153(2) :517–529, 2004.

J. D. Hamilton and R. Susmel. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics*, 64(1-2) :307–333, 1994.

S. M. Hartzmark and A. B. Sussman. Do investors value sustainability ? a natural experiment examining ranking and fund flows. *The Journal of Finance*, 74(6) :2789–2837, 2019.

C. R. Harvey, J. C. Liechty, M. W. Liechty, and P. Müller. Portfolio selection with higher moments. *Quantitative Finance*, 10(5) :469–485, 2010.

K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

R. Heede and N. Oreskes. Potential emissions of co2 and methane from proved reserves of fossil fuels : An alternative analysis. *Global Environmental Change*, 36 :12–20, 2016.

A. Heinen and A. Valdesogo. Asymmetric capm dependence for large dimensions : The canonical vine autoregressive copula model. *Available at SSRN 1297506*, 2008.

J. A. Hernandez. Vine copula modelling of dependence and portfolio optimization with application to mining and energy stock return series from the australian market. 2015.

J. M. Hernández-Lobato and R. Adams. Probabilistic backpropagation for scalable learning of bayesian neural networks. In *International conference on machine learning*, pages 1861–1869. PMLR, 2015.

M. K. Hess. What drives markov regime-switching behavior of stock markets ? the swiss case. *International Review of Financial Analysis*, 12(5) :527–543, 2003.

P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu. Temporal convolutional neural (tcn) network for an effective weather forecasting using time-series data from the local weather station. *Soft Computing*, 24 :16453–16482, 2020.

K. W. Hipel and A. I. McLeod. *Time series modelling of water resources and environmental systems*. Elsevier, 1994.

M. Hirschberger, R. E. Steuer, S. Utz, M. Wimmer, and Y. Qi. Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research*, 61(1) :169–183, 2013.

B. Hirschl. International renewable energy policy—between marginalization and initial approaches. *Energy Policy*, 37(11) :4407–4416, 2009.

S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8) :1735–1780, 1997.

A. G. Hoepner, I. Oikonomou, Z. Sautner, L. T. Starks, and X. Zhou. Esg shareholder engagement and downside risk. 2018.

M. Holt, R. J. Campbell, and M. B. Nikitin. Fukushima nuclear disaster, 2012.

H. Hong and M. Kacperczyk. The price of sin : The effects of social norms on markets. *Journal of financial economics*, 93(1) :15–36, 2009.

L. Hu. Dependence patterns across financial markets : a mixed copula approach. *Applied financial economics*, 16(10) :717–729, 2006.

W. Huang, K. K. Lai, Y. Nakamori, and S. Wang. Forecasting foreign exchange rates with artificial neural networks : A review. *International Journal of Information Technology & Decision Making*, 3(01) :145–165, 2004.

W. Huang, Y. Nakamori, and S.-Y. Wang. Forecasting stock market movement direction with support vector machine. *Computers & operations research*, 32(10) :2513–2522, 2005.

K. Hussein and M. Omran. Ethical investment revisited : evidence from dow jones islamic indexes. *The journal of Investing*, 14(3) :105–126, 2005.

I. Ioannou and G. Serafeim. The impact of corporate social responsibility on investment recommendations : Analysts' perceptions and shifting institutional logics. *Strategic management journal*, 36(7) :1053–1081, 2015.

R. Jagannathan and T. Ma. Risk reduction in large portfolios : Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4) :1651–1683, 2003.

P. Jana, T. Roy, and S. Mazumder. Multi-objective possibilistic model for portfolio selection with transaction cost. *Journal of computational and applied mathematics*, 228(1) :188–196, 2009.

C. M. Jarque and A. K. Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3) :255–259, 1980.

W. Ji and K. C. Chee. Prediction of hourly solar radiation using a novel hybrid model of arma and tdnn. *Solar energy*, 85(5) :808–817, 2011.

H. Joe. Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. *Lecture Notes-Monograph Series*, pages 120–141, 1996.

H. Joe. *Multivariate models and multivariate dependence concepts*. CRC press, 1997.

E. Jondeau and M. Rockinger. The copula-garch model of conditional dependencies : An international stock market application. *Journal of international money and finance*, 25(5) :827–853, 2006a.

E. Jondeau and M. Rockinger. Optimal portfolio allocation under higher moments. *European Financial Management*, 12(1) :29–55, 2006b.

P. Jorion. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative analysis*, 21(3) : 279–292, 1986.

P. Jorion. Value at risk. 2000.

D. Kahneman and A. Tversky. Prospect theory : An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making : Part I*, pages 99–127. World Scientific, 2013.

I. Kakouris and B. Rustem. Robust portfolio optimization with copulas. *European Journal of Operational Research*, 235(1) :28–37, 2014.

J. S. Kamdem, R. B. Essomba, and J. N. Berinyuy. Deep learning models for forecasting and analyzing the implications of covid-19 spread on some commodities markets volatilities. *Chaos, Solitons & Fractals*, 140 : 110215, 2020.

R. S. Kaplan and K. Ramanna. How to fix esg reporting. *Available at SSRN 3900146*, 2021.

A. N. Kia, M. Fathian, and M. Gholamian. Using mlp and rbf neural networks to improve the prediction of exchange rate time series with arima. *International Journal of Information and Electronics Engineering*, 2(4) : 543–546, 2012.

S.-J. Kim and J. Sheen. Interventions in the yen–dollar spot market : A story of price, volatility and volume. *Journal of Banking & Finance*, 30(11) :3191–3214, 2006.

W. C. Kim, F. J. Fabozzi, P. Cheridito, and C. Fox. Controlling portfolio skewness and kurtosis without directly optimizing third and fourth moments. *Economics Letters*, 122(2) :154–158, 2014.

F. Klaassen. Improving garch volatility forecasts with regime-switching garch. In *Advances in Markov-switching models*, pages 223–254. Springer, 2002.

F. H. Knight. Risk, uncertainty and profit. 1921.

R. Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995.

P. Koizol and M. Kunisch. Analysis of copula functions and applications to credit risk management. In *Financial Management Workshop, Ulm, USA*, 2005.

P. N. Kolm, R. Tütüncü, and F. J. Fabozzi. 60 years of portfolio optimization : Practical challenges and current trends. *European Journal of Operational Research*, 234(2) :356–371, 2014.

S. Kotsantonis and G. Serafeim. Four things no one will tell you about esg data. *Journal of Applied Corporate Finance*, 31(2) :50–58, 2019.

A. Krogh and J. Vedelsby. Neural network ensembles, cross validation, and active learning. *Advances in neural information processing systems*, 7, 1994.

P. Krokhmal, J. Palmquist, and S. Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of risk*, 4 :43–68, 2002.

W. H. Kruskal and W. A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260) :583–621, 1952.

M. Kumar and M. Thenmozhi. Forecasting stock index returns using arima-svm, arima-ann, and arima-random forest hybrid models. *International Journal of Banking, Accounting and Finance*, 5(3) :284–308, 2014.

D. Kurowicka and R. M. Cooke. *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons, 2006.

L. Kurtz. Socially responsible investment and shareholder activism. 2008.

L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters. Noise dressing of financial correlation matrices. *Physical review letters*, 83(7) :1467, 1999.

L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud. Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3(03) :391–397, 2000.

C. G. Lamoureux and W. D. Lastrapes. Persistence in variance, structural change, and the garch model. *Journal of Business & Economic Statistics*, 8(2) :225–234, 1990.

N. Lawrence and A. Hyvärinen. Probabilistic non-linear principal component analysis with gaussian process latent variable models. *Journal of machine learning research*, 6(11), 2005.

J. Le Maux and E. Le Saout. The performance of sustainability indexes. *Finance India*, 18 :737, 2004.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324, 1998.

O. Ledoit and M. Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5) :603–621, 2003.

O. Ledoit and M. Wolf. Nonlinear shrinkage of the covariance matrix for portfolio selection : Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12) :4349–4388, 2017.

D. D. Lee and R. W. Faff. Corporate sustainability performance and idiosyncratic risk : A global perspective. *Financial Review*, 44(2) :213–237, 2009.

T.-H. Lee and X. Long. Copula-based multivariate garch model with uncorrelated dependent errors. *Journal of Econometrics*, 150(2) :207–218, 2009.

I. Leifer, W. J. Lehr, D. Simecek-Beatty, E. Bradley, R. Clark, P. Dennison, Y. Hu, S. Matheson, C. E. Jones, B. Holt, et al. State of the art satellite and airborne marine oil spill remote sensing : Application to the bp deepwater horizon oil spill. *Remote Sensing of Environment*, 124 :185–209, 2012.

H. Levy and H. M. Markowitz. Approximating expected utility by a function of mean and variance. *The American Economic Review*, 69(3) :308–317, 1979.

K. Li, A. Sarkar, and Z. Wang. Diversification benefits of emerging markets subject to portfolio constraints. *Journal of empirical finance*, 10(1-2) :57–80, 2003.

T. Li, M. Hua, and X. Wu. A hybrid cnn-lstm model for forecasting particulate matter (pm2. 5). *Ieee Access*, 8 : 26933–26940, 2020.

A. Lim, J. Shanthikumar, and G.-Y. Vahn. Fragility of cvar in portfolio optimization. 2009.

K. V. Lins, H. Servaes, and A. Tamayo. Social capital, trust, and firm performance : The value of corporate social responsibility during the financial crisis. *the Journal of Finance*, 72(4) :1785–1824, 2017.

T. J. Linsmeier and N. D. Pearson. Risk measurement : An introduction to value at risk. Technical report, 1996.

G. Liu, W. Long, X. Zhang, and Q. Li. Detecting financial data dependence structure by averaging mixture copulas. *Econometric Theory*, 35(4) :777–815, 2019.

I. E. Livieris, E. Pintelas, and P. Pintelas. A cnn–lstm model for gold price time-series forecasting. *Neural computing and applications*, 32(23) :17351–17360, 2020.

A. W. Lo and A. C. MacKinlay. Stock market prices do not follow random walks : Evidence from a simple specification test. *The review of financial studies*, 1(1) :41–66, 1988.

A. W. Lo, H. Mamaysky, and J. Wang. Foundations of technical analysis : Computational algorithms, statistical inference, and empirical implementation. *The journal of finance*, 55(4) :1705–1765, 2000.

A. W. Lo, C. Petrov, and M. Wierzbicki. It's 11 pm—do you know where your liquidity is ? : The mean–variance–liquidity frontier. In *The world of risk management*, pages 47–92. World Scientific, 2006.

J. Longerstaey and M. Spencer. Riskmetricstm—technical document. *Morgan Guaranty Trust Company of New York : New York*, 51 :54, 1996.

F. Longin and B. Solnik. Is the correlation in international equity returns constant : 1960–1990 ? *Journal of international money and finance*, 14(1) :3–26, 1995.

F. Longin and B. Solnik. Extreme correlation of international equity markets. *The journal of finance*, 56(2) : 649–676, 2001.

F. M. Longin. The asymptotic distribution of extreme stock market returns. *Journal of business*, pages 383–408, 1996.

R. K. Y. Low. Vine copulas : modelling systemic risk and enhancing higher-moment portfolio optimisation. *Accounting & Finance*, 58 :423–463, 2018.

W. Lu, J. Li, Y. Li, A. Sun, and J. Wang. A cnn-lstm-based model to forecast stock prices. *Complexity*, 2020, 2020.

F. Luan, W. Zhang, and Y. Liu. Robust international portfolio optimization with worst-case mean-lpm. *Mathematical Problems in Engineering*, 2022, 2022.

E. Lundström and C. Svensson. Including esg concerns in the portfolio selection process : An mcdm approach, 2014.

D. J. MacKay et al. Introduction to gaussian processes. *NATO ASI series F computer and systems sciences*, 168 : 133–166, 1998.

D. MacKenzie and T. Spears. 'the formula that killed wall street' : The gaussian copula and modelling practices in investment banking. *Social Studies of Science*, 44(3) :393–417, 2014.

B. G. Malkiel. *A random walk down Wall Street : including a life-cycle guide to personal investing*. WW Norton & Company, 1999.

C. A. Mallin, B. Saadouni, and R. J. Briston. The financial performance of ethical investment funds. *Journal of Business Finance and Accounting*, 22 :483–483, 1995.

B. Mandelbrot. New methods in statistical economics. *Journal of political economy*, 71(5) :421–440, 1963.

B. Mandelbrot et al. The variation of certain speculative prices. *The Journal of Business*, 36 :394–394, 1963.

B. B. Mandelbrot. The variation of certain speculative prices. In *Fractals and scaling in finance*, pages 371–418. Springer, 1997.

H. Markowitz. Harry m. markowitz. *Portfolio selection, Journal of Finance*, 7(1) :77–91, 1952a.

H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1) :77–91, 1952b.

H. Markowitz. The utility of wealth. *Journal of political Economy*, 60(2) :151–158, 1952c.

H. M. Markowitz. Portfolio selection : Efficient diversification of investments. *Cowles Foundation Monograph*, 16, 1959.

H. M. Markowitz. Portfolio selection. efficient diversification of investments. oxford, 1991.

H. M. Markowitz. The general mean-variance portfolio selection problem. *Philosophical Transactions of the Royal Society of London. Series A : Physical and Engineering Sciences*, 347(1684) :543–549, 1994.

L. Martellini. Towards the design of better equity benchmarks. Technical report, EDHEC Working Paper, 2009.

C. McKenzie and K. Hipel. The gamma arima (garima) model for irregularly observed time series. *Journal of Time Series Analysis*, 15(5) :541–553, 1994.

A. J. McNeil, R. Frey, and P. Embrechts. Quantitative risk management : Concepts, techniques and tools. 2005.

M. Meinshausen, N. Meinshausen, W. Hare, S. C. Raper, K. Frieler, R. Knutti, D. J. Frame, and M. R. Allen. Greenhouse-gas emission targets for limiting global warming to 2 c. *Nature*, 458(7242) :1158–1162, 2009.

H. Mejdoub and B. Mounira. A multivariate analysis for risk capital estimation in insurance industry : Vine copulas. *Asian Development Policy Review*, 5(2) :100–119, 2017.

M. R. Melchiori. Which archimedean copula is the right one ? *YieldCurve, October*, 2003.

E. Michalena and J. M. Hills. Renewable energy issues and implementation of european energy policy : The missing generation ? *Energy policy*, 45 :201–216, 2012.

R. O. Michaud. The markowitz optimization enigma : Is 'optimized'optimal ? *Financial analysts journal*, 45(1) : 31–42, 1989.

A. Min and C. Czado. Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics*, 8(4) :511–546, 2010.

A. Min and C. Czado. Bayesian model selection for d-vine pair-copula constructions. *Canadian Journal of Statistics*, 39(2) :239–258, 2011.

S. Mittnik, F. J. Fabozzi, S. M. Focardi, S. T. Rachev, and T. Jašić. *Financial econometrics : from basics to advanced modeling techniques*. John Wiley & Sons, 2007.

R. A. Moghadam and M. Keshmirpour. Hybrid arima and neural network model for measurement estimation in energy-efficient wireless sensor networks. In *International Conference on Informatics Engineering and Information Science*, pages 35–48. Springer, 2011.

O. Morales-Napoles. Counting vines. In *Dependence modeling : Vine copula handbook*, pages 189–218. World Scientific, 2010.

S. Mukherjee, E. Osuna, and F. Girosi. Nonlinear prediction of chaotic time series using support vector machines. In *Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Signal Processing Society Workshop*, pages 511–520. IEEE, 1997.

K.-R. Müller, A. J. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik. Predicting time series with support vector machines. In *International Conference on Artificial Neural Networks*, pages 999–1004. Springer, 1997.

J. W. Muteba Mwamba and S. M. Mwambi. Assessing market risk in brics and oil markets : An application of markov switching and vine copula. *International Journal of Financial Studies*, 9(2) :30, 2021.

J. E. Nash and J. V. Sutcliffe. River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology*, 10(3) :282–290, 1970.

A. Navon and Y. Keller. Financial time series prediction using deep learning. *arXiv preprint arXiv :1711.04174*, 2017.

W. I. A. W. M. Nawi, M. S. Lola, R. Zakariya, N. H. Zainuddin, A. A. K. Abd Hamid, E. Aruchunan, and N. S. A. Nazzrol. Improved of forecasting sea surface temperature based on hybrid arima and support vector machines models. *Malaysian Journal of Fundamental and Applied Sciences*, 17(5) :609–620, 2021.

D. N. Nawrocki. Optimal algorithms and lower partial moment : ex post results. *Applied Economics*, 23(3) : 465–470, 1991.

D. N. Nawrocki. The characteristics of portfolios selected by n-degree lower partial moment. *International Review of Financial Analysis*, 1(3) :195–209, 1992.

D. N. Nawrocki. A brief history of downside risk measures. *The Journal of Investing*, 8(3) :9–25, 1999.

D. B. Nelson. Conditional heteroskedasticity in asset returns : A new approach. *Econometrica : Journal of the Econometric Society*, pages 347–370, 1991.

D. M. Nelson, A. C. Pereira, and R. A. de Oliveira. Stock market's price movement prediction with lstm neural networks. In *2017 International joint conference on neural networks (IJCNN)*, pages 1419–1426. IEEE, 2017.

C. C. Nguyen and M. I. Bhatti. Copula model dependency between oil prices and stock markets : Evidence from china and vietnam. *Journal of International Financial Markets, Institutions and Money*, 22(4) :758–773, 2012.

W. Ogryczak and A. Ruszczynski. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13(1) :60–78, 2002.

T. Okimoto. New evidence of asymmetric dependence structures in international equity markets. *Journal of financial and quantitative analysis*, 43(3) :787–815, 2008.

C. Ordóñez, F. S. Lasheras, J. Roca-Pardiñas, and F. J. de Cos Juez. A hybrid arima–svm model for the study of the remaining useful life of aircraft engines. *Journal of Computational and Applied Mathematics*, 346 :184–191, 2019.

B. N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio. N-beats : Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv :1905.10437*, 2019.

S. Pafka, M. Potters, and I. Kondor. Exponential weighting and random-matrix-theory-based filtering of financial covariance matrices for portfolio optimization. *arXiv preprint cond-mat/0402573*, 2004.

P.-F. Pai and C.-S. Lin. A hybrid arima and support vector machines model in stock price forecasting. *Omega*, 33(6) :497–505, 2005.

F. Pan, H. Zhang, and M. Xia. A hybrid time-series forecasting model using extreme learning machines. In *2009 Second International Conference on Intelligent Computation Technology and Automation*, volume 1, pages 933–936. IEEE, 2009.

A. Panibratov, J. Fedoritenko, D. Dubova, and D. Seleznev. Impact of the eu carbon policy on the globalization and esg agenda. *BRICS Journal of Economics*, 3(1) :53–71, 2022.

L. Pástor, R. F. Stambaugh, and L. A. Taylor. Dissecting green returns. *Journal of Financial Economics*, 146(2) : 403–424, 2022.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International economic review*, 47(2) :527–556, 2006.

L. H. Pedersen, S. Fitzgibbons, and L. Pomorski. Responsible investing : The esg-efficient frontier. *Journal of Financial Economics*, 142(2) :572–597, 2021.

D. Pelletier. Regime switching for dynamic correlations. *Journal of econometrics*, 131(1-2) :445–473, 2006.

M. H. Pesaran. *Time series and panel data econometrics*. Oxford University Press, 2015.

G. C. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization*, pages 272–281. Springer, 2000.

V. Plerou, P. Gopikrishnan, B. Rosenow, L. N. Amaral, and H. E. Stanley. A random matrix theory approach to financial cross-correlations. *Physica A : Statistical Mechanics and its Applications*, 287(3-4) :374–382, 2000.

G. G. Polak, D. F. Rogers, and D. J. Sweeney. Risk management strategies via minimax portfolio optimization. *European Journal of Operational Research*, 207(1) :409–419, 2010.

M. E. Porter. Towards a dynamic theory of strategy. *Strategic management journal*, 12(S2) :95–117, 1991.

Y. Qi, R. E. Steuer, and M. Wimmer. An analytical derivation of the efficient surface in portfolio selection with three criteria. *Annals of Operations Research*, 251(1) :161–177, 2017.

A. G. Quaranta and A. Zaffaroni. Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance*, 32(10) :2046–2056, 2008.

V. Rajalakshmi et al. Hybrid time-series forecasting models for traffic flow prediction. *Promet-Traffic&Transportation*, 34(4) :537–549, 2022.

C. E. Rasmussen, C. K. Williams, et al. *Gaussian processes for machine learning*, volume 1. Springer, 2006.

L. Renneboog, J. Ter Horst, and C. Zhang. Socially responsible investments : Institutional aspects, performance, and investor behavior. *Journal of banking & finance*, 32(9) :1723–1742, 2008.

P. M. Rikhotso and B. D. Simo-Kengne. Dependence structures between sovereign credit default swaps and global risk factors in brics countries. *Journal of Risk and Financial Management*, 15(3) :109, 2022.

A. V. Rocha and F. Cribari-Neto. Beta autoregressive moving average models. *Test*, 18(3) :529, 2009.

R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7) :1443–1471, 2002.

R. T. Rockafellar, S. Uryasev, et al. Optimization of conditional value-at-risk. *Journal of risk*, 2 :21–42, 2000.

M. Roondiwala, H. Patel, and S. Varma. Predicting stock prices using lstm. *International Journal of Science and Research (IJSR)*, 6(4) :1754–1756, 2017.

F. Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.

F. Rosenblatt. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. Technical report, Cornell Aeronautical Lab Inc Buffalo NY, 1961.

A. Rudd. Social responsibility and portfolio performance. *California Management Review*, 23(4) :55–61, 1981.

D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *nature*, 323(6088) :533–536, 1986.

J. Sadefo Kamdem et al. Downside risk and kappa index of non-gaussian portfolio with lpm. Technical report, 2011.

M. Salahi, F. Mehrdoust, and F. Piri. Cvar robust mean-cvar portfolio optimization. *International Scholarly Research Notices*, 2013, 2013.

F. Salmon. The formula that killed wall street. *Significance*, 9(1) :16–20, 2012.

D. A. Sauer. The impact of social-responsibility screens on investment performance : Evidence from the domini 400 social index and domini equity mutual fund. *Review of Financial Economics*, 6(2) :137–149, 1997.

B. Schandl, K. Klamroth, and M. M. Wiecek. Norm-based approximation in bicriteria programming. *Computational Optimization and Applications*, 20(1) :23–42, 2001.

D. Schirmacher and E. Schirmacher. Multivariate dependence modeling using pair-copulas. Technical report, Citeseer, 2008.

A. B. Schmidt. Optimal esg portfolios : an example for the dow jones index. *Journal of Sustainable Finance & Investment*, pages 1–7, 2020.

A. B. Schmidt. Optimal esg portfolios : an example for the dow jones index. *Journal of Sustainable Finance & Investment*, 12(2) :529–535, 2022.

P. Schönbucher. Taken to the limit : simple and not-so-simple loan loss distributions. *The Best of Wilmott*, 1 : 143–160, 2002.

S. Schueth. Socially responsible investing in the united states. *Journal of business ethics*, 43(3) :189–194, 2003.

G. Schwarz. Estimating the dimension of a model. *The annals of statistics*, pages 461–464, 1978.

R. Sen, H.-F. Yu, and I. S. Dhillon. Think globally, act locally : A deep neural network approach to high-dimensional time series forecasting. *Advances in neural information processing systems*, 32, 2019.

G. Serafeim and A. Yoon. Which corporate esg news does the market react to ? *Financial Analysts Journal*, pages 1–20, 2022.

H. Servaes and A. Tamayo. The impact of corporate social responsibility on firm value : The role of customer awareness. *Management science*, 59(5) :1045–1061, 2013.

S. P. Sethi. Dimensions of corporate social performance : An analytical framework. *California management review*, 17(3) :58–64, 1975.

W. F. Sharpe. Capital asset prices : A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3) :425–442, 1964.

H. Shefrin. *Beyond greed and fear : Understanding behavioral finance and the psychology of investing*. Oxford University Press on Demand, 2002.

R. J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends ?, 1980.

A. Silvennoinen and T. Teräsvirta. Multivariate garch models. In *Handbook of financial time series*, pages 201–229. Springer, 2009.

P. Simon. *Too big to ignore : the business case for big data*, volume 72. John Wiley & Sons, 2013.

M. Sklar. Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8 :229–231, 1959.

A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3) :199–222, 2004.

M. K. So and C. Y. Yeung. Vine-copula garch model with dynamic conditional dependence. *Computational Statistics & Data Analysis*, 76 :655–671, 2014.

R. Sparkes. *Socially responsible investment : A global revolution*. John Wiley & Sons, 2003.

R. K. Srivastava, K. Greff, and J. Schmidhuber. Training very deep networks. *Advances in neural information processing systems*, 28, 2015.

M. Statman. *Finance for normal people : how investors and markets behave*. Oxford University Press, 2017.

R. E. Steuer, M. Wimmer, and M. Hirschberger. Overviewing the transition of markowitz bi-criterion portfolio selection to tri-criterion portfolio selection. *Journal of Business Economics*, 83(1) :61–85, 2013.

J. Stöber and C. Czado. Regime switches in the dependence structure of multidimensional financial data. *Computational Statistics & Data Analysis*, 76 :672–686, 2014.

B. K. Stone. A general class of three-parameter risk measures. *The Journal of Finance*, 28(3) :675–685, 1973.

I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27, 2014.

G. Szegö. Measures of risk. *European Journal of Operational Research*, 163(1) :5–19, 2005.

D. Tasche. Expected shortfall and beyond. *Journal of Banking & Finance*, 26(7) :1519–1533, 2002.

F. E. Tay and L. Cao. Application of support vector machines in financial time series forecasting. *omega*, 29(4) : 309–317, 2001.

A. S. Temur, M. Akgün, and G. Temur. Predicting housing sales in turkey using arima, lstm and hybrid models. 2019.

T. Teräsvirta, C.-F. Lin, and C. W. Granger. Power of the neural network linearity test. *Journal of time series analysis*, 14(2) :209–220, 1993.

R. H. Thaler and L. Ganser. Misbehaving : The making of behavioral economics. 2015.

M. L. Tiku, W.-K. Wong, D. C. Vaughan, and G. Bian. Time series models in non-normal situations : Symmetric innovations. *Journal of Time Series Analysis*, 21(5) :571–596, 2000.

N. Topaloglou, H. Vladimirou, and S. A. Zenios. Cvar models with selective hedging for international asset allocation. *Journal of Banking & Finance*, 26(7) :1535–1561, 2002.

R. R. Trippi and D. DeSieno. Trading equity index futures with a neural network. *Journal of Portfolio Management*, 19 :27–27, 1992.

P. K. Trivedi, D. M. Zimmer, et al. Copula modeling : an introduction for practitioners. *Foundations and Trends® in Econometrics*, 1(1) :1–111, 2007.

Y. K. Tse and A. K. C. Tsui. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3) :351–362, 2002.

I. Usta and Y. M. Kantar. Mean-variance-skewness-entropy measures : A multi-objective approach for portfolio selection. *Entropy*, 13(1) :117–133, 2011.

S. Utz, M. Wimmer, M. Hirschberger, and R. E. Steuer. Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. *European Journal of Operational Research*, 234(2) :491–498, 2014.

S. Utz, M. Wimmer, and R. E. Steuer. Tri-criterion modeling for constructing more-sustainable mutual funds. *European Journal of Operational Research*, 246(1) :331–338, 2015.

V. N. Vapnik. The nature of statistical learning. *Theory*, 1995.

A. Vidal and W. Kristjanpoller. Gold volatility prediction using a cnn-lstm approach. *Expert Systems with Applications*, 157 :113481, 2020.

S. Viviers and N. Eccles. 35 years of socially responsible investing (sri) research-general trends over time. *South African Journal of Business Management*, 43(4) :1–16, 2012.

J. Von Neumann and O. Morgenstern. Theory of games and economic behavior, 2nd rev. 1947.

J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton university press, 2007.

Q. H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica : Journal of the Econometric Society*, pages 307–333, 1989.

A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang. Phoneme recognition using time-delay neural networks. *IEEE transactions on acoustics, speech, and signal processing*, 37(3) :328–339, 1989.

G. T. Walker. On periodicity in series of related terms. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 131(818) :518–532, 1931.

R. Wan, S. Mei, J. Wang, M. Liu, and F. Yang. Multivariate temporal convolutional network : A deep neural networks approach for multivariate time series forecasting. *Electronics*, 8(8) :876, 2019.

M. Welling. Support vector regression. *Department of Computer Science, University of Toronto, Toronto (Kanada)*, 2004.

M. Whitehouse. How a formula ignited market that burned some big investors. *The Wall Street Journal*, 12 (18) :1, 2005.

E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327, 1958.

F. Wilcoxon. Individual comparisons of grouped data by ranking methods. *Journal of economic entomology*, 39 (2) :269–270, 1946.

C. K. Williams. Prediction with gaussian processes : From linear regression to linear prediction and beyond. *Learning in graphical models*, pages 599–621, 1998.

C. K. Williams and C. E. Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

A. Wilson and R. Adams. Gaussian process kernels for pattern discovery and extrapolation. In *International conference on machine learning*, pages 1067–1075. PMLR, 2013.

M. Wimmer. Esg-persistence in socially responsible mutual funds. *J. Mgmt. & Sustainability*, 3 :9, 2013.

D. Würtz, Y. Chalabi, W. Chen, and A. Ellis. *Portfolio optimization with R/Rmetrics*. Rmetrics, 2009.

R. Xiong, E. P. Nichols, and Y. Shen. Deep learning stock volatility with google domestic trends. *arXiv preprint arXiv :1512.04916*, 2015.

D. Xu, Q. Zhang, Y. Ding, and H. Huang. Application of a hybrid arima–svr model based on the spi for the forecast of drought—a case study in henan province, china. *Journal of Applied Meteorology and Climatology*, 59(7) :1239–1259, 2020.

D. Xu, Q. Zhang, Y. Ding, and D. Zhang. Application of a hybrid arima-lstm model based on the spei for drought forecasting. *Environmental Science and Pollution Research*, 29(3) :4128–4144, 2022.

P. Xue-Kun Song. Multivariate dispersion models generated from gaussian copula. *Scandinavian Journal of Statistics*, 27(2) :305–320, 2000.

Z. Yan, Z. Chen, G. Consigli, J. Liu, and M. Jin. A copula-based scenario tree generation algorithm for multi-period portfolio selection problems. *Annals of operations research*, 292(2) :849–881, 2020.

B. Yang, Z. Cai, C. M. Hafner, G. Liu, et al. Time-varying mixture copula models with copula selection. Technical report, University of Kansas, Department of Economics, 2021.

H. Yao, J. Huang, Y. Li, and J. E. Humphrey. A general approach to smooth and convex portfolio optimization using lower partial moments. *Journal of Banking & Finance*, 129 :106167, 2021.

J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson. Understanding neural networks through deep visualization. *arXiv preprint arXiv :1506.06579*, 2015.

B. Yu, H. Yin, and Z. Zhu. Spatio-temporal graph convolutional networks : A deep learning framework for traffic forecasting. *arXiv preprint arXiv :1709.04875*, 2017a.

F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv :1511.07122*, 2015.

F. Yu, V. Koltun, and T. Funkhouser. Dilated residual networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 472–480, 2017b.

W. Yue and Y. Wang. A new fuzzy multi-objective higher order moment portfolio selection model for diversified portfolios. *Physica A : Statistical Mechanics and its Applications*, 465 :124–140, 2017.

G. U. Yule. Vii. on a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636-646) :267–298, 1927.

C. Zhang, H. Wei, X. Zhao, T. Liu, and K. Zhang. A gaussian process regression based hybrid approach for short-term wind speed prediction. *Energy conversion and management*, 126 :1084–1092, 2016.

G. P. Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50 : 159–175, 2003.

T. Zhang and A. Benveniste. On the convergence properties of the gradient method for training multilayer feedforward neural networks. *Neural networks*, 5(2) :159–175, 1992.

W.-G. Zhang, Y.-J. Liu, and W.-J. Xu. A possibilistic mean-semivariance-entropy model for multi-period portfolio selection with transaction costs. *European Journal of Operational Research*, 222(2) :341–349, 2012.

Z. Zhang and T. Zhang. Has the dependence structure of the brics exchange rates changed after the financial crisis ? evidence from r-vine copula model. *International Journal of Emerging Markets*, 2021.

B. Zhu, J. Chevallier, B. Zhu, and J. Chevallier. Carbon price forecasting with a hybrid arima and least squares support vector machines methodology. *Pricing and Forecasting Carbon Markets : Models and Empirical Analyses*, pages 87–107, 2017.

S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations research*, 57(5) :1155–1168, 2009.